

Symmetrie –
Einführung in die harmonische Analysis



Einleitung

Als Leitthema dieser Vorlesung könnte man „Symmetrie“ wählen. Wir beginnen mit einem geometrischen Beispiel, dazu stelle man sich eine Figur F in der Ebene vor. Für gewisse Bewegungen T der Ebene (technisch: Elemente der $SO(2)$) wird dann F in sich überführt werden. Die Gesamtheit dieser T bildet eine Gruppe, die *Symmetriegruppe* von F .

Als Variante kann man auch *Symmetrien in der Zeit* studieren. Genauer: Gegeben sei ein $f : \mathbb{R} \rightarrow \mathbb{R}$, dabei fassen wir die Variable oft als Zeit auf. Eine Zahl p wird *Periode von f* genannt, wenn $f(t + p) = f(t)$ für alle t gilt. Die Menge aller Perioden bildet offensichtlich eine Untergruppe von $(\mathbb{R}, +)$, und im Fall stetiger f ist sie auch abgeschlossen.

Angenommen, f hat eine nichttriviale Periode p . Typische Fragen sind dann:

- Was sind die „einfachsten“ p -periodischen Funktionen?
- Kann man alle p -periodischen Funktionen aus diesen einfachen Vertretern zusammensetzen?

In der Theorie der Fourieranalyse sind diese Fragen der Ausgangspunkt für eine reichhaltige Theorie mit vielfältigen Anwendungen. Wir werden sehen, wie sich beim Lösen von Differentialgleichungsproblemen genau diese Probleme ergeben und die Grundzüge der Theorie kennen lernen.

Es zeigt sich, dass die wichtigen Gruppen im Hintergrund der Fourieranalyse die Gruppen $(\mathbb{Z}, +)$ und $(\mathbb{R}, +)$ sind. Ersetzt man sie durch andere Gruppen, so ergibt sich ein Ansatz, mit dem man sehr allgemeine Fragen zum Thema „Symmetrie“ untersuchen kann. Diese Theorie ist im vorigen Jahrhundert intensiv entwickelt worden, man spricht von *abstrakter harmonischer Analysis*. Dabei zeigt sich, dass man ohne ein Mindestmaß an topologischer Information nicht auskommt¹⁾.

Das Besondere an diesen so genannten *topologischen Gruppen* ist, dass es in vielen Fällen ein Maß gibt, das mit der Gruppenstruktur verträglich ist, das *Haarsche Maß*. Es spielt bei allen Untersuchungen eine wichtige Rolle. Ziel ist es herauszuarbeiten, wie eine Gruppe als Symmetriegruppe auftreten kann und welche Teile der „klassischen“ Theorie ein Analogon haben. Das ist erwartungsgemäß unterschiedlich schwierig, je nachdem, welche topologischen und gruppentheoretischen Eigenschaften man voraussetzt. Wir werden behandeln:

- Endliche kommutative Gruppen
- Endliche Gruppen
- Kompakte Gruppen
- Lokalkompakte abelsche Gruppen

¹⁾Für unsere Zwecke muss man sich allerdings nicht vorher mit Topologie beschäftigt haben, es wird es reichen, sich mit metrischen Räumen auszukennen.

- Lie-Gruppen

Für einige exemplarische Beispiele wird es zu den theoretischen Ergebnissen Anwendungen aus verschiedenen Bereichen der Mathematik geben.

E. Behrends, Oktober 2013.

Bei der Vorbereitung dieser Vorlesung wurde die nachstehende **Literatur** verwendet:

Behrends, Ehrhard: Introduction to Markov Chains.
Vieweg, 2000.

Behrends, Ehrhard: Elementare Stochastik.
Springer Spektrum, 2012.

Chandrasekharan, Komaravolu: Classical Fourier transforms.
Springer, 1980.

Deitmar, Anton: A first Course in Harmonic Analysis.
Springer, 2002.

Duistermark, J. und Kolk, J.: Lie Groups.
Springer, 2000.

Folland, G.: Fourier Analysis.
AMS, 2009.

Heuser, Harro: Gewöhnliche Differentialgleichungen.
Teubner, 1989.

Hewitt, Edwin and Ross, Kenneth: Abstract Harmonic Analysis I und II.
Springer, 1963.

Kosmann-Schwarzbeck, Y.: Groups and Symmetries.
Springer, 2010.

Procesi, C: Lie Groups.
Springer, 2008.

Vretblad, Anders: Fourier Analysis and Its Applications.
Springer, 2000.

Kapitel 1

Klassische Fourieranalyse

In diesem Kapitel leiten wir einige grundlegende Ergebnisse zu Fourierreihen und zur Fouriertransformation her. Sie sind in der Analysis von großer Bedeutung, bei uns haben sie auch die Funktion, die dann folgenden Untersuchungen zu motivieren.

1.1 Warum Fourierreihen?

Die hier vorgestellten Ideen stammen aus dem 18. Jahrhundert. Damals wurden immer neue Probleme mit analytischen Methoden in Angriff genommen. Zum Beispiel wurde versucht, das *Verhalten einer schwingenden Saite* zu beschreiben. Der Ansatz ist wie folgt:

Wir stellen uns die Saite auf der x -Achse eingespannt vor, zwischen 0 und π . Zu irgendeinem Zeitpunkt t bezeichne mit $u(x, t)$ die Auslenkung an der Stelle x zur Zeit t , wobei $0 \leq x \leq \pi$.

Wir setzen voraus, dass $u_0 = u(\cdot, 0)$ (Auslenkung zur Zeit $t = 0$) und $v_0 = \frac{\partial u}{\partial t}(\cdot, 0)$ (Auslenkungsgeschwindigkeit zur Zeit $t = 0$) bekannt sind; es ist klar, dass u_0 und v_0 bei $x = 0$ und $x = \pi$ verschwinden müssen.

Unter Verwendung einfacher physikalischer Ideen („Kraft gleich Masse mal Beschleunigung“, ...), die man auf „sehr kleine“ Stücke der Saite anwendet, gelangt man dann zur folgenden mathematischen Problemstellung, wenn man sich auf eine idealisierte Situation beschränkt (Saite homogen, nur kleine Auslenkungen, Saite schwingt nur in einer Dimension, Konstanten geeignet umskaliert):

Gesucht ist eine Funktion $u : [0, \pi] \times [0, \infty[\rightarrow \mathbb{R}$ mit den folgenden Eigenschaften:

- $u(0, t) = u(\pi, t) = 0$ für alle t .

- $u(x, 0) = u_0(x)$ für alle x .
- $\frac{\partial u}{\partial t}(x, 0) = v_0(x)$ für alle x .
- $\frac{\partial^2 u}{\partial t^2}(x, t) = \frac{\partial^2 u}{\partial x^2}(x, t)$ für alle x, t (Schwingungsgleichung).

Die – am Ende erfolgreiche – Lösungsstrategie sieht so aus:

- Schritt 1: Suche Lösungen der Form $u(x, t) = A(x)B(t)$ (stehende Wellen); kümmere Dich dabei nicht um u_0, v_0 .
Angenommen, so wurden u_1, u_2, \dots gefunden.
- Schritt 2: Beachte, dass aus Linearitätsgründen auch $a_1 u_1 + a_2 u_2 + \dots$ für reelle a_n Lösung ist (bei der u_0, v_0 immer noch nicht berücksichtigt sind.)
- Schritt 3: Wähle die a_n so geschickt, dass auch noch die Bedingungen mit u_0, v_0 erfüllt sind.

Schritt 1: Die entscheidende Beobachtung ist die folgende: Wenn $A(x)B(t)$ Lösung der Schwingungsgleichung ist, muss notwendig

$$\frac{A''}{A}(x) = \frac{B''}{B}(t) \text{ für alle } x, t$$

gelten, und beide Seiten der Gleichung müssen gleich einer Konstante λ sein.

Nichttriviale Lösungen, die auch noch der Bedingung $A(0) = A(\pi) = 0$ genügen, entstehen nur für $\lambda = \lambda_n = -n^2$ für $n \in \mathbb{N}$, und das führt zu

$$A_n(x) = \sin(nx), \quad B_n(t) = \alpha_n \cos(nt) + \beta_n \sin(nt);$$

dabei sind $\alpha_n, \beta_n \in \mathbb{R}$.

Schritt 2: Wir machen nun den allgemeinen Ansatz

$$u(x, t) = \sum_{n=1}^{\infty} \sin(nx) (\alpha_n \cos(nt) + \beta_n \sin(nt)).$$

Dabei sind die α_n, β_n so geschickt zu wählen, dass gilt:

- Die Reihenkonvergenz ist für unsere Zwecke genügend gutartig; es soll zum Beispiel Summation und Differentiation vertauschbar sein.
- Die u_0 -Bedingung ist erfüllt, d.h. $u_0(x) = \sum_n \alpha_n \sin(nx)$ für alle x .
- Die v_0 -Bedingung ist erfüllt, d.h. $v_0(x) = \sum_n n \beta_n \sin(nx)$ für alle x .

Damit sind wir auf ein bei solchen Ansätzen *typisches Problem* gestoßen. Etwas allgemeiner kann es so formuliert werden:

Gegeben sei eine Funktion $f : [-\pi, \pi] \rightarrow \mathbb{R}$ mit $f(-\pi) = f(\pi)$.
Stelle f als Reihe über die Funktionen $1, \cos(x), \cos(2x), \dots, \sin(x), \sin(2x), \dots$ dar.

Das ist der Ausgangspunkt der Fourieranalyse.

1.2 Fourierreihen: einige grundlegende Ergebnisse

Die einzige Chance!

Gegeben sei ein 2π -periodisches $f : \mathbb{R} \rightarrow \mathbb{R}$. Wir werden im Folgenden annehmen, dass es auf $[-\pi, \pi]$ definiert ist (jedes Intervall der Länge 2π erfüllt den gleichen Zweck.) Mal angenommen, es wäre möglich, f wie gewünscht zu schreiben:

$$f(x) = a_0 + \sum_{n=1}^{\infty} (a_n \cos(nx) + b_n \sin(nx)).$$

Bemerkenswerter Weise ist es dann möglich, die a_n, b_n zu rekonstruieren:

Lemma 1.2.1. *Notwendig gilt dann*

- $a_0 = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) dx.$
- $a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos(nx) dx.$
- $b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin(nx) dx.$

Definitionen

Das vorstehende Lemma legt folgende Definition nahe:

Definition 1.2.2. $f : [-\pi, \pi] \rightarrow \mathbb{R}$ sei *stückweise stetig*. Unter den Fourierkoeffizienten von f verstehen wir die Zahlen

- $a_0 := \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) dx.$
- $a_n := \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos(nx) dx.$
- $b_n := \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin(nx) dx.$

Und wir sagen, dass f bei x in eine Fourierreihe entwickelbar ist, wenn

$$f(x) = a_0 + \sum_{n=1}^{\infty} (a_n \cos(nx) + b_n \sin(nx))$$

gilt. Mit P_n bezeichnen wir die n -te Partialsumme dieser Reihe (P_n ist also eine Funktion).

Doch Achtung: Die a_n, b_n sind zwar immer definiert, es ist aber völlig offen, ob die zugehörige Reihe konvergiert. Schlimmer: Selbst wenn sie konvergiert, muss sie nicht gegen $f(x)$ konvergieren.

Wir wollen noch eine *Fairnessbedingung* motivieren. Die Fourierkoeffizienten sind doch durch Integrale definiert. Folglich werden sich die gleichen Koeffizienten ergeben, wenn wir f an einer Stelle abändern: Der Wert an den Sprungstellen hat keinen Einfluss. Deswegen ist auch nicht zu erwarten, dass f entwickelbar ist, wenn wir an den Sprungstellen beliebige Werte zulassen. Die Lösung:

Definition 1.2.3. f sei stückweise stetig. Unter f_μ verstehen wir die punktweise durch

$$f_\mu(x) := \frac{\lim_{h \rightarrow 0^+} f(x+h) + f(x-h)}{2}$$

definierte Funktion. f_μ ist wohldefiniert und wieder 2π -periodisch, stimmt an den Stetigkeitspunkten von f mit f überein und hat an den Sprungstellen als Wert den Mittelwert zwischen rechts- und linksseitigem Limes.

f heie normalisiert, wenn $f = f_\mu$ gilt.

Ergebnisse

Es wird wichtig sein, Partialsummen der Fourierreihe als Faltungsintegral darzustellen.

Definition 1.2.4. Für $n \in \mathbb{N}$ verstehen wir unter dem Dirichlet-Kern die Funktion

$$D_n(x) := \frac{1}{2\pi} (1 + 2 \cos(x) + 2 \cos(2x) + \cdots + 2 \cos(nx)).$$

Die Faltung $f * D_n$ ist für ein stückweise stetiges $f : [-\pi, \pi]$ punktweise durch

$$(f * D_n)(x) := \int_{-\pi}^{\pi} f(t) D_n(x-t) dt$$

erklärt.

Lemma 1.2.5. (i) $f * D_n$ hat bei $-\pi$ und π den gleichen Wert, kann also wieder als 2π -periodische Funktion aufgefasst werden.

(ii) $f * D_n = P_n$ (vgl. 1.2.2).

(iii) Wenn man sich f 2π -periodisch auf \mathbb{R} fortgesetzt denkt, gilt

$$(f * D_n)(x) = \int_{-\pi}^{\pi} f(x-t) D_n(t) dt$$

für alle x .

(iv) $\int_{-\pi}^{\pi} D_n(t) dt = 1$.

(v) Für alle x mit $x \notin 2\pi\mathbb{Z}$ gilt

$$D_n(x) = \frac{\sin((2n+1)x/2)}{2\pi \sin(x/2)}.$$

Satz 1.2.6. (Riemann-Lemma) Es sei $g : [c, d] \rightarrow \mathbb{R}$ stückweise stetig. Dann gilt

$$\int_c^d g(t) \sin(rt) dt \xrightarrow{r \rightarrow \infty} 0.$$

1. Beweis: Glatt approximieren.
2. Beweis: Zunächst für Treppenfunktionen, dann Linearkombinationen und Approximation.

Es folgt ein *erstes Hauptergebnis*. Dabei ist bemerkenswert, dass die Konvergenz der Fourierreihe bei einem x_0 überraschender Weise nur von den lokalen Eigenschaften von f bei x_0 abhängt.

Satz 1.2.7. $f : [-\pi, \pi] \rightarrow \mathbb{R}$ sei 2π -periodisch, stückweise stetig und normiert. Weiter sei $x_0 \in [-\pi, \pi]$. Eine der beiden folgenden Bedingungen sei erfüllt:

(i) f ist bei x_0 nach links und nach rechts Lipschitz-stetig im folgenden Sinn: Es soll ein L so existieren, dass für $t > 0$ stets

$$|f(x_0 + t) - f(x_0^+)| \leq Lt, \quad |f(x_0 - t) - f(x_0^-)| \leq Lt$$

gilt; dabei bezeichnen $f(x_0^+)$ und $f(x_0^-)$ den rechts- und linksseitigen Limes der Funktion f .

(ii) Es gibt ein Intervall $[x_0 - \varepsilon, x_0 + \varepsilon]$, auf dem f von beschränkter Variation ist.

Dann gilt $\lim_{n \rightarrow \infty} P_n(x_0) = f(x_0)$, d.h., f ist bei x_0 in seine Fourierreihe entwickelbar.

Bemerkungen: 1. Insbesondere folgt, dass f bei allen x in seine Fourierreihe entwickelbar ist, wenn f stückweise stetig differenzierbar ist.

2. Der Beweis zu (i) zeigt, dass die Konvergenzgeschwindigkeit der Reihe nur von der Lipschitzkonstanten L abhängt. Lipschitzabbildungen – insbesondere also stetig differenzierbare Funktionen – haben also eine gleichmäßig konvergente Fourierreihe.

Beispiele für Fourierreihen-Entwicklungen

0. Für Funktionen des Typs $\cos(mx)$ und $\sin(mx)$ ist die Entwicklung trivial.

1. Sei f die „Rechteckfunktion“. Sie ist 0 bei $x = \pm\pi$ und $x = 0$ und hat den Wert -1 bzw. 1 auf $] -\pi, 0 [$ bzw. $] 0, \pi [$. Es ist dann leicht nachzurechnen, dass gilt:

$$a_n = 0 \text{ für } n = 0, 1, \dots; \quad b_n = 0 \text{ für } n = 2, 4, 6, \dots; \quad b_n = 4/(n\pi) \text{ für } n = 1, 3, 5, \dots$$

Kurz:

$$f(x) = \frac{4}{\pi} (\sin(x) + (1/3) \sin(3x) + (1/5) \sin(5x) + \dots)$$

(punktweise Konvergenz).

Als kleine Anwendung erhalten wir durch Auswerten bei $\pi/2$ die interessante Formel

$$\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} \pm \dots$$

2. Definiere $f(x) := |x|$ auf $[-\pi, \pi]$ und setze periodisch fort (das ist die *Sägezahnfunktion*). Wir erhalten

$$f(x) = \frac{\pi}{2} - \frac{4}{\pi} \left(\cos(x) + \frac{1}{3^2} \cos(3x) + \frac{1}{5^2} \cos(5x) + \dots \right)$$

und – durch Auswertung bei π –

$$\frac{\pi^2}{8} = 1 + \frac{1}{3^2} + \frac{1}{5^2} + \dots$$

3. Definiert man $f(x)$ durch x^2 bzw. $(x + 2\pi)^2$ bzw. $2\pi^2$ für x in $]0, \pi]$ bzw. $[-\pi, 0[$ bzw. $\{0\}$, bestimmt dann die Fourierreiheentwicklung und setzt danach speziell $x = 0$ ein, so erhält man

$$\frac{\pi^2}{6} = 1 + \frac{1}{2^2} + \frac{1}{3^2} + \dots$$

Durch Multiplikation ergibt sich noch eine Formel für die inversen geradzahigen Quadrate:

$$\frac{\pi^2}{24} = \frac{1}{2^2} + \frac{1}{4^2} + \frac{1}{6^2} + \dots$$

Anwendungen

Erste – eher zahlentheoretische – Anwendungen haben wir schon kennen gelernt.

1. Wir kommen zum Ausgangsproblem, der Schwingungsgleichung zurück. Mit der im vorigen Abschnitt beschriebenen Strategie wollen wir das folgende Problem lösen:

Finde $u : [0, \pi] \times \mathbb{R} \rightarrow \mathbb{R}$ mit

- $u(0, t) = u(\pi, t) = 0$ für alle t .
- $u(x, 0) = \chi_{]0, \pi[}(x)$ für alle x .
- $(\partial u / \partial t)(x, 0) = 0$ für alle x .
- $\partial^2 u / \partial t^2 = \partial^2 u / \partial x^2$.

Zunächst setzen wir die auftretenden Funktionen schiefssymmetrisch auf $[-\pi, \pi]$ fort. Unter Verwendung der im vorstehenden Unterabschnitt zu „1.“ bewiesenen Ergebnisse liefert die Durchführung der Strategie, dass

$$u(x, t) = \frac{4}{\pi} \sum_{n=1}^{\infty} \frac{1}{(2n+1)} \sin((2n+1)x) \cos((2n+1)t)$$

die gewünschten Eigenschaften hat.

2. Wir behandeln ein Wärmeleitungsproblem auf analoge Weise. Gesucht ist eine Funktion $u : [0, \pi] \times [0, \infty[\rightarrow \mathbb{R}$ mit

- $u(0, t) = u(\pi, t) = 0$ für alle t .
- $u(x, 0) = x$ für $0 \leq x < \pi$, $u(\pi, 0) = 0$.
- $\partial u / \partial t = 3 \partial^2 u / \partial x^2$.

Dieses Problem tritt auf, wenn man die Temperaturentwicklung in einem Stab der Länge π beschreiben möchte. Dabei bezeichnet $u(x, t)$ die Temperatur bei x zur Zeit t , die Ende des Stabs werden permanent auf Null gekühlt, am Anfang gibt es eine spezielle Temperaturverteilung, und der Faktor 3 ergibt sich aus den Materialkonstanten.

Schritt 1: Spezielle Lösungen suchen. Wir setzen wieder $u(x, t) = A(x)B(t)$ an. Das führt auf $(A''/A)(x) = (B'/(3B))(t)$. Beide Funktionen müssen gleich einer Konstanten λ sein. Wieder kommen aufgrund der Randbedingung nur die $\lambda = \lambda_n = -n^2$ in Frage, die zugehörigen A -Funktionen sind $\sin(nx)$.

Wenn $\lambda = \lambda_n$ ist, muss noch $\lambda_n = B'/(3B)$ gelöst werden. Wir erhalten (bis auf eine Konstante) $B(t) = \exp(-3n^2t)$.

Schritt 2: Linearkombinationen bilden. Wir wissen schon, dass alle Funktionen des Typs

$$u(x, t) = \sum_n a_n \sin(nx) \exp(-3n^2t)$$

folgende Eigenschaften haben: Sie verschwinden am Rand und lösen die Wärmeleitungsgleichung. Die zweite Bedingung läuft darauf hinaus, dass $\sum_n a_n \sin(nx)$ auf $0 \leq x < \pi$ gleich x und bei π gleich 0 sein soll. Das ist durch Auffinden der Fourierreihe der zugehörigen Funktion leicht zu realisieren. Wir erhalten als Endergebnis:

$$u(x, t) = \sum_{n=1}^{\infty} \frac{2(-1)^{n+1}}{n\pi} e^{-3n^2t} \sin(nx).$$

Eine komplexe Formulierung

Für spätere Zwecke wird es sinnvoll sein, sich an die Eulersche Gleichung $e^{ix} = \cos x + i \sin x$ zu erinnern und die vorstehenden Ergebnisse entsprechend umzuformulieren.

Es sei $f : [-\pi, \pi] \rightarrow \mathbb{C}$ eine stückweise stetige Funktion. Für $n \in \mathbb{Z}$ (Achtung: nicht nur $n \in \mathbb{N}$) betrachten wir die Funktionen $x \mapsto e^{inx}$, und f soll als Reihe über geeignete Vielfache dieser Funktionen dargestellt werden.

Angenommen, das geht: $f(x) = \sum_{n \in \mathbb{Z}} a_n e^{inx}$. Wenn man mit e^{imx} multipliziert und integriert, so folgt $a_m = (1/2\pi) \int_{-\pi}^{\pi} f(x) e^{imx} dx =: \hat{f}(m)$: Das sind die *komplexen Fourierkoeffizienten*. Die $\hat{f}(n)$ hängen mit den gewöhnlichen Fourierkoeffizienten von Real- bzw. Imaginärteil von f auf einfache Weise zusammen, es

handelt sich wirklich nur um eine Umschreibung. Insbesondere ist wieder richtig, dass punktweise stetig differenzierbare f punktweise als ihre Fourierreihe dargestellt werden können.

Ergänzungen zur Erweiterung der Allgemeinbildung

1. *Cesàro-Konvergenz*: Es ist eine Standardübungsaufgabe der Analysis zu zeigen, dass aus $a_n \rightarrow a$ auch $(a_1 + \dots + a_n)/n \rightarrow a$ folgt. Die passende Definition dazu: Eine Folge (a_n) heißt Cesàro-konvergent gegen a , wenn $(a_1 + \dots + a_n)/n \rightarrow a$ gilt. Das vorstehende Ergebnis besagt dann, dass aus der Konvergenz die Cesàro-Konvergenz folgt. Die Umkehrung ist offensichtlich nicht richtig. Wichtig wird diese Definition besonders in der Theorie der Fourierreihen. Es gilt nämlich der

Satz von Fejér: $f : [-\pi, \pi]$ sei periodisch und stetig. Dann ist f der gleichmäßige Cesàro-Limes der Partialsummen der Fourierreihe:

$$\frac{P_0 + \dots + P_n}{n+1} \rightarrow f \text{ (gleichmäßig).}$$

Beweisidee: Erstens muss man wissen, was eine *Diracfolge* ist. Das ist eine Folge (K_n) nichtnegativer stetiger Funktionen auf \mathbb{R} , deren Integral 1 ist und die im folgenden Sinn immer stärker bei 0 konzentriert ist: Für jedes positive δ gilt

$$\lim_n \left(\int_{-\infty}^{-\delta} + \int_{\delta}^{\infty} \right) K_n(x) dx = 0.$$

Für so eine Folge gilt zweitens: Ist $f : [a, b] \rightarrow \mathbb{R}$ stetig, so konvergieren die $K_n * f$ gleichmäßig auf $[a, b]$ gegen f . (Beweis z.B. in Behrends, Analysis 2, Satz 7.1.3.)

Und drittens schließlich: Definiert man

$$F_n := \frac{D_0 + D_1 + \dots + D_n}{n+1},$$

so ist (F_n) eine Diracfolge. Die F_n heißen die *Fejér-Kerne*¹⁾. Beachte, dass $F_n * f = (P_0 + \dots + P_n)/(n+1)$. □

2. *Das Gibbs-Phänomen* Angenommen, f hat eine Sprungstelle bei x_0 , etwa bei $x_0 = 0$. Dann ist nicht zu erwarten, dass die P_n gleichmäßig gegen f konvergieren, denn gleichmäßige Limes stetiger Funktionen sind wieder stetig. Es wäre allerdings möglich, dass die Konvergenz auf den Teilintervallen des Definitionsbereichs, auf denen f stetig ist, „gut“ ist.

Als Beispiel betrachten wir die Funktion $f(x) := x$, der Wert bei $\pm\pi$ soll gleich Null sein. Skizziert man die Partialsummen, so sieht man, dass sie die Funktion f auch für große n in der Nähe von 0 immer deutlich über- bzw. untertreffen.

¹⁾Explizit ist übrigens $F_n(x) = [\sin((n+1)x/2) / \sin(x/2)]^2 / (n+1)$.

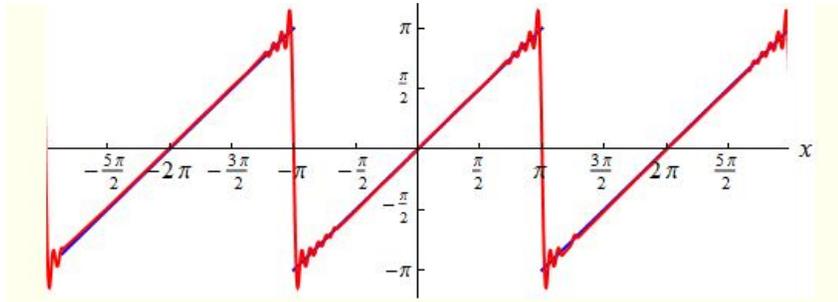


Illustration zum Gibbs-Phänomen.

Das ist typisch: Es ist *immer* so, dass in der Nähe einer Unstetigkeitsstelle x_0 auf den Intervallen $]x_0 - \varepsilon, x_0[$ und $]x_0, x_0 + \varepsilon[$ die gleichmäßige Konvergenz auf quantitative Weise verletzt ist. Die P_n entfernen sich oszillierend in der Nähe von x_0 von $f(x_0^-)$ bzw. $f(x_0^+)$. Sie über- bzw. untertreffen f um etwa 9 Prozent. Das ist das *Gibbs-Phänomen*.

3. L^2 -Konvergenz. Bisher haben wir untersucht, ob die P_n punktweise oder gleichmäßig gegen f konvergent sind. Es gibt aber eine Fülle weiterer Konvergenzbegriffe, hier soll ein Ergebnis zur L^2 -Konvergenz erwähnt werden. Es gilt: Ist $f : [-\pi, \pi] \rightarrow \mathbb{R}$ eine L^2 -Funktion, so konvergieren die P_n in der L^2 -Norm gegen die Funktion f .

Damit das eine sinnvolle Aussage ist, muss zunächst bemerkt werden, dass f auch eine L^1 -Funktion ist: Deswegen sind die Fourierkoeffizienten definiert.

Die Aussage selbst folgt schnell aus der Theorie der Hilberträume, denn die Funktionen $1, \cos(x), \cos(2x), \dots, \sin(x), \sin(2x), \dots$ bilden ein vollständiges Orthogonalsystem in $L^2[-\pi, \pi]$. (Die Orthogonalitätsrelationen wurden schon ausgenutzt. Die Vollständigkeit folgt – z.B. – daraus, dass wir schon gezeigt haben, dass alle C^1 -Funktionen dargestellt werden können, und die bilden einen dichten Unterraum.)

1.3 Historisches

Die Fourieranalyse spielt eine bedeutende Rolle in der historischen Entwicklung der Mathematik. Hier die wichtigsten Punkte:

- Im 18. Jahrhundert zeigte sich bei der Lösung partieller Differentialgleichungen, dass es wichtig sein könnte, eine periodische Funktion in eine Reihe über Sinus- und Cosinusterme zu entwickeln (vgl. Abschnitt 1.1).
- Fourier entwickelte dazu in seiner „Théorie de la chaleur“ („Theorie der Wärme“) einen erfolgversprechenden Ansatz. Allerdings war er der Meinung, dass *alle* Funktionen entwickelbar sind.

Aus heutiger Sicht ist es nicht verwunderlich, dass seine Ergebnisse gegenwärtigen Strenge-Maßstäben nicht standhalten, denn die grundlegen-

den Begriffe, die für die Theorie wichtig sind (Integral, Konvergenz, Funktion) waren noch nicht präzise eingeführt.

- Einige Jahrzehnte später gab es die ersten belastbaren Ergebnisse von Dirichlet. Einige haben wir im vorigen Abschnitt kennen gelernt.
- Es war sehr verwirrend für die Fachwelt, als du Bois-Raymond 1876 bewies, dass es stetige Funktionen f gibt, für die die Fourierreihen-Partialsummen an manchen Stellen nicht gegen f konvergieren.
- Georg Cantor ging von einem Problem aus der Theorie der Fourierreihen aus, als er seine Mengenlehre entwickelte. (Genauer: 1869 bewies er einen Eindeutigkeitssatz, er zeigte, dass aus

$$0 = a_0 + \sum_n (a_n \cos(nx) + b_n \sin(nx))$$

für alle x folgt, dass alle a_n, b_n verschwinden; im Fall gleichmäßiger Konvergenz ist das klar. Kann man das auch zeigen, wenn die Reihe nicht für alle x Null ist? Wie groß kann die „Ausnahmemenge“ sein?)

- Der weitestgehende sinnvolle Ansatz, Fourierreihen zu untersuchen, sind doch sicher L^1 -Funktionen, denn dafür lassen sich die P_n definieren. Konvergieren sie gegen die Funktion? Dazu gibt es zwei wichtige Ergebnisse:
 - Kolmogoroff zeigte durch ein Gegenbeispiel 1926, dass es L^1 -Funktionen gibt, bei denen punktweise Konvergenz an keiner Stelle vorliegt.
 - Das Carleson-Hunt-Theorem (1966) besagt, dass bessere Integrierbarkeitsbedingungen fast sichere Konvergenz erzwingen. Genauer: Ist f so, dass $f \in L^p[-\pi, \pi]$ für ein $p > 1$, so gibt es eine Lebesgue-Nullmenge N in $[-\pi, \pi]$, derart dass für $x \notin N$ die $P_n(x)$ punktweise gegen $f(x)$ konvergieren.

Die wichtigste Person in diesem Zusammenhang ist sicher *Fourier*.



Er hatte ein bemerkenswert interessantes und abwechslungsreiches Leben (was gar nicht so untypisch für die Zeit während und nach der französischen Revolution war). Hier die wichtigsten Stationen:

- Lebensdaten: 1768 bis 1830.
- Mit Napoleon in Ägypten: Archäologische Bestandsaufnahme. (Daher ist sein Name auch unter Ägyptologen immer noch bekannt.)
- Präfekt in Grenoble.
- Direktor des statistischen Büros in Paris.
- Mitglied der Akademie der Wissenschaften.

1.4 Fouriertransformation

Wir wollen nun entsprechende Ergebnisse für Funktionen herleiten, die auf \mathbb{R} definiert sind.

Erste Definitionen und Ergebnisse

Diesmal betrachten wir gleich die komplexe Variante.

Definition 1.4.1. *Es sei $f : \mathbb{R} \rightarrow \mathbb{C}$ integrierbar. Für $\alpha \in \mathbb{R}$ definieren wir*

$$\hat{f}(\alpha) := \int_{-\infty}^{\infty} f(x) e^{i\alpha x} dx,$$

und unter $\mathcal{F}(f) : \mathbb{R} \rightarrow \mathbb{C}$ verstehen wir die Funktion $\alpha \mapsto \frac{1}{\sqrt{2\pi}} \hat{f}(\alpha)$. Sie heißt die Fouriertransformation von f .

Dabei ist die Normalisierung mit $\frac{1}{\sqrt{2\pi}}$ willkürlich. In der Literatur findet man die Definition auch mit den Faktoren 1 oder $\frac{1}{2\pi}$.

Es handelt sich also um eine Abbildung, die einer Funktion eine Funktion zuordnet. Eine ihrer wichtigsten Eigenschaften ist die Tatsache, dass sie – wie die Laplacetransformation – Differentiationsprobleme in algebraische Probleme verwandelt. (Die Laplacetransformation wird in meinem Buch zur Analysis 2 besprochen.)

Zunächst stellen wir einige leicht nachzuprüfende Eigenschaften zusammen:

Lemma 1.4.2. *(i) \mathcal{F} ist eine lineare Abbildung.*

(ii) $\mathcal{F}(f)$ ist stets eine stetige beschränkte Funktion. Genauer gilt $\|\mathcal{F}(f)\|_{\infty} \leq \|f\|_1$. Es folgt, dass \mathcal{F} stetig ist, wenn man den Urbildraum mit der L^1 -Norm und den Bildraum mit der L^{∞} -Norm versieht.

(iii) (Translationen) Sei $h \in \mathbb{R}$. Wir definieren f_h durch $x \mapsto f(x+h)$. Dann ist $\hat{f}_h(\alpha) = (\hat{f}(\alpha)) e^{-ih\alpha}$.

(iv) (Umskalierungen) Sei $r \in \mathbb{R} \setminus \{0\}$. Definiere f_{*r} durch $x \mapsto f(rx)$. Dann gilt

$$\widehat{f_{*r}}(\alpha) = \frac{1}{|r|} \widehat{f}\left(\frac{\alpha}{r}\right).$$

Es folgt: Mit f ist auch $\mathcal{F}(f)$ symmetrisch bzw. schiefssymmetrisch.

(v) Für $f, g \in L^1 \mathbb{R}$ gilt $\int_{\mathbb{R}} \widehat{f}(x)g(x) dx = \int_{\mathbb{R}} f(x)\widehat{g}(x) dx$.

Für große α ist $e^{i\alpha x}$ eine stark oszillierende Funktion. Deswegen ist zu erwarten, dass es auch hier ein Riemann-Lebesgue-Lemma gibt:

Lemma 1.4.3. $\widehat{f}(\alpha) \rightarrow 0$ für $|\alpha| \rightarrow \infty$.

Beweis: Sei g eine stetig differenzierbare Funktion, die außerhalb eines Intervalls $[-R, R]$ verschwindet. Mit partieller Integration folgt die Aussage für \widehat{g} . Beachte noch, dass die Menge dieser g in $L^1(\mathbb{R})$ dicht liegt.

(Alternativer Beweis: Die Aussage ist klar, wenn g charakteristische Funktion eines Intervalls ist, und die Linearkombinationen solcher Funktionen liegen dicht.) \square

Beispiele

1. Ist f die charakteristische Funktion von $[-1, 1]$, so ist $\widehat{f}(\alpha) = 2(\sin \alpha)/\alpha$. Man beachte, dass das keine L^1 -Funktion ist.

2. Sei $f(x) = 1 - |x|$ für $|x| \leq 1$ und sonst gleich Null. Dann ist

$$\widehat{f}(\alpha) = \left(\frac{\sin(\alpha/2)}{\alpha/2}\right)^2.$$

3. Setze $f(x) = e^{-|x|}$. Dann ist $\widehat{f}(\alpha) = 2/(1 + \alpha^2)$.

4. Für $f(x) = e^{-x^2/2}$ ist $\widehat{f}(\alpha) = \sqrt{2\pi}e^{-\alpha^2/2}$. Die Funktion f ist also eine Eigenfunktion der Fouriertransformation.

Fouriertransformation und Differenzierbarkeit

Diese Frage hat zwei Aspekte: Was weiß man, wenn f differenzierbar ist, und was lässt sich über die Differenzierbarkeit von $\mathcal{F}f$ sagen?

Satz 1.4.4. $f \in L^1$ sei so, dass auch $x \mapsto xf(x)$ zum L^1 gehört. Dann ist $\mathcal{F}f$ differenzierbar und es gilt $(\widehat{f})'(\alpha) = \int_{\mathbb{R}} ix f(x)e^{i\alpha x} dx$. Anders ausgedrückt: Definiert man $g(x) := ix f(x)$, so ist $(\widehat{f})'(\alpha) = \widehat{g}(\alpha)$.

Satz 1.4.5. f sei stetig differenzierbar, und f und f' sollen zum L^1 gehören. Dann gilt $\widehat{f}'(\alpha) = -i\alpha \widehat{f}(\alpha)$.

Analoge Ergebnisse lassen sich zeigen, wenn es um höhere Ableitungen geht.

Es gibt eine interessante Folgerung. Angenommen, f ist k -mal stetig differenzierbar und die Funktionen $f, f', \dots, f^{(k)}$ sind integrierbar. Satz 1.4.4 besagt dann, dass die Fouriertransformation von $f^{(k)}$ gleich $(i\alpha)^k$ mal der Fouriertransformation von f ist. Also ist $(i\alpha)^k \hat{f}(\alpha)$ eine beschränkte Funktion, oder anders ausgedrückt: \hat{f} geht mindestens so schnell gegen Null wie $c\alpha^{-k}$. Eine hohe Differenzierbarkeitsordnung von f bewirkt also schnelles Abfallen der Fouriertransformation im Unendlichen.

Eine Anwendungsstrategie

Die Ergebnisse des vorigen Abschnitts begründen die Bedeutung der Fouriertransformation beim Lösen von linearen gewöhnlichen Differentialgleichungen. Hier ein Beispiel:

- Mal angenommen, wir wollen die Differentialgleichung $4y''' - 3y'' + 4y' + y = g$ lösen, wobei g eine genügend glatte L^1 -Funktion ist.
- Wendet man darauf die Fouriertransformation an, so ergibt sich eine Gleichung des Typs „Polynom in α mal $\mathcal{F}(y)$ gleich $\mathcal{F}(g)$ “. Damit ist $\mathcal{F}(y)$ bekannt, und nun muss man „nur“ noch y aus $\mathcal{F}(y)$ rekonstruieren. Dieses Problem greifen wir im nächsten Unterabschnitt auf.

Wir fassen zusammen: **Mit Hilfe der Fouriertransformation kann man Differentialgleichungsprobleme in Probleme der linearen Algebra transformieren.**

Die Inversionsformel

Wie rekonstruiert man f aus $\mathcal{F}(f)$? Im Allgemeinen ist $\mathcal{F}(f)$ nicht integrierbar, deswegen kann es keine Integralformel geben, in der das Standard-Integral auftaucht. Wir werden jedoch sehen, dass für „vernünftige“ f gilt:

$$f(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \hat{f}(\alpha) e^{-i\alpha x} d\alpha.$$

Das ist die *Fourier-Inversionsformel*²⁾.

Die Präzisierung ist etwas komplizierter. Wir fixieren $f \in L^1$, konstruieren \hat{f} und betrachten für jedes $R > 0$ und $x \in \mathbb{R}$ das Integral

$$S_R(x) := \frac{1}{2\pi} \int_{-R}^R \hat{f}(\alpha) e^{-i\alpha x} d\alpha.$$

²⁾Die Formel kann man auch als

$$f(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} (\mathcal{F}f)(\alpha) e^{-i\alpha x} d\alpha$$

schreiben. Damit taucht beide Male (Transformation und Inverse) der gleiche Faktor $1/\sqrt{2\pi}$ auf. Manche Autoren wählen andere Faktoren, das Produkt ist aber immer $1/2\pi$.

Wenn eine Funktion g integrierbar ist, gilt sicher $\lim_{R \rightarrow \infty} \int_{-R}^R g(\alpha) d\alpha = \int_{\mathbb{R}} g(\alpha) d\alpha$, aber der Limes kann auch für nicht-integrierbare Funktionen (wie etwa $1/\alpha$) existieren.

Wir wollen untersuchen, wann die $S_R(x)$ gegen $f(x)$ konvergieren. Falls \hat{f} integrierbar ist, entspricht das der Fourier-Inversionsformel.

Lemma 1.4.6.

$$S_R(x) - f(x) = \frac{2}{\pi} \int_0^\infty \left(\frac{f(x+t) + f(x-t)}{2} - f(x) \right) \frac{\sin(Rt)}{t} dt.$$

Wir wollen Bedingungen dafür finden, dass $S_R(x) - f(x)$ für $R \rightarrow \infty$ gegen Null geht. Die Situation ist sehr ähnlich zu der entsprechenden in Abschnitt 1.2. Wieder beweisen wir zwei Ergebnisse.

Satz 1.4.7. *Wir verwenden die vorstehenden Bezeichnungen und fixieren ein $x \in \mathbb{R}$. Jede der folgenden Bedingungen impliziert, dass $\lim_R S_R(x) = f(x)$ gilt:*

(i) *f ist stückweise stetig und genügt bei x nach links und nach rechts einer Lipschitzbedingung.*

(ii) *f ist stückweise stetig, normalisiert und auf einer Umgebung von x von beschränkter Variation.*

1.5 Verallgemeinerungen

Die bisher behandelten Ergebnisse zur Fourieranalyse sind in viele Richtungen verallgemeinert worden. In diesem Abschnitt gibt es einige Stichpunkte dazu.

Der Schwartz-Raum und Distributionen

Die Fouriertransformation hat zwei große Nachteile.

- Für jede L^1 -Funktion f auf \mathbb{R} kann man die Fouriertransformation definieren, aber $\mathcal{F}f$ muss nicht zum L^1 gehören.
- Für die Anwendungen wäre es interessant, auch Funktionen behandeln zu können, die nicht zum L^1 gehören.

Ein möglicher Ausweg besteht darin, zu Distributionen überzugehen. Dazu verschafft man sich zunächst einen Raum von besonders gutartigen Funktionen, den *Schwartz-Raum* \mathcal{S} . Der besteht aus allen $\phi \in C^\infty \mathbb{R}$, für die alle Ableitungen im Unendlichen im folgenden Sinn schnell gegen Null gehen: Für alle natürlichen Zahlen k, l soll $\lim_{x \rightarrow \infty} x^l \phi^{(k)}(x) = 0$. (Als Beispiel denke man an e^{-x^2} .)

Dann ist die Fouriertransformation eine Bijektion von \mathcal{S} nach \mathcal{S} , und die Inversionsformel gilt für alle diese Funktionen. Sie kann auf eine größere Klasse von Funktionen erweitert werden. Motiviert wird der Ansatz durch Lemma

1.4.2(v), danach ist doch

$$\int_{\mathbb{R}} (\mathcal{F}\phi)(x)\psi(x) dx = \int_{\mathbb{R}} \phi(x)(\mathcal{F}\psi)(x) dx.$$

Um zu verstehen, wie es weitergeht, muss man an die Grundidee der *Distributionentheorie* erinnern. In dieser Theorie werden Funktionen dadurch verallgemeinert, dass sie als Abbildungen auf Räumen „sehr gutartiger“ Funktionen interpretiert werden. Zum Beispiel wird eine stetige Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$ aufgefasst als Abbildung $\phi \mapsto \int f\phi$, wobei ϕ alle C^∞ -Funktionen mit kompaktem Träger durchläuft.

Ist f differenzierbar, so folgt mit partieller Integration schnell, dass $\int f'\phi = -\int f\phi'$, und das motiviert die folgende Definition: Für *beliebige* f soll die Ableitung (im Distributionensinn) die Abbildung $\phi \mapsto -\int f\phi'$ sein. Damit sind alle stetigen Funktionen (sogar beliebig oft) differenzierbar, die Ableitungen werden aber nicht immer als Funktionen zu interpretieren sein³⁾.

Zurück zur Fourieranalyse. Eine *temperierte Distribution* ist eine lineare Abbildung auf \mathcal{S} , die bezüglich einer geeigneten Topologie auf \mathcal{S} stetig ist. Beispiele ergeben sich durch Funktionen f , die „nicht zu schnell“ wachsen: Es soll ein n geben so dass $x \mapsto f(x)/(1 + |x|^n)$ beschränkt bleibt. Dann erzeugt f eine temperierte Distribution durch $\phi \mapsto \int f\phi$.

Und alle temperierten Distributionen haben eine Fouriertransformation: Ist T so eine Distribution, so definiere $\mathcal{F}T$ durch $\phi \mapsto T(\mathcal{F}\phi)$. Lemma 1.4.2.(v) garantiert dann, dass für Funktionen, die eine „klassische“ Fouriertransformation haben, die neue mit der alten Definition verträglich ist.

Durch diesen Zugang ist sichergestellt, dass so gut wie alle in den Anwendungen wichtigen Funktionen eine Fouriertransformation (im Distributionensinn) besitzen. Der Nachteil ist allerdings, dass man oft als Lösung eines Problems eine Distribution erhält. Wenn man eine klassische Funktion braucht, sind noch Regularitätsüberlegungen anzuschließen.

Die L^2 -Fouriertransformation

Wie kann man Eigenschaften von L^1 -Funktionen (wichtig bei der Definition der Fouriertransformation) mit denen von L^2 -Funktionen (Methoden der euklidischen Räume verfügbar) kombinieren? Die Lösung besteht in einer Anwendung des *Plancherel-Theorems*. Es besagt, dass die Abbildung $f \mapsto \mathcal{F}f$ auf \mathcal{S} (versehen mit der L^2 -Norm) eine Isometrie ist. Da L^2 vollständig ist und \mathcal{S} dicht liegt, kann \mathcal{F} durch ein Standard-Fortsetzungsargument auf L^2 fortgesetzt werden. Für $f \in L^2$ ist die *L^2 -Fouriertransformation* definiert als L^2 -Limes der Folge $(\mathcal{F}f_n)$, wobei die (f_n) eine Folge in \mathcal{S} mit Limes f sind.

Sturm-Liouville-Differentialgleichungen

³⁾Hier sind viele Feinheiten übersprungen worden. Besonders aufwändig ist die Präzisierung der Feinheit, dass Distributionen *stetige* Abbildungen auf dem Raum der ϕ sein sollen.

Fourierreihen wurden dadurch motiviert, partielle Differentialgleichungen durch Betrachtung von Funktionen in getrennten Veränderlichen zu lösen. Wichtig waren dann die nichttrivialen Lösungen von $y'' + \lambda y = 0$, $y(\pm\pi) = 0$. Die Theorie der Fourierreihen nutzt dann die Tatsache aus, dass es „viele“ derartige Lösungen gibt.

In der Sturm-Liouville-Theorie wird das wesentlich verallgemeinert. Da geht es um die folgende Randwertaufgabe:

Gegeben seien ein Intervall $[a, b]$, stetige Funktionen $p, q, r : [a, b] \rightarrow \mathbb{R}$ (p soll auch stetig differenzierbar sein und es soll $r(x) > 0$ für alle x gelten) sowie Zahlen $\alpha_1, \alpha_2, \beta_1, \beta_2$ mit $(\alpha_1, \alpha_2) \neq (0, 0) \neq (\beta_1, \beta_2)$.

Man definiert $Ly := (py')' + qy$ für zweimal differenzierbare Funktionen $y : [a, b] \rightarrow \mathbb{R}$, und gesucht sind Lösungen y des Problems

$$\begin{aligned} Ly + \lambda r(x)y &= 0. \\ \alpha_1 y(a) + \alpha_2 y'(a) &= 0. \\ \beta_1 y(b) + \beta_2 y'(b) &= 0. \end{aligned}$$

Insbesondere interessieren diejenigen λ , für die es nichttriviale Lösungen gibt. Solche λ heißen *Eigenwerte*.

Das obige Ausgangsproblem ist natürlich ein einfacher Spezialfall.

Hier die wichtigsten Ansätze und Ergebnisse:

- Auf $C[a, b]$ führen wir ein inneres Produkt durch

$$\langle f, g \rangle := \int_a^b f(x)g(x)r(x) dx$$

ein. Bemerkenswerter Weise führt das auf die Frage, Eigenwerte eines symmetrischen Operators zu finden. *Hier* spielt die merkwürdige Form von L eine wichtige Rolle. Damit kann man eine in der Funktionalanalysis entwickelte Verallgemeinerung des Satzes von der Hauptachsentransformation anwenden (*Spektraltheorie für kompakte Operatoren*).

- Die Eigenwerte λ sind reell und einfach, sie streben gegen unendlich. Eigenfunktionen zu verschiedenen Eigenwerten sind orthogonal.
- Jedes $f \in C[a, b]$ lässt sich in eine Reihe über Eigenfunktionen entwickeln. Das gilt sowohl bzgl. der durch das Skalarprodukt definierten Konvergenz als auch – bei genügend glatten Funktionen – für die gleichmäßige Konvergenz.

Eine ausführliche Darstellung findet man in Heuser, *Gewöhnliche Differentialgleichungen*, Kapitel VI.

Kapitel 2

Harmonische Analysis: Einführung

In der abstrakten harmonischen Analysis wird herausgearbeitet, dass Fourierreihen und Fouriertransformation eigentlich nur Spezialfälle eines viel allgemeineren gruppentheoretischen Ansatzes sind. Dieser Zugang wird in diesem Kapitel durch die Bereitstellung der wichtigsten Definitionen und grundlegenden Ergebnisse vorbereitet.

Im nächsten Kapitel wird die Theorie dann für einige wichtige Klassen von Gruppen weiterentwickelt.

2.1 Topologische Gruppen

Wir werden uns mit Objekten beschäftigen, die einerseits eine Gruppenstruktur tragen und bei denen man gleichzeitig sinnvoll über Konvergenz, Stetigkeit von Funktionen usw. reden kann. Es wäre naheliegend, mit Gruppen zu beginnen, die gleichzeitig eine Topologie tragen.

Man soll das Folgende aber auch verstehen können, wenn man sich nicht mit topologischen Räumen auskennt. Deswegen gibt es den folgenden *Kompromiss*:

Wenn im Folgenden von einem topologischen Raum die Rede ist, darf man sich einen metrischen Raum darunter vorstellen: Die Theorie der metrischen Räume wird als bekannt vorausgesetzt.

Das schränkt die Untersuchungen zwar etwas ein, wirklich wesentliche Klassen von Gruppen bleiben aber nicht unberücksichtigt.

Hier die fundamentale

2.1.1 Definition: Es sei (G, \circ) eine Gruppe, und gleichzeitig sei auf G eine Topologie \mathcal{T} definiert. Wir sprechen von einer *topologischen Gruppe*, wenn die Abbildungen $(x, y) \rightarrow x \circ y$ (von $G \times G$ nach G) und $x \mapsto x^{-1}$ (von G nach G) stetig sind; dabei ist $G \times G$ mit der Produkttopologie zu versehen.

Erste Bemerkungen und Beispiele:

0. Wir verwenden die übliche Konvention: Ist G kommutativ, so bezeichnen wir die innere Komposition mit $+$ und das neutrale Element mit 0 ; für allgemeine G heißt das neutrale Element e .

1. Es ist leicht zu sehen, dass man die beiden Forderungen auch zusammenfassen kann: $(x, y) \mapsto xy^{-1}$ soll stetig sein.

2. Versieht man eine beliebige Gruppe mit der diskreten Topologie/Metrik, so entsteht eine topologische Gruppe. Gruppen, die höchstens abzählbar sind, werden wir immer mit der diskreten Topologie versehen. Das gilt insbesondere für die Gruppe $(\mathbb{Z}, +)$.

3. Untergruppen sind wieder topologische Gruppen.

4. Der \mathbb{R}^n zusammen mit der Addition und der euklidischen Topologie ist eine topologische Gruppe.

5. $(\mathbb{C} \setminus \{0\}, \cdot)$ ist eine topologische Gruppe. Eine wichtige Rolle wird die Untergruppe der z mit $|z| = 1$ spielen.

6. Sei X ein normierter Raum. Dann ist $(X, +)$ in der durch die Norm induzierten Metrik eine topologische Gruppe.

7. Sei $G = GL(n, \mathbb{K})$ die multiplikative Gruppe der invertierbaren $n \times n$ -Matrizen über \mathbb{R} oder \mathbb{C} . G trägt eine natürliche Topologie, wenn wir G als Teilmenge des \mathbb{R}^{n^2} (bzw. des \mathbb{C}^{n^2}) auffassen. G ist eine topologische Gruppe. (Das liegt daran, dass man Matrizenmultiplikation und das Berechnen der Inversen explizit durch die Matrixkomponenten ausdrücken kann.) Allgemeiner lässt sich zeigen, dass die invertierbaren Operatoren auf einem Banachraum bezüglich der Operatornorm eine topologische Gruppe bilden.

Die folgenden Untergruppen spielen eine besonders wichtige Rolle:

- $SL(n, \mathbb{K})$: die Matrizen mit Determinante Eins.
- $O(n, \mathbb{R})$: die reellen orthogonalen Matrizen.
- $SO(n, \mathbb{R})$: dto., aber mit Determinante Eins.
- $U(n, \mathbb{C})$: die komplexen unitären Matrizen.
- $SU(n, \mathbb{C})$: dto, mit Determinante Eins.

8. Üblicherweise macht eine „natürliche“ Topologie auf einer Gruppe G diese zu einer topologischen Gruppe. Es ist aber leicht, Topologien zu finden, bei der die Forderungen aus 2.1.1 verletzt sind. Am Einfachsten ist es, es so einzurichten, dass verschiedene Elemente aus G verschiedene topologische Eigenschaften haben. Fasse etwa $(\mathbb{R}, +)$ auf als disjunkte Vereinigung aus $] -\infty, 0 [$ (natürliche Topologie), $\{0\}$ und $] 0, \infty [$ (natürliche Topologie)¹). Dann gilt $1 + 1/n \rightarrow 1$,

¹Man kann die zugehörige Metrik so definieren: Es ist $d(x, y)$ in der Regel der übliche Abstand. Wenn sich x, y allerdings in zwei verschiedenen der drei Anteile befinden und der übliche Abstand kleiner als 1 ist, soll er gleich 1 sein.

und wenn man -1 addiert, sollte $1/n \rightarrow 0$ herauskommen. Das stimmt aber in diesem metrischen Raum nicht!

Wir werden uns nach und nach davon überzeugen, dass topologische Gruppen „überall irgendwie gleich“ aussehen. Insbesondere sind die wichtigsten Eigenschaften schon in der Nähe des neutralen Elements e abzulesen. Zum Beispiel:

- Kennt man alle Umgebungen von e , so kennt man auch alle Umgebungen für jedes $g \in G$: Man muss die e -Umgebungen nur verschieben.
- $F : G \rightarrow H$ sei ein Gruppenmorphismus zwischen topologischen Gruppen G und H . Dann ist F genau dann stetig, wenn F bei e stetig ist.

2.2 Das Haarsche Maß

Das Ziel

Sei (G, \circ) eine topologische Gruppe. Dann sind in natürlicher Weise die Borelmengen in G definiert: Das sind die Elemente in der σ -Algebra, die von den offenen Mengen erzeugt ist. Wir wollen Borelmengen messen, dabei sollen Mengen, die durch Translation auseinander hervorgehen, das gleiche Maß haben. Maße, die das leisten, heißen *Haarsche Maße* auf G (Details folgen später).

Die richtigen Maße

Sei μ ein Maß auf den Borelmengen von G : Dann heiße μ ein *Borelmaß*. μ soll auch noch verträglich mit der topologischen Struktur auf G sein:

Definition 2.2.1. Ein Maß auf den Borelmengen von G heißt *regulär*, wenn gilt:

- *Kompakte Teilmengen haben ein endliches Maß.*
- *Für jede Borelmenge B ist $\mu(B)$ das Supremum der $\mu(K)$, wenn K alle kompakten Teilmengen von B durchläuft.*
- *Für jede Borelmenge B ist $\mu(B)$ das Infimum der $\mu(O)$, wenn O alle offenen Obermengen von B durchläuft.*

Die üblicherweise verwendeten Maße sind regulär (insbesondere zum Beispiel das Borel-Lebesguemaß auf dem \mathbb{R}^n), es gibt aber auch nichtreguläre Maße.

Maße und Funktionale

Seit dem vorigen Jahrhundert ist bekannt, dass es *zwei gleichberechtigte Zugänge zur Maßtheorie* gibt:

- Erstens der Standardzugang, bei dem Maße den Ausgangspunkt bilden. Mit Maßen definiert man dann ein Integral $f \mapsto \int_G f d\mu$, das ist eine lineare Abbildung mit gewissen Monotonie- und Stetigkeitseigenschaften.
- Zweitens kann man das auch umkehren. Ist L ein lokalkompakter Raum und $I : C_{00}L \rightarrow \mathbb{R}$ eine monotone lineare Abbildung, so kann man ein (eindeutig bestimmtes) reguläres Borelmaß μ so finden, dass $I(f) = \int_L f d\mu$ für alle f gilt; dabei bezeichnet $C_{00}(L)$ die stetigen Funktionen mit kompaktem Träger.

In der Funktionalanalysis wird die Tatsache, dass beide Zugänge gleichwertig sind, der *Satz von Riesz* genannt.

Um mit solchen Maßen zu arbeiten, braucht man „viele kompakte Umgebungen“. Damit werden im Folgenden einige der in den Beispielen genannten Gruppen nicht weiter behandelt werden können. Zum Beispiel hat ein unendlich dimensionaler normierter Raum überhaupt keine kompakten Nullumgebungen.

Das Haarmaß²⁾

Definition 2.2.2. *Es sei (G, \circ) eine lokalkompakte topologische Gruppe und μ ein reguläres Borelmaß auf G .*

(i) *Es heißt ein rechtes Haarmaß, wenn es nichttrivial ist und wenn gilt: Für jede Borelmenge B und jedes $g \in G$ gilt $\mu(B) = \mu(B_g)$; dabei bezeichnet B_g die Menge $\{h \circ g \mid h \in B\}$.*

(ii) *Linke Haarmaße werden analog definiert. Da soll $\mu(B) = \mu({}_g B)$ gelten, wobei ${}_g B := \{g \circ h \mid h \in B\}$.*

(iii) *Eine lineare positive Abbildung $H : C_{00}(G) \rightarrow \mathbb{R}$ heißt rechtes Haarintegral, wenn stets $H(f_g) = H(f)$ gilt; dabei ist f_g durch $h \mapsto f(h \circ g)$ definiert.*

(iv) *Linke Haarintegrale werden analog definiert: f und ${}_g f$ (das ist die Abbildung $h \mapsto f(g \circ h)$) haben den gleichen Wert.*

Stillschweigend wurde in dieser Definition ausgenutzt, dass Translationen von Borelmengen wieder Borelmengen sind. Das ist mit den üblichen Techniken (siehe etwa Abschnitt 1.5 in meinem Buch „Elementare Stochastik“) leicht zu beweisen: Die Stetigkeit von $h \mapsto g \circ h$ und $h \mapsto h \circ g$ spielt dabei eine wichtige Rolle. Auch muss bemerkt werden, dass mit f auch f_g und ${}_g f$ zu $C_{00}(G)$ gehören. Im Folgenden kann auf solche Feinheiten nur ausnahmsweise eingegangen werden.

Beispiele

1. Ist (G, \circ) diskret topologisiert, so ist das zählende Maß ein linkes und ein rechtes Haarmaß. Positive Vielfache von einem Haarmaß sind wieder Haarmaß. Um wenigstens in Spezialfällen zu erreichen, dass das Haarmaß eindeutig bestimmt ist, vereinbart man meist, dass G Maß Eins hat, wenn G kompakt ist. Im

vorliegenden Fall heißt das, dass man im Fall endlicher G das durch die Anzahl der Elemente von G geteilte zählende Maß als *das* Haarmaß betrachtet.

2. Auf dem $(\mathbb{K}^n, +)$ ist das Lebesguemaß linkes und rechtes Haarmaß. Allgemein gilt offensichtlich, dass „links“ und „rechts“ für kommutative Gruppen zusammenfallen.

3. Sei (Γ, \cdot) die multiplikative Gruppe der $z \in \mathbb{C}$ mit $|z| = 1$. Das durch 2π geteilte eindimensionale Lebesguemaß ist Haarmaß.

4. Nun wird ein allgemeines Verfahren zur Erzeugung von Haarmaßen beschrieben, das auf dem Transformationssatz für Integrale beruht (siehe z.B. Behrends, „Elementare Stochastik“, Anhang zur Analysis).

Sei dazu (G, \circ) eine topologische Gruppe, die topologisch als offene Teilmenge eines \mathbb{R}^n aufgefasst werden kann. Links- und Rechtsmultiplikation sollen differenzierbare Abbildungen sein, wobei noch gilt:

- Die Jacobideterminante der Abbildung $h \mapsto g \circ h$ hängt – bei beliebigen h_0 – nur von g ab; wir bezeichnen sie mit λ_g .
- Die Jacobideterminante der Abbildung $h \mapsto h \circ g$ hängt – bei beliebigen h_0 – nur von g ab; wir bezeichnen sie mit ρ_g .

Dann gilt:

- $B \mapsto \int_B (1/|\lambda_x|) dx$ ist ein linkes Haarmaß auf G .
Äquivalent: $f \mapsto \int_G (f(x)/|\lambda_x|)$ ist linkes Haarintegral.
- $B \mapsto \int_B (1/|\rho_x|) dx$ ist ein rechtes Haarmaß auf G .
Äquivalent: $f \mapsto \int_G (f(x)/|\rho_x| dx)$ ist rechtes Haarintegral.

Es handelt sich dabei nur um eine Anwendung des Transformationssatzes. Hier einige Spezialfälle:

4a) Betrachte $G = (\mathbb{R} \setminus \{0\}, \cdot)$. Es ist $\lambda_x = \rho_x = x$, und folglich ist

$$f \mapsto \int_G \frac{f(x)}{|x|} dx$$

Links- und Rechts-Haarintegral. Qualitativ heißt das, dass die Länge anders als in \mathbb{R} gemessen wird. In der Nähe der 0 wird sie gestreckt, und „weit draußen“ wird sie gestaucht.

4b) Wir fassen $(\mathbb{C} \setminus \{0\}, \cdot)$ als offene Teilmenge des \mathbb{R}^2 auf. Die \mathbb{C} -Multiplikation mit einem (x, y) hat bei (x', y') die Jacobideterminante $x^2 + y^2$, und wir erhalten das Haarintegral

$$f \mapsto \int_G \frac{f(x + iy)}{x^2 + y^2} dx dy.$$

4c) Auf $GL(n, \mathbb{R})$ ist

$$f \mapsto \int_G \frac{f(A)}{|\det(A)|^n} dA$$

ein Links- und Rechts-Haarintegral; integriert wird bzgl. des Lebesguemaßes auf dem \mathbb{R}^n .

Existenz

Wir wollen nun zeigen, dass Haarmaße auf lokalkompakten Gruppen stets existieren. Wir konzentrieren uns auf den Fall von linken Haarmaßen, für rechte kann analog argumentiert werden.

Sei also (G, \circ) eine lokalkompakte Gruppe. In der Funktionalformulierung kann das Problem wie folgt umgeschrieben werden:

$C_{00}^+(G)$ bezeichne die nichtnegativen Funktionen in $C_{00}(G)$. Finde eine nichttriviale Abbildung $H : C_{00}^+(G) \rightarrow [0, \infty[$ mit folgenden Eigenschaften:

- $H(f + g) = Hf + Hg$;
- $H(af) = aHf$ für alle $a \geq 0$.
- $H(gf) = Hf$ für alle g, f .

So ein H kann durch $H(f - g) := Hf - hg$ leicht zu einer linearen Abbildung auf $C_{00}G$ fortgesetzt werden, und wegen der Äquivalenz Maße – Integrale ist damit ein linkes Haarmaß gefunden.

Es ist alles andere als offensichtlich, wie so etwas zu konstruieren ist. Die Idee soll am Beispiel $(\mathbb{R}, +)$ motiviert werden: Wie können wir das Lebesguemaß konstruieren, wobei nur die Topologie und die Gruppenstruktur ausgenutzt werden dürfen? Man kann so vorgehen:

- Fixiere eine „Eichfunktion“ $f_0 \in C_{00}^+\mathbb{R}$: Die soll später Integral Eins haben.
- Wähle eine „stark bei 0 konzentrierte“ Funktion $\psi \in C_{00}^+\mathbb{R}$. Der Einfachheit halber wollen wir uns ψ als charakteristische Funktion von $[-\varepsilon, \varepsilon]$ vorstellen.
- Wenn wir dann ein Haarintegral hätten, so könnten wir f_0 gut durch eine Funktion des Typs $\sum_{i=1}^n c_i(x_i\psi)$ approximieren, und deswegen müsste $H(f_0)$ etwa gleich $(\sum_i c_i)H(\psi)$ sein.

Für beliebiges f mache man es ebenso: $Hf \approx (\sum_j d_j)H(\psi)$, wobei die Funktion $\sum_{j=1}^m d_j(y_j\psi)$ eine gute Approximation an f ist.

Und zusammen heißt das (weil ja $H(f_0) = 1$ gelten soll), dass Hf näherungsweise $(\sum_j d_j)/(\sum_i c_i)$ sein sollte.

- Das ist eine Größe, die man nur durch Translationen und Approximationen ausrechnen kann. Leider wird sie nicht exakt linear von f abhängen. Die Linearität wird aber immer besser, je stärker ψ bei Null konzentriert ist, und deswegen muss die Konstruktion noch mit einem Limesprozess kombiniert werden.

Mit dieser Strategie – sie stammt von A. Weil – wird es allgemein klappen. Hier sind die wichtigsten Schritte der Konstruktion.

1. *Schritt:* Fixiere $f_0, \psi \in C_{00}^+(G)$. Später soll f_0 Integral Eins haben, und man sollte sich vorstellen, dass ψ bei e stark konzentriert ist. Genauer soll das bedeuten, dass wir das folgende Verfahren mit verschiedenen ψ durchführen werden, und im k -ten Durchlauf soll ψ seinen Träger in U_k haben, wo $U_1 \supset U_2 \supset \dots$ eine Basis der Umgebungen von e ist. Im metrischen Fall etwa könnte der Träger von ψ in der Kugel um e mit dem Radius $1/n$ enthalten sein.

Wir definieren dann, für $f \in C_{00}^+(G)$, die Zahl $(f : \psi)$ als das Infimum der Zahlen $\sum_{i=1}^n c_i$, wobei die c_i nichtnegativ sind und g_i existieren mit $\sum_{i=1}^n c_i (g_i \psi) \geq f$. Man muss sich natürlich vergewissern, dass das wohldefiniert ist. Man beweist dann:

- $\|f\| \leq (f : \psi) \|\psi\|$ (dabei bezeichnet $\|\cdot\|$ die Supremumsnorm).
- $(gf : \psi) = (f : g\psi) = (f : \psi)$ für alle g .
- $(af : \psi) = a(f : \psi)$ für $a \geq 0$.
- $(f_1 + f_2 : \psi) \leq (f_1 : \psi) + (f_2 : \psi)$.
- $(f : \psi) \leq (f : \phi)(\phi : \psi)$.

Unser Kandidat für eine Approximation an das Haarintegral ist dann

$$I_\psi : f \mapsto \frac{(f : \psi)}{(f_0 : \psi)}.$$

Damit ist klar:

- $I_\psi(gf) = I_\psi(f)$.
- $I_\psi(af) = aI_\psi(f)$ für $a \geq 0$.
- Aus $f_1 \leq f_2$ folgt $I_\psi(f_1) \leq I_\psi(f_2)$.
- $I_\psi(f_1 + f_2) \leq I_\psi(f_1) + I_\psi(f_2)$.
- $I_\psi(f_0) = 1$.

•

$$\frac{1}{(f_0 : f)} \leq I_\psi(f) \leq (f : f_0).$$

Es ist dann fundamental, dass I_ψ beinahe linear ist, wenn der Träger von ψ klein genug ist. Genauer:

- Für alle $\varepsilon > 0$ und alle f_1, f_2 existiert eine Nullumgebung U von e mit der folgenden Eigenschaft: Wenn der Träger von ψ in U enthalten ist, dann gilt $I_\psi(f_1 + f_2) \geq I_\psi(f_1) + I_\psi(f_2) - \varepsilon$.

2. *Schritt:* Nun wählen wir eine Basis der e -Umgebungen $U_1 \supset U_2 \supset \dots$ und Funktionen ψ_n mit Träger in U_n . Der Schluss des Beweises läuft im Wesentlichen darauf hinaus zu zeigen, dass man $I(f)$ als Grenzwert der $I_{\psi_n}(f)$ definieren kann. Nach Konstruktion ist das ein Haarintegral.

Eindeutigkeit

Haar-Maße sind im Wesentlichen eindeutig bestimmt. Es gilt:

Satz 2.2.3. *Es sei (G, \circ) eine lokalkompakte Gruppe, und I_1, I_2 seien links-Haar-Integrale. Dann gibt es ein $c > 0$ mit $I_2 = cI_1$. Ein analoges Ergebnis gilt für Rechts-Haarintegrale.*

Beweis: Alle Funktionen werden in $C_{00}^+(G)$ sein. Wenn sie nichttrivial sind, ist jedes Haar-Integral dann strikt positiv.

1. Fixiere so ein f und definiere $\nabla_f : G \rightarrow \mathbb{R}^+$ durch $g \mapsto I_2(f_g)/I_1(f)$. ∇_f ist stetig und strikt positiv.

2. Behauptung: Für jede Funktion $h \in C_{00}^+(G)$ gilt

$$I_2(\check{h}) = I_1(\nabla_f h);$$

dabei ist $\check{h}(g) := h(g^{-1})$.

Beweis dazu: Definiere \check{I}_2 durch $f \mapsto I_2(\check{f})$. Das ist ein Rechts-Haarintegral. $\mu_1, \mu_2, \check{\mu}_2$ seien die zu I_1, I_2, \check{I}_2 gehörigen Maße auf G . In den folgenden Umformungen nutzen wir (neben dem Satz von Fubini) aus:

- Die Definition von μ_1 und $\check{\mu}_2$.
- Die Rechtsinvarianz von $\check{\mu}_2$.
- Die Linksinvarianz von μ_1 .
- Die Definition von $\check{\mu}_2$.
- Die Definition von ∇_f

$$\begin{aligned} I_1(f)I_2(\check{h}) &= \int \int_{G^2} f(g)h(g') d\mu_1(g) d\check{\mu}_2(g') \\ &= \int \int_{G^2} f(g)h(g'g) d\check{\mu}_2(g') d\mu_1(g) \\ &= \int \int_{G^2} f((g')^{-1}g)h(g) d\mu_1(g) d\check{\mu}_2(g') \\ &= \int_G h(g) \int_G f((g')^{-1}g) d\check{\mu}_2(g') d\mu_1(g) \\ &= \int_G h(g) \int_G f(g'g) d\mu_2(g') d\mu_1(g) \\ &= \int_G h(g)I_2(f_g) d\mu_1(g) \\ &= I_1(f)I_1(h\nabla_f). \end{aligned}$$

Nun muss man nur noch $I_1(f)$ kürzen.

3. Es folgt, dass ∇_f nicht von f abhängt, d.h. es existiert eine stetige Funktion ∇ , so dass $I_2(f_g) = \nabla(g)I_1(f)$ für alle f gilt. Auswertung bei $g = e$ ergibt die Behauptung. \square

Die Modularfunktion

Und wie hängen Links- und Rechts-Haarintegrale zusammen? Das klärt der folgende

Satz 2.2.4. *Sei (G, \circ) eine lokalkompakte Gruppe. Dann gibt es einen eindeutig bestimmten stetigen Gruppenhomomorphismus $\Delta_G : G \rightarrow (\mathbb{R}^+ \setminus \{0\}, +)$ mit der folgenden Eigenschaft.*

- Ist $I : C_0(G) \rightarrow \mathbb{R}$ ein Links-Haarintegral, so ist $f \mapsto I(f/\Delta_G)$ ein Rechts-Haarintegral und umgekehrt.

Δ_G heißt die Modularfunktion von G , und G heißt unimodular, wenn $\Delta_G = 1$ gilt.

Beweis: I sei ein beliebiges Links-Haarintegral.

1. Fixiere f und definiere $\Delta_G : G \rightarrow \mathbb{R}^+$ durch $g \mapsto I(f_{g^{-1}})/I(f)$.
2. Die Definition hängt nicht von f ab. (Nutze dazu aus, dass $f \mapsto I(f_{g_0^{-1}})$ ein Links-Haarintegral ist.)
3. Δ_G ist stetig.
4. Δ_G ist ein Gruppenhomomorphismus.
5. Es gilt $I(f) = I(\check{f}/\Delta_G)$ für alle f .
6. $I(\check{f}) = I(f/\Delta_G)$ für alle f . Damit fehlt noch die Eindeutigkeit.
7. Δ_G ist eindeutig bestimmt. (Δ sei ein weiterer Kandidat. Nutze aus, dass $f \mapsto I(f/\Delta)$, $I(f/\Delta_G)$ Rechts-Haarintegrale sind.) \square

Satz 2.2.5. *Kompakte (insbesondere endliche) Gruppen sind unimodular.*

Beweis: $\Delta_G(G)$ ist eine kompakte Untergruppe von (\mathbb{R}^+, \cdot) . Es gibt aber nur eine, nämlich $\{1\}$. Fertig!³⁾ \square

Bemerkung: Das obige Beispiel 4c zeigt, dass auch „komplizierte“ Gruppen unimodular sein können. Hier ein Beispiel für eine nicht-unimodulare Gruppe:

Sei G die Menge aller (x, y) mit reellen x, y und $x \neq 0$. Das Produkt sei so erklärt:

$$(x, y)(u, v) = (xu, xv + y).$$

(Das ist die gewöhnliche Matrizenmultiplikation, wenn man die (x, y) mit Matrizen $\begin{pmatrix} x & y \\ 0 & 1 \end{pmatrix}$ identifiziert.) Hier ist $\Delta_G(x, y) = 1/|x|$.

³⁾Das ist nach Meinung des Dozenten einer der elegantesten Beweise der Mathematik.

Faltungen und die Gruppenalgebra

(G, \circ) sei eine lokalkompakte Gruppe, und μ sei ein linkes Haarmaß. Für $f, h \in L^1(G, \mu)$ erklären wir dann die *Faltung* $f * h$ und die *Involution* f^* durch

- $(f * h)(g) := \int_G f(\tilde{g})h(\tilde{g}^{-1}g) d\mu(\tilde{g})$.
- $f^*(g) := \overline{f(g^{-1})}$.

Damit wird der $L^1(G)$ zu einer Banach-Algebra mit Involution: Das ist die zu G gehörige *Gruppenalgebra*. Sie spielt eine wichtige Rolle, wenn man funktionalanalytische Techniken und Konzepte anwenden möchte.

2.3 Darstellungen und Charaktere

Eine Gruppe ist ein abstraktes Gebilde. Wenn man – zum Beispiel – Symmetrien untersuchen möchte, muss man die Gruppe durch eine Gruppe von Abbildungen „konkret“ darstellen. Die Darstellungstheorie topologischer Gruppen untersucht, auf welche Weise das möglich ist. Die Fragestellung ist deswegen fundamental wichtig, weil man natürlich wissen möchte, wie sich eine abstrakte Gruppe „konkretisieren“ kann.

Das Thema „Symmetrie“ kann auf diese Weise auch einheitlich behandelt werden. Ist eine Gruppe durch gewisse Abbildungen „dargestellt“, so wird man ein Element des zugehörigen Raumes symmetrisch nennen, wenn es unter allen diesen Abbildungen invariant ist. Hier einige Beispiele:

- Es soll ein mechanisches System untersucht werden, bei dem m Massenpunkte eine Rolle spielen. Es kann durch m Funktionen y_1, \dots, y_m beschrieben werden. Sind die Punkte nicht unterscheidbar, so gibt wird die Permutationsgruppe über m Elementen als Symmetriegruppe auftreten. Sie wird durch Vertauschung der entsprechenden Funktionen dargestellt. (Als Beispiel denke man an ein Doppelpendel.)
- Ist $y : \mathbb{R} \rightarrow \mathbb{R}$ eine Lösung von $y^{(n)} + a_{n-1}y^{(n-1)} + \dots + a_1y' + a_0y = 0$, so auch $t \mapsto y(t + t_0)$ für jedes t_0 . Das gibt in natürlicher Weise Anlass zu einer Darstellung von $(\mathbb{R}, +)$ im Lösungsraum dieser homogenen Differentialgleichung.

Darstellungen (Definitionen)

Hier die Definition für beliebige Gruppen:

Definition 2.3.1. *Es sei (G, \circ) eine Gruppe und V ein Vektorraum. Eine Darstellung von G über V ist dann ein Gruppenhomomorphismus $\pi : G \rightarrow [V]_{iso}$, wobei $[V]_{iso}$ die Isomorphismen (also die bijektiven linearen Abbildungen) von V nach V bezeichnet.*

Das ist für unsere Zwecke zu allgemein. Wir benötigen eine Definition, die die Topologie von G berücksichtigt und die „geometrischer“ ist:

Definition 2.3.2. *Es sei (G, \circ) eine topologische Gruppe⁴⁾ und H ein komplexer Hilbertraum⁵⁾. Wir bezeichnen mit $U(H)$ die Menge der unitären Abbildungen auf H , das sind die linearen $T : H \rightarrow H$ mit $TT^* = T^*T = Id$.*

(i) *Eine unitäre Darstellung von G über H ist dann ein Gruppenhomomorphismus von G nach $U(H)$, für den $g \mapsto (\pi(g))(x)$ für jedes $x \in H$ eine stetige Abbildung von G nach H ist.*

(ii) *Ist H eindimensional, so entsprechen die unitären T gerade den z mit $|z| = 1$. Deswegen können unitäre Darstellungen in diesem Fall mit den stetigen Gruppenhomomorphismen $\chi : G \rightarrow \Gamma$ identifiziert werden; dabei steht Γ für die multiplikative Gruppe der $z \in \mathbb{C}$ mit $|z| = 1$. Solche χ heißen Charaktere von G .*

Wir werden uns im Folgenden mit den folgenden Problemen beschäftigen, um die Frage „Wie kann eine Gruppe wirken?“ zu beantworten:

- Wann sind zwei Darstellungen als „im Wesentlichen gleich“ anzusehen?
- Was sind die „einfachsten“ Darstellungen, und kann man alle Darstellungen aus einfachen zusammensetzen?
- Gibt es „genügend viele“ Darstellungen?

Beispiele:

0. Ist $\pi(g) = Id$ für alle g , so spricht man von der *trivialen Darstellung*

1. Charaktere spielen – besonders im Fall kommutativer G – eine wichtige Rolle.

Hier einige Beispiele für Charaktere

1a. Zunächst betrachten wir die Gruppe Γ (die übrigens isomorph zur Gruppe $[-\pi, \pi]$ ist, wenn wir sie mit der Addition modulo 2π versehen). Es gilt: Charaktere sind genau die Abbildungen $z \mapsto z^n$, wobei $n \in \mathbb{Z}$.

1b. Nur untersuchen wir $(\mathbb{Z}, +)$ mit der diskreten Topologie. Hier gilt: χ ist genau dann Charakter, wenn χ von der Form $n \mapsto z^n$ für ein geeignetes $z \in \Gamma$ ist. (Das ist leicht einzusehen: Setze $z := \chi(1)$.)

1c. Auf der Gruppe (\mathbb{Z}_n, \cdot) sind genau die Abbildungen $k \mapsto z^k$ die Charaktere, wobei $z := e^{2\pi i/n}$.

⁴⁾Fast immer werden wir uns auf lokalkompakte Gruppen konzentrieren.

⁵⁾Das ist ein \mathbb{C} -Vektorraum, der mit einem Skalarprodukt $\langle \cdot, \cdot \rangle$ versehen ist und bezüglich der davon induzierten Norm vollständig ist. Zum Kennenlernen kann man an den \mathbb{C}^n mit dem Skalarprodukt $\sum_j x_j \overline{y_j}$ denken.

1d. Die Charaktere auf $(\mathbb{R}, +)$ sehen so aus: Es sind die Abbildungen $x \mapsto e^{iyx}$ mit $y \in \mathbb{R}$.

1e. Die Abbildung, die einer Permutation das Signum dieser Permutation in $\{-1, 1\}$ zuordnet, ist ein Charakter auf der Gruppe der Permutationen in n Elementen.

Die Menge der Charaktere ist offensichtlich wieder eine abelsche Gruppe bzgl. der punktweisen Multiplikation. Sie wird mit \hat{G} bezeichnet, man spricht von der *dualen Gruppe*. Versieht man \hat{G} mit der Topologie der gleichmäßigen Konvergenz auf Kompakta, so ergibt sich wieder eine topologische Gruppe.

2. G sei lokalkompakt, und μ sei ein linkes Haarmaß. Die *linksreguläre Darstellung* ist dann wie folgt definiert:

- Der Hilbertraum ist der komplexe $L^2(G, \mu)$.
- Für $g \in G$ ist $\pi(g)$ durch $\pi(g)(f) := g^{-1}f$ erklärt.

Das ist eine unitäre Darstellung; dabei ist es wichtig, dass μ linksinvariant ist.

Äquivalente Darstellungen

Wann sind zwei Darstellungen im Wesentlichen gleich? Sicher dann, wenn die zugrunde liegenden Hilberträume isomorph sind und die $\pi(g)$ modulo dieser Isomorphie gleich sind. Genauer:

Definition 2.3.3. (π_1, H_1) und (π_2, H_2) seien unitäre Darstellungen von (G, \circ) . π_1, π_2 heißen äquivalent, wenn es einen Hilbertraumisomorphismus $V : H_1 \rightarrow H_2$ so gibt, dass $\pi_2(g) \circ V = V \circ \pi_1(g)$ für alle g gilt.

Insbesondere ist für jede unitäre Darstellung (π, H) und jede unitäre Abbildung $V : H \rightarrow H$ die Darstellung $g \mapsto V\pi(g)V^{-1}$ zu π äquivalent.

Konstruktion weiterer Darstellungen

Zwei Methoden, aus schon bekannten Darstellungen eine neue zu gewinnen, spielen eine wichtige Rolle.

Methode 1: Produkte: (π_1, H_1) und (π_2, H_2) seien unitäre Darstellungen von (G, \circ) . Wir definieren einen Hilbertraum H als Produkt-Hilbertraum $H_1 \times H_2$ und eine Darstellung π durch $g \mapsto \pi_1(g) \times \pi_2(g)$. Das ist, wie leicht zu sehen, wieder eine unitäre Darstellung, die *Produktdarstellung*.

Methode 2: Einschränkungen: (π, H) sei eine unitäre Darstellung von G , und $L \subset H$ sei ein invarianter Unterraum: für alle g bildet $\pi(g)$ diesen Unterraum auf sich ab. Definiert man dann (π_L, L) durch $g \mapsto \pi(g)|_L$, so ist das wieder eine unitäre Darstellung.

Irreduzibilität

Eine Darstellung heißt *irreduzibel*, wenn sie nicht aus einfacheren Darstellungen zusammengesetzt ist⁶⁾. Hier die präzisere

Definition 2.3.4. (π, G) heißt irreduzibel, wenn es nicht möglich ist, Darstellungen $(\pi_1, H_1), (\pi_2, H_2)$ mit $H_1 \neq \{0\} \neq H_2$ zu finden, so dass (π, H) äquivalent zur Summe dieser Darstellungen ist.

Das ist sicher eine wichtige Definition. Sie ist aber recht unhandlich, und deswegen ist es gut zu wissen, dass man Irreduzibilität besser charakterisieren kann:

Satz 2.3.5. (Schursches Lemma) Die folgenden Aussagen sind für eine unitäre Darstellung äquivalent:

(i) (π, H) ist irreduzibel.

(ii) In H gibt es keine nichttrivialen invarianten Unterräume.

(iii) Ist $T : H \rightarrow H$ linear und stetig und gilt $T \circ \pi(g) = \pi(g) \circ T$, so ist T ein Vielfaches der Identität. Kurz: Der Kommutator der Menge $\pi(G)$ ist trivial.

Beweis: (i) \Rightarrow (ii): Sei $H' \subset H$ ein invarianter Unterraum. Wir wollen zeigen, dass L trivial (also gleich $\{0\}$ oder H) ist. In einem ersten Schritt zeigt man, dass auch L^\perp invariant ist: Das folgt daraus, dass die $\pi(g)$ unitär sind. Wäre nun L nicht trivial, so würde man π als nichttriviale Summe der Einschränkungen von π auf L und L^\perp darstellen können. Widerspruch.

(ii) \Rightarrow (iii): Dieser Beweisteil ist schwieriger. Wir definieren R als den Kommutator von $\pi(G)$, also als die Gesamtheit der T , die mit allen $\pi(g)$ kommutieren. Es ist leicht zu sehen, dass R eine unter der punktweisen Konvergenz abgeschlossene Unteralgebra mit Einheit der Gesamtheit aller linear stetigen $T : H \rightarrow H$ ist und dass mit jedem T auch T^* in R liegt.

Schritt 1: Wenn P eine orthogonale Projektion in R ist (also $P = P^* = P^2$), so ist $P \in \{0, \text{Id}\}$.

Beweis dazu: Sei L der Bildraum von P . Dann ist L ein invarianter Unterraum, also $L \in \{\{0\}, H\}$ nach Voraussetzung. Fertig.

Schritt 2: Wenn alle Projektionen in R trivial sind, so sind alle $T \in R$ trivial.

Beweis dazu: Da mit T auch T^* in R liegt, dürfen wir annehmen, dass $T = T^*$ gilt. Wir konzentrieren uns zunächst auf den Fall endlich-dimensionaler H , denken also etwa an den \mathbb{C}^n .

Sei A eine selbstadjungierte Matrix. Wir können sie als Diagonalmatrix mit reellen Einträgen schreiben. Durch (vorläufigen) Übergang von A zu $B := A + r\text{Id}$ mit einem „großen“ r dürfen wir annehmen, dass alle Eigenwerte positiv sind. Wir sortieren sie der Größe nach: $0 < \lambda_1 < \dots < \lambda_k$. Die Matrix B/λ_k ist diagonal, die letzten Einträge sind 1, die anderen liegen zwischen 0 und 1. Folglich konvergieren die Potenzen gegen eine Projektion P_k , sie projiziert auf den Unterraum, der zu den Einsen gehört.

⁶⁾Ein etwas gewagter Vergleich: Wenn Darstellungen Zahlen entsprechen, so entsprechen die irreduziblen Darstellungen gerade den Primzahlen.

Es folgt: P_k liegt auch in R , und $B - \lambda_k P_k$ hat da, wo eben noch die λ_k standen, nur noch Nullen. So setzen wir das Verfahren fort, am Ende haben wir P_1, \dots, P_k mit $B = \lambda_1 P_1 + \dots + \lambda_k P_k$ gefunden, wo alle P_κ in R liegen. Es ist

$$A = B - r\text{Id} = \lambda_1 P_1 + \dots + \lambda_k P_k - r\text{Id},$$

dabei sind nach dem ersten Schritt alle P_κ trivial. Damit ist A ein Vielfaches der Identität.

Der Beweis im allgemeinen Fall ist im Wesentlichen genauso, setzt allerdings viel Funktionalanalysis voraus. Wesentlich ist hier, dass in einer punktweise abgeschlossenen selbstadjungierten Unteralgebra mit Einheit der Operatoren auf H (das sind die so genannten von-Neumann-Algebren) jeder selbstadjungierte Operator als Integral über geeignete Projektionen, die in dieser Algebra liegen, dargestellt werden kann: Das ist der *Spektralsatz für selbstadjungierte Operatoren*, er wird hier auf den Kommutator angewendet.

(iii) \Rightarrow (i): Ist (π, H) *nicht* irreduzibel, so können wir nach geeigneter Identifizierung annehmen, dass H als $L \times L^\perp$ mit invarianten und nichttrivialen Unterräumen L, L^\perp geschrieben werden kann. Dann ist aber die Projektion auf L ein Element im Kommutator, das nicht von der Form λId ist. \square

Kapitel 3

Darstellungstheorie spezieller Gruppen

Wir werden nun die bisherigen Überlegungen für einige ausgewählte Klassen von Gruppen vertiefen. Es fängt sehr elementar mit endlichen abelschen Gruppen an. Dieser Abschnitt wurde deswegen aufgenommen, weil viele wichtige Ideen schon an dieser einfachen Situation erklärt werden können. Nach und nach werden wir uns dann immer kompliziertere Situationen vornehmen.

3.1 Endliche abelsche Gruppen

In diesem Abschnitt ist G eine endliche abelsche Gruppe. Da sie abelsch ist, schreiben wir die innere Komposition als $+$ und das neutrale Element als 0 . (Standardbeispiele sind die \mathbb{Z}_n und endliche Produkte dieser Gruppen, und in der Algebra beweist man, dass jede endliche abelsche Gruppe diese Form hat.)

Charaktere

Sei zunächst G zyklisch, also $G = \{0, a, 2a, \dots, (n-1)a\}$ für geeignete n und a . Ist χ Charakter, so ist χ durch den Wert $\chi(1)$ eindeutig bestimmt, und das muss eine n -te Einheitswurzel sein. Umgekehrt kann man auch χ durch Vorgabe so einer Einheitswurzel definieren. Kurz:

- Es gibt genau n Charaktere.
- (\hat{G}, \cdot) ist isomorph zur multiplikativen Gruppe der n -ten Einheitswurzeln und folglich wieder zyklisch. Damit sind G und \hat{G} isomorph (wenn auch nicht auf kanonische Weise).

Da \hat{G} explizit beschrieben werden kann, lassen sich Aussagen leicht ablesen, die im allgemeinen Fall nur wesentlich aufwändiger gezeigt werden können. Zum Beispiel: Für jedes $b \neq 0$ in G gibt es ein χ mit $\chi(g) \neq 1$. Es folgt leicht: Charaktere trennen die Punkte von G .

Es ist leicht zu sehen, dass $\widehat{G_1 \times G_2} = \hat{G}_1 \times \hat{G}_2$. Es folgt: Ist G von der Form

$$\{0, a_1, 2a_1, \dots, (n_1 - 1)a_1\} \times \dots \times \{0, a_r, 2a_r, \dots, (n_r - 1)a_r\},$$

so kann man die Charaktere auf G wie folgt beschreiben.

Für jeden Charakter χ gibt es eine n_1 -Einheitswurzel ω_1 , eine n_2 -te Einheitswurzel ω_2, \dots und eine n_r -te Einheitswurzel ω_r , so dass stets

$$\chi(k_1 a_1, k_2 a_2, \dots, k_r a_r) = (\omega_1)^{k_1} \dots (\omega_r)^{k_r}$$

gilt.

Insbesondere sind G und \hat{G} wieder isomorph. Wir erinnern noch einmal daran, dass alle endlichen abelschen G diese Form haben, \hat{G} ist also immer bekannt.

Wir beweisen nun einige allgemeine Eigenschaften, die schon an unsere ersten Ergebnisse zur Fourieranalyse erinnern.

Lemma 3.1.1. *Sei $(G, +)$ eine endliche abelsche Gruppe, $\#G$ sei die Anzahl der Elemente in G .*

(i) *Für $\chi \in \hat{G}$ mit $\chi \neq 1$ ist $\sum_g \chi(g) = 0$.*

(ii) *Für $\chi_1, \chi_2 \in \hat{G}$ mit $\chi_1 \neq \chi_2$ ist $\sum_g \chi_1(g) \overline{\chi_2(g)} = 0$.*

(iii) *Sei X_G der $\#G$ -dimensionale Raum der Funktionen von G nach \mathbb{C} . Definiere auf X_G ein Skalarprodukt durch*

$$\langle f_1, f_2 \rangle := \frac{1}{\#G} \sum_g f_1(g) \overline{f_2(g)}.$$

Dann ist \hat{G} eine Orthonormalbasis in X_G .

Beweis: Lemma 15.3 in [Be1]. □

Korollar 3.1.2. (i) *Jedes $f \in X_G$ kann als $\sum_{\chi \in \hat{G}} \langle f, \chi \rangle \chi$ geschrieben werden.*

(ii) *\hat{G} trennt die Punkte aus G .*

Beweis: Cor. 15.4 in [Be1]. □

Faltung

Für $f_1, f_2 \in X_G$ wird die *Faltung* von f_1 mit f_2 punktweise durch

$$(f_1 * f_2)(g) := \frac{1}{\#G} \sum_{h \in G} f_1(h) f_2(g - h)$$

erklärt. Sie hat wieder die üblichen Linearitätseigenschaften, und $(X_G, +, *)$ wird die *Gruppenalgebra* genannt. Die Faltung wird wichtig werden, wenn man sich um die additive Überlagerung unabhängiger Zufallsvariablen kümmert¹⁾.

Die Fouriertransformation

Nun kommen wir zur Fouriertransformation. Man sollte sich nachträglich den gruppentheoretischen Hintergrund der Überlegungen aus dem ersten Kapitel klar machen.

Definition 3.1.3. Für $f \in X_G$ definieren wir die Fouriertransformation von f durch

$$\hat{f} : \hat{G} \rightarrow \mathbb{C}, \chi \mapsto \frac{1}{\#G} \sum_g f(g)\chi(g).$$

Es gilt dann:

Satz 3.1.4. (i) $f \mapsto \hat{f}$ ist linear.

(ii) Die Fouriertransformation eines Charakters ist die Indikatorfunktion von $\{\chi\}$.

(iii) $f \mapsto \hat{f}$ ist surjektiv und folglich bijektiv. Explizit gilt für $f \in X_G$ die folgende Inversionsformel:

$$f(g) = \sum_{\chi} \hat{f}(\chi)\overline{\chi(g)}$$

für alle g .

(iv) Es gilt die Plancherelformel:

$$\langle f_1, f_2 \rangle = \sum_{\chi} \hat{f}_1(\chi)\overline{\hat{f}_2(\chi)}.$$

Insbesondere gilt

$$\frac{1}{\#G} \sum_g |f(g)|^2 = \sum_{\chi} |\hat{f}(\chi)|^2$$

für alle $f \in X_G$. Anders ausgedrückt: Bei richtiger Normierung ist die Fouriertransformation ein Hilbertraum-Isomorphismus.

(v) Aus der Faltung wird das Produkt:

$$\widehat{f_1 * f_2}(\chi) = \hat{f}_1(\chi) \cdot \hat{f}_2(\chi).$$

Beweis: 15.3 bis 15.7 in [Be1]. □

Beliebige unitäre Darstellungen

So, wie Zahlen aus Primzahlen zusammengesetzt sind, können beliebige Darstellungen aus irreduziblen Darstellung aufgebaut werden:

¹⁾Das ist in der elementaren Stochastik genauso.

Satz 3.1.5. (i) Die irreduziblen Darstellungen entsprechen den Charakteren.

(ii) Sei (π, H) eine unitäre Darstellung von G . Dann zerfällt H in eine direkte Summe von abgeschlossenen Unterräumen H_χ , mit $\chi \in \hat{G}$. Alle H_χ sind invariant, und auf H_χ ist die Einschränkung von $\pi(g)$ das $\chi(g)$ -fache des identischen Operators auf H_χ .

(iii) Jede unitäre Darstellung π von G zerfällt in eine orthogonale Summe von eindimensionalen Darstellungen, wovon jede zu einem Charakter äquivalent ist. Kurz: π ist direkte Summe von irreduziblen Darstellungen.

(iv) Die Zerlegung ist im Wesentlichen eindeutig. Zu π ist eindeutig festgelegt, in welcher Vielfachheit die einzelnen χ darin auftreten.

Beweis: (i) Eindimensionale Darstellungen (d.h., Charaktere) sind trivialerweise irreduzibel. Für die Umkehrung wenden wir das Schursche Lemma 2.3.5 an. π sei eine irreduzible Darstellung auf dem Hilbertraum H . Nach dem Lemma ist der Kommutator gleich $\mathbb{C} \cdot \text{Id}$. Da die Gruppe kommutativ ist, enthält er $\pi(G)$: Zu jedem g gibt es also ein $\chi(g) \in \mathbb{C}$, so dass $\pi(g)$ das $\chi(g)$ -fache der Identität ist. Notwendig ist χ ein Charakter, und wenn der Hilbertraum mehr als eindimensional wäre, wäre leicht ein nichttriviale Zerlegung von π anzugeben.

(ii) Für $\chi \in \hat{G}$ definieren wir

$$H_\chi := \left\{ x \mid x = \frac{1}{\#G} \sum_g \bar{\chi}(g) \pi(g)x \right\}.$$

Dann lässt sich leicht zeigen:

- Jedes H_χ ist ein abgeschlossener Unterraum.
- Diese Räume stehen paarweise senkrecht aufeinander.
- H ist die direkte Summe der H_χ .
- Jedes H_χ ist invariant unter allen $\pi(g)$, und die Einschränkung von $\pi(g)$ ist das $\chi(g)$ -fache des identischen Operators.

(Im Beweis wird noch ausgenutzt, dass $\sum_\chi \chi(g)$ für alle $g \neq 0$ gleich Null ist.)

(iii) Das folgt sofort aus (ii): Man muss die H_χ nur noch in eine orthogonale Summe eindimensionaler Räume zerlegen.

(iv) Das liegt daran, dass die H_χ direkt mit Hilfe der χ konstruiert wurden. \square

Bemerkungen: 1. Man kann das Ergebnis so umformulieren: Bezüglich der Darstellung in einer geeigneten Orthonormalbasis sind alle $\pi(g)$ gleichzeitig orthogonal, und auf der Diagonalen stehen die $\chi(g)$. Testen Sie das an der \mathbb{Z}_2 , die durch Bewegungen auf dem \mathbb{R}^2 dargestellt werden soll. (Denken Sie an die Symmetriegruppen von Opel- und Renault-Logo).

2. Testen wir das Verfahren an der rechtsregulären Darstellung: $H = L^2(G)$, und $\pi(g)f = f_g$. Wir bemerken, dass $\chi \in H_\chi$ gilt, und da diese Räume orthogonal

aufeinander stehen, es genau $\#G$ davon gibt und H die gleiche Dimension hat, muss stets $H_\chi = \mathbb{C}\chi$ gelten. Anders ausgedrückt: Die natürliche Basis für diese Darstellung ist \hat{G} . Es ist auch zu betonen, dass alle χ in der rechtsregulären Darstellung zu finden sind. \square

3.2 Beliebige endliche Gruppen

Wir verzichten nun auf die Voraussetzung der Kommutativität und schreiben wieder \circ für die innere Komposition und e für das neutrale Element. Im Folgenden ist also (G, \circ) eine beliebige endliche Gruppe²⁾.

Zunächst mache man sich klar, dass man mit Charakteren nicht auskommen wird, denn Charaktere können zwischen $g \circ h$ und $h \circ g$ nicht unterscheiden. Wir werden uns also um höherdimensionale Darstellungen kümmern müssen. Allzu groß können sie nicht sein, wir beweisen – etwas allgemeiner als hier eigentlich erforderlich – ein Resultat für kompakte Gruppen:

Satz 3.2.1. *Irreduzible Darstellungen kompakter Gruppen sind endlichdimensional.*

Beweis: Sei μ ein zu Eins normiertes linkes Haarmaß auf G und (π, H) eine irreduzible Darstellung. Wir fixieren ein $y \in H$ und betrachten die Abbildung

$$(x_1, x_2) \mapsto \langle x_1, x_2 \rangle_y := \int_G \langle \pi(g)y, x_2 \rangle \overline{\langle \pi(g)y, x_1 \rangle} d\mu(g)$$

von $H \times H$ nach \mathbb{C} . Diese Abbildung hat fast alle Eigenschaften eines Skalarprodukts, es kann allerdings nicht garantiert werden, dass aus $\langle x, x \rangle_y = 0$ stets $x = 0$ folgt. Nun brauchen wir den

Satz von Lax-Milgram: Solche Abbildungen sind – falls sie stetig sind – notwendig von der Form

$$(x_1, x_2) \mapsto \langle Bx_1, x_2 \rangle$$

für einen geeigneten linearen stetigen Operator $B : H \rightarrow H$.

In unserem Fall heißt das, dass es ein $B_y : H \rightarrow H$ gibt, so dass stets

$$\langle x_1, x_2 \rangle_y = \langle B_y x_1, x_2 \rangle$$

gilt.

Aus der Tatsache, dass π Darstellung ist und dass μ invariant ist, lässt sich folgern, dass B_y mit allen $\pi(g)$ kommutiert. Es gibt also wegen des Schurschen Lemmas ein $a_y \in \mathbb{C}$ mit $B_y = a_y \text{Id}$. Es folgt, indem man $x_1 = x_2 = x$ setzt,

$$\int_G \langle \pi(g)y, x \rangle \overline{\langle \pi(g)y, x \rangle} d\mu(g) = a_y \|x\|^2.$$

²⁾Wir folgen hier im Wesentlichen Kapitel 16 in [Be1].

Im Integral darf g durch g^{-1} ersetzt werden, da G unimodular ist. Man kann deswegen die Rollen von x und y vertauschen und es folgt $a_y \|x\|^2 = a_x \|y\|^2$ für alle x, y . Das geht nur, wenn es ein a so gibt, dass $a_y = a \|y\|^2$ für alle y gilt. Wählt man ein $x \neq 0$, so ist $g \mapsto \langle \pi(g)x, x \rangle$ bei e strikt positiv, deswegen muss $\langle x, x \rangle_x$ strikt positiv sein, und das zeigt $a > 0$.

Nun kann die Endlichdimensionalität von H gezeigt werden. Es seien x_1, \dots, x_n orthonormale Vektoren. Wir haben schon

$$\int |\langle \pi(g)y, x \rangle|^2 d\mu(g) = a \|x\|^2 \|y\|^2$$

für alle x, y gezeigt. Insbesondere ist $\int |\langle \pi(g)x_k, x_1 \rangle|^2 d\mu(g) = a$ für $k = 1, \dots, n$. Andererseits sind – für jedes g – die Vektoren $\pi(g)x_1, \dots, \pi(g)x_n$ orthonormal, denn $\pi(g)$ ist unitär. Wegen der Besselschen Ungleichung heißt das, dass $\sum_k |\langle \pi(g)x_k, x_1 \rangle|^2 \leq \|x_1\|^2 = 1$ gilt, und so folgt

$$\begin{aligned} a \cdot n &= \sum_k \int_G |\langle \pi(g)x_k, x_1 \rangle|^2 d\mu(g) \\ &= \int_G \sum_k |\langle \pi(g)x_k, x_1 \rangle|^2 d\mu(g) \\ &\leq 1. \end{aligned}$$

Damit ist klar, dass nicht beliebig große n auftreten können, d.h., H ist endlichdimensional. \square

Ab hier sei G wieder endlich. Wie viele irreduzible Darstellungen gibt es? Wir werden zeigen, dass sie alle als Teildarstellungen der linksregulären Darstellung auftreten, und es wird folgen, dass es nur endlich viele geben kann. Wir beginnen mit einer wichtigen

Definition 3.2.2. (π_1, H_1) und (π_2, H_2) seien irreduzible Darstellungen von G . Wir identifizieren H_1 bzw. H_2 mit dem \mathbb{C}^{d_1} bzw. \mathbb{C}^{d_2} , die $\pi_1(g)$ bzw. die $\pi_2(g)$ können daher als Matrizen geschrieben werden.

Wir sagen, dass eine $d_2 \times d_1$ -Matrix A die Darstellungen π_1 und π_2 verbindet, wenn

$$A\pi_1(g) = \pi_2(g)A$$

für alle g gilt³⁾.

Lemma 3.2.3. (i) A ist die Nullmatrix, oder es gilt $d_1 = d_2$. In diesem Fall ist A ein nichtnegatives Vielfaches einer unitären Matrix.

(ii) π_1, π_2 sind genau dann äquivalent, wenn es eine nichttriviale Verbindungsmatrix gibt.

Beweis: Kor 16.4 in [Be2]. \square

³⁾Das ist offensichtlich eine Verallgemeinerung des Begriffs „Äquivalenz“ für Darstellungen.

Lemma 3.2.4. π_1, π_2 seien wie vorstehend. Für jede $d_2 \times d_1$ -Matrix A verbindet die Matrix

$$\tilde{A} := \frac{1}{\#G} \sum_g \pi_2(g^{-1}) A \pi_1(g)$$

die Darstellungen π_1, π_2 .

Folglich ist \tilde{A} die Nullmatrix, wenn π_1, π_2 nicht äquivalent sind, und andernfalls ist \tilde{A} ein nichtnegatives Vielfaches einer unitären Matrix. Im Fall $\pi_1 = \pi_2$ ist \tilde{A} von der Form $c \text{Id}$, dabei ist $c = \text{tr}(A) / \dim H_1$. (Hier bezeichnet $\text{tr}(A)$ die Spur von A .)

Beweis: Lemma 16.5 in [Be2]. □

Wir bezeichnen wieder mit X_G den Raum der komplexwertigen Funktionen von G nach \mathbb{C} mit dem üblichen Skalarprodukt $(\langle f_1, f_2 \rangle := (\sum_g f_1(g) \overline{f_2(g)}) / \#G)$.

Sei π eine irreduzible unitäre Darstellung auf dem (notwendig endlichdimensionalen) Hilbertraum H . Nach geeigneter Basiswahl kann H mit \mathbb{C}^n und jedes $\pi(g)$ mit einer unitären Matrix A_g^π identifiziert werden. Bezeichne mit $f_{j,k}^\pi : G \rightarrow \mathbb{C}$ die Abbildung, die einem g den (j, k) -Eintrag von A_g^π zuordnet. (Wenn π d_π -dimensional ist, so sind das d_π^2 Funktionen.) Bemerkenswert sind dann die Orthogonalitätsrelationen, die den entsprechenden Relationen für Charaktere entsprechen:

Satz 3.2.5. π, π_1, π_2 seien Darstellungen von G , und π_1, π_2 seien nichtäquivalent.

(i) Ist $(j, k) \neq (j', k')$, so sind $f_{j,k}^\pi, f_{j',k'}^\pi$ orthogonal.

(ii) Sind π und π' nicht äquivalent, so sind $f_{j,k}^\pi, f_{j',k'}^{\pi'}$ für beliebige (j, k, j', k') orthogonal.

(iii) $\langle f_{j,k}^\pi, f_{j,k}^\pi \rangle = 1/d_\pi$ für alle j, k

Beweis: Lemma 16.6 in [Be2]. □

Das vorige Lemma besagt, dass die Funktionen $\sqrt{d_\pi} f_{j,k}^\pi$ ein Orthonormalsystem bilden, wenn π alle irreduziblen Darstellungen und i, j alle Indizes von 1 bis d_π durchläuft. Der folgende Satz zeigt, dass sogar eine Orthonormalbasis vorliegt, und in diesem Sinn spielen die $\sqrt{d_\pi} f_{j,k}^\pi$ die gleiche Rolle wie die Charaktere im vorigen Abschnitt.

Satz 3.2.6. (Peter-Weyl-Theorem) G sei eine endliche Gruppe, \hat{G} sei eine Menge von Darstellungen, die für jede irreduzible Darstellung genau einen Vertreter enthält.

(i) Die $\sqrt{d_\pi} f_{j,k}^\pi$ bilden eine Orthonormalbasis.

(ii) $\sum d_\pi^2 = \#G$.

(iii) Für jedes $f \in X_G$ ist

$$f = \sum_{\pi \in \hat{G}} d_\pi \langle f, f_{j,k}^\pi \rangle f_{j,k}^\pi.$$

(iv) Für jedes g gibt es eine irreduzible Darstellung π mit $\pi(g) \neq \text{Id}$.

Beweis: Theorem 16.11 in [Be1]. □

3.3 Exkurs: Markovketten

In diesem Abschnitt soll gezeigt werden, wie sich die Darstellungstheorie in der Theorie der Markovketten anwenden lässt. Er lehnt sich stark an mein Buch [Be2] an.

Wir beginnen mit einigen Vorbereitungen.

Markovketten

$S = \{s_0, \dots, s_{n-1}\}$ sei eine endliche Menge. Eine *Markovkette auf S* ist gegeben durch

- eine Startverteilung $p = (p_0, \dots, p_{n-1})$ (also $p_i \geq 0$, $\sum_i p_i = 1$);
- eine stochastische Matrix $P = (p_{i,j})_{i,j=0,\dots,n-1}$; es gilt also stets $p_{i,j} \geq 0$ sowie $\sum_j p_{i,j} = 1$ für alle i .

Die Idee: Ein Zufallsspaziergänger startet auf S gemäß p . Der jeweils nächste Schritt wird so festgesetzt: Befindet er sich in i , so geht er mit Wahrscheinlichkeit $p_{i,j}$ nach j .

Der Zusammenhang zur linearen Algebra ist durch die folgende Formel gegeben: Die Wahrscheinlichkeit, den Spaziergänger nach dem n -ten Schritt in i anzutreffen, ist die i -te Komponente des Vektors pP^n .

Ergodische Markovketten

Ergodische Ketten sind solche, bei denen für ein geeignetes n alle Komponenten von P^n strikt positiv sind. Die Interpretation: Egal, wo der Spaziergänger startet, nach n Schritten könnte er überall anzutreffen sein.

Die Gleichgewichtsverteilung

Zu jeder ergodischen Kette gehört eine *Gleichgewichtsverteilung*. Das ist ein stochastischer Zeilenvektor π , für den $\pi P = \pi$ gilt. Das bedeutet: Wird der Spaziergänger mit der Verteilung π gestartet, so ist er immer – nach beliebig vielen Schritten – in den einzelnen Zuständen i mit Wahrscheinlichkeit π_i zu finden.

Die Matrizen P^n konvergieren gegen eine Matrix W , in der alle Zeilen gleich π sind. Aus verschiedenen Gründen ist es interessant zu wissen, wie schnell diese Konvergenz vor sich geht. Das ist ausführlich in meinem Buch [Be2] dargestellt.

Methoden, um die Konvergenzgeschwindigkeit zu berechnen

Der zweite Teil von [Be2] ist der Darstellung von verschiedenen Ansätzen gewidmet, diese Frage zu beantworten.

Zunächst ist zu klären, wie man den Abstand von Wahrscheinlichkeitsmaßen definieren sollte. Der folgende Ansatz wird hier verwendet werden:

Definition 3.3.1. *Es seien P und Q Wahrscheinlichkeitsmaße auf einer endlichen Menge S . Der Variationsabstand zwischen P und Q ist dann definiert durch*

$$\|P - Q\| := \sup_{\Delta \subset S} |P(\Delta) - Q(\Delta)|.$$

Es lässt sich dann zeigen⁴⁾:

Lemma 3.3.2. $\|P - Q\| = (\sum_{i \in S} |P(\{i\}) - Q(\{i\})|)/2$, d.h. der Variationsabstand ist im Wesentlichen der l^1 -Abstand.

Markovketten, die durch Maße auf Gruppen definiert sind

Es sei (G, \circ) eine endliche Gruppe, wir wollen spezielle Markovketten auf G betrachten. Dabei ist ein Wahrscheinlichkeitsmaß \mathbb{P} auf G gegeben, und die Kette ist wie folgt definiert:

- Gestartet wird bei e .
- Befindet man sich bei g , so erzeuge ein zufälliges Element g' gemäß \mathbb{P} . Gehe dann von g zu $g' \circ g$.

Beispiele: Zufallsspaziergänge auf \mathbb{Z}_n , Mischen im Fall von Permutationsgruppen.

Wir wollen annehmen, dass die Kette wirklich mischt, dass also der Träger von \mathbb{P} nicht in einer echten Untergruppe liegt. Dann ist sie ergodisch, und die Gleichgewichtsverteilung ist die Gleichverteilung.

Das Hauptproblem

Interessant ist dann die Frage, wie schnell die Kette gegen die Gleichverteilung konvergiert. Wie oft muss man zum Beispiel mischen, damit alle Kartenpermutationen in etwa die gleiche Wahrscheinlichkeit haben?

Die Wahrscheinlichkeitsverteilung des Spaziergängers nach k Schritten ist das k -fache Faltungsprodukt von \mathbb{P} mit sich. Die Frage ist also, wie schnell der Variationsabstand zur Gleichverteilung gegen Null geht

Der kommutative Fall

Jetzt sei G kommutativ. Wir wollen zeigen, wie man die Frage nach der Konvergenzgeschwindigkeit mit Hilfe der Charaktere beantworten kann.

⁴⁾Beweis in [Be2], Lemma 13.3.

Definition 3.3.3. Sei \mathbb{P} ein Wahrscheinlichkeitsmaß auf G . Unter der Fouriertransformation von \mathbb{P} verstehen wir dann die durch

$$\hat{\mathbb{P}}(\chi) := \sum_g \chi(g) \mathbb{P}(\{g\})$$

definierte Funktion $\hat{\mathbb{P}} : \hat{G} \rightarrow \mathbb{C}$.

Beispiele sind auf Seite 122/123 von [Be2] zu finden.

Lemma 3.3.4. Bezeichne mit U die Gleichverteilung auf G .

(i) $\mathbb{P} = U$ genau dann, wenn $\hat{\mathbb{P}}$ beim trivialen Charakter gleich Eins ist und sonst verschwindet.

(ii) $\|\mathbb{P}_1 - \mathbb{P}_2\| \leq (\sum_{\chi \in \hat{G}} |\hat{\mathbb{P}}_1(\chi) - \hat{\mathbb{P}}_2(\chi)|^2)^{1/2} / 2$.

(iii) Insbesondere ist $\|\mathbb{P} - U\| \leq (\sum_{\chi \neq \chi_{\text{triv}}} |\hat{\mathbb{P}}(\chi)|^2)^{1/2} / 2$.

(iv) Sei \mathbb{P}^{*n} das n -fache Faltungsprodukt von \mathbb{P} mit sich. Dann ist $\widehat{\mathbb{P}^{*n}} = (\hat{\mathbb{P}})^n$.

Beweis: Lemam 15.8, prop. 15.10 in [Be2]. \square

Fasst man alles zusammen, so heißt das: Der Abstand der Verteilung nach n Schritten zur Gleichgewichtsverteilung (das ist hier die Gleichverteilung) ist abschätzbar durch $(\sum_{\chi \neq \chi_{\text{triv}}} |\hat{\mathbb{P}}(\chi)|^{2n})^{1/2} / 2$. Es reicht also, die Zahlen $\hat{\mathbb{P}}(\chi)$ für die nichttrivialen χ auszurechnen. Sie liegen garantiert im komplexen Einheitskreis⁵⁾, und je kleiner sie betragsmäßig sind, umso besser wird die Konvergenz sein.

Die ganze Wahrheit steht in Proposition 15.11 in [Be]: Wenn der Träger von \mathbb{P} nicht in einer Untergruppe gefangen ist, ist die Kette ergodisch, sie konvergiert gegen die Gleichverteilung, und quantitativ ist alles durch die Zahlen $|\hat{\mathbb{P}}(\chi)|$ beschreibbar. Auf den Seiten 138 ff. gibt es konkrete Beispiele.

Der Fall beliebiger endlicher Gruppen

Nun sei (G, \circ) eine beliebige endliche Gruppe. Wir wissen schon, dass G nur endlich viele inäquivalente irreduzible Darstellungen π hat und dass für die Dimensionen $\sum_{\pi} d_{\pi}^2 = \#G$ gilt. Wir fixieren so ein System von irreduziblen Darstellungen: Alle kommen vor, und jede bis auf Äquivalenz nur einmal. Wir nennen es wieder \hat{G} , so eine Auswahl heißt ein *Dual* von G .

Definition 3.3.5. $f : G \rightarrow \mathbb{C}$ sei eine Funktion und \mathbb{P} sei ein Wahrscheinlichkeitsmaß auf G . Die Fouriertransformation von f bzw. \mathbb{P} ist dann eine Familie von Matrizen: Jedem $\pi \in \hat{G}$ wird die Matrix $\hat{f}(\pi)$ bzw. $\hat{\mathbb{P}}(\pi)$ zugeordnet. Dabei ist

$$\hat{f}(\pi) := \frac{1}{\#G} \sum_g f(g) \pi(g), \quad \hat{\mathbb{P}}(\pi) := \sum_g \mathbb{P}(\{g\}) \pi(g).$$

⁵⁾Als Konvexkombinationen von Zahlen auf dem Rand.

Auf Seite 155 in [Be1] wird ein Beispiel berechnet, und in Lemma 16.14 wird gezeigt:

- Die Fouriertransformation der (Konjugierten der) Koordinatenfunktionen der $\pi \in \hat{G}$ können explizit und einfach berechnet werden.
- Die Fouriertransformation der Gleichverteilung U hat bei der trivialen Darstellung den Wert 1, alle anderen $\hat{U}(\pi)$ verschwinden.

In Proposition 16.15 steht die Fourier-Inversionsformel: f bzw. \mathbb{P} kann aus \hat{f} bzw. $\hat{\mathbb{P}}$ leicht zurückgewonnen werden. Es lässt sich auch zeigen, dass die Fouriertransformation Faltungen von Maßen in die Matrixprodukte der Fouriertransformation transformiert. Das ist insbesondere dann interessant, wenn man wissen möchte, nach welchen Wahrscheinlichkeiten ein Zufallsspaziergänger nach n Schritten auf der Gruppe zu finden ist, wenn er sich in jedem Schritt von \mathbb{P} leiten lässt: von g nach $h \circ g$, wo h gemäß \mathbb{P} ausgesucht ist.

Das Analogon zu Lemma 3.3.4 ist

Lemma 3.3.6. *Sei \mathbb{P} ein Wahrscheinlichkeitsmaß auf G .*

(i) \mathbb{P} ist genau dann die Gleichverteilung, wenn $\hat{\mathbb{P}}$ bei allen nichttrivialen Darstellungen verschwindet.

(ii) Für den Variationsabstand zur Gleichverteilung U gilt:

$$\|\mathbb{P} - U\| \leq \frac{1}{2} \sum_{\pi \neq \pi_{\text{triv}}} d_{\pi} \text{tr}(\hat{\mathbb{P}}(\pi) \hat{\mathbb{P}}(\pi)^*),$$

dabei steht tr für die Spur einer Matrix.

Beweis: Vgl. Lemma 16.17 in [Be2]. □

Wieder ist es so, dass die Verteilung des Spaziergängers nach k Schritten durch das k -fache Faltungsprodukt von \mathbb{P} mit sich gegeben ist. Und die Fouriertransformation bildet Faltungen auf (Matrix-)Produkte ab.

Und das bedeutet:

Um zu untersuchen, wie schnell eine Markovkette gegen die Gleichgewichtsverteilung (das ist hier wieder die Gleichverteilung) konvergiert, kann man wie folgt vorgehen:

- Bestimme die nichttrivialen irreduziblen Darstellungen π der Gruppe G .
- Berechne für $k \in \mathbb{N}$ und $g \in G$ die Spur von $(\hat{\mathbb{P}}(\pi))^k (\hat{\mathbb{P}}(\pi)^*)^k$.

Je nachdem, wie schnell diese Zahlen mit $k \rightarrow \infty$ gegen Null gehen, wird die Konvergenz mehr oder weniger schnell sein.

Das kann recht mühsam sein. Deutlich einfacher ist es, wenn die Matrizen $\hat{\mathbb{P}}(\pi)$ selbstadjungiert sind. Dann sind sie o.B.d.A. diagonal, und das Konvergenzverhalten wird durch den Betrag der Eigenwerte determiniert⁶⁾. Die Voraussetzung ist dann erfüllt, wenn $\mathbb{P}(\{g\}) = \mathbb{P}(\{g^{-1}\})$ für alle g gilt, wenn also \mathbb{P} *symmetrisch* ist.

3.4 Kompakte Gruppen

Sei (G, \circ) eine kompakte topologiesche Gruppe. Wir wissen dann schon:

- G ist unimodular.
- Jede irreduzible Darstellung ist endlich-dimensional.

Um G analysieren zu können, benötigen wir das folgende Theorem. Es besagt, dass es – auch im Fall beliebiger lokalkompakter Gruppen – „genügend viele“ irreduzible Darstellungen gibt:

Theorem 3.4.1. (*Theorem von Gelfand-Raikov, 1940*) *Es sei (G, \circ) eine lokalkompakte Gruppe. Für jedes $g \neq e$ gibt es dann eine irreduzible Darstellung π mit $\pi(g) \neq Id$.*

Beweis: Der Beweis ist überraschend schwierig. Wir werden das Ergebnis hier als „black box“ verwenden und nur eine Beweisidee präsentieren. Hier sind die wichtigsten Schritte:

- Eine Banach- $*$ -Algebra ist eine komplexe Banachalgebra mit Involution $*$. Beispiele: $C_{\mathbb{C}}K$, alle Operatoren auf einem Hilbertraum oder die Gruppenalgebra einer lokalkompakten topologischen Gruppe. Der Begriff „Darstellung“ wird in naheliegender Weise definiert.
- Ist A eine Banach- $*$ -Algebra, so soll P_A die Menge derjenigen stetigen linearen Funktionale p bezeichnen, für die gilt:
 - Für alle $x \in A$ ist $p(x^*x) \geq 0$.
 - Für alle x ist $p(x^*) = \overline{p(x)}$.
 - Für alle x ist $|p(x)|^2 \leq p(x^*x)$.
- Wichtig ist dann die Aussage: Wenn es für ein $x \in A$ ein $p \in P_A$ mit $p(x) \neq 0$ gibt, so gibt es auch eine irreduzible Darstellung π von A mit $\pi(x) \neq 0$. (Die Hauptidee dabei: Jedes p liefert eine Darstellung π , das ist die so genannte GNS-Konstruktion.)
- Das wird wie folgt zusammengeführt: Die relevante Banach- $*$ -Algebra ist die Gruppenalgebra von G ; die Existenz geeigneter p folgt aus einer Kombination der Sätze von Hahn-Banach und Krein-Milman; Darstellungen der Gruppenalgebra liefern Darstellungen der Gruppe.

⁶⁾Beachte: Auch die k -te Potenz ist diagonal, und auf der Diagonale stehen die k -ten Potenzen der Eigenwerte.

(Einzelheiten im Buch von Hewitt-Ross.) \square

Etwas vereinfacht lässt sich dann sagen, dass die Ergebnisse aus Abschnitt 3.2 auf den Fall kompakter Gruppen übertragen werden können:

1. G sei kompakt. \hat{G} enthalte aus jeder irreduziblen Darstellung einen Vertreter.
2. μ sei ein Haarsches Maß auf G . Der $L^2(G, \mu)$ ist dann ein Hilbertraum. Es gelten dann die Orthogonalitätsrelationen:

- Sind π_1, π_2 nicht äquivalent, so sind Funktionen des Typs

$$g \mapsto \langle x, \pi_1(g)y \rangle, \quad g \mapsto \langle x', \pi_2(g)y' \rangle$$

orthogonal.

- Sei $\pi \in \hat{G}$. Damit ist π eine Darstellung auf einem endlich-dimensionalen Hilbertraum, die $\pi(g)$ sind also Matrizen $(f_{i,j}^\pi)_{i,j=1,\dots,d_\pi}$.

Damit gibt π Anlass zu den $(d_\pi)^2$ stetigen Funktionen $g \mapsto f_{i,j}^\pi$. Diese Funktionen sind paarweise orthogonal, ihre L^2 -Norm ist $1/d_\pi$.

- Die Menge der $\sqrt{d_\pi} f_{i,j}^\pi$ (alle $\pi \in \hat{G}$, jeweils alle Koordinatenfunktionen) ist eine Orthonormalbasis des $L^2(G, \mu)$. Das ist das *Peter-Weyl-Theorem* für kompakte Gruppen⁷⁾.
- Für jedes $f \in L^2$ ist damit

$$f = \sum_{\pi} \sum_{i,j=1,\dots,d_\pi} d_\pi \langle f, f_{i,j}^\pi \rangle f_{i,j}^\pi.$$

Dabei ist die Summe im L^2 -Sinn zu verstehen.

3.5 Kommutative lokalkompakte Gruppen

In diesem Abschnitt sei $(G, +)$ eine abelsche lokalkompakte Gruppe. Wie bisher sei \hat{G} die Menge der Charaktere. Aufgrund des Satzes von Gelfand-Raikov wissen wir, dass \hat{G} die Punkte trennt.

\hat{G} ist eine abelsche Gruppe. In diesem Abschnitt wollen wir uns um die folgenden Sachverhalte kümmern:

- \hat{G} kann in natürlicher Weise wieder zu einer lokalkompakten Gruppe gemacht werden. Folglich kann man $\hat{\hat{G}}$ – die duale Gruppe der dualen Gruppe bilden.
- G und $\hat{\hat{G}}$ sind isomorph. Genauer: Die naheliegende Abbildung $g \mapsto (\chi \mapsto \chi(g))$ ist ein Isomorphismus topologischer Gruppen. (Kommutative Gruppen verhalten sich also so wie reflexive Banachräume.) Das ist der *Dualitätssatz von Pontryagin/van Kampen*.

⁷⁾Beachte: Im Fall $d_\pi > 1$ sind die $f_{i,j}^\pi$ nicht eindeutig festgelegt.

- Es gibt ein Analogon zur Fouriertransformation.

Außerdem sollen einige Folgerungen aus diesen Ergebnissen besprochen werden.

 \hat{G} als topologische Gruppe

Wir versehen \hat{G} mit der Topologie der gleichmäßigen Konvergenz auf kompakten Teilmengen von G . Da wir Topologie hier nicht voraussetzen wollten, setzen wir die Untersuchungen nur für eine Teilklasse fort:

- G ist durch eine Metrik δ topologisiert.
- Es gibt kompakte Teilmengen K_m von G mit der folgenden Eigenschaft: Es ist $K_1 \subset K_2 \subset \dots$, und jede kompakte Teilmenge von G ist in einem K_m enthalten. Für separable lokalkompakte metrisierbare Gruppen geht das immer.

Für zwei beschränkte $f_1, f_2 : G \rightarrow \mathbb{C}$ sei dann

$$d_m(f_1, f_2) := \max_{g \in K_m} |f_1(g) - f_2(g)|,$$

und $d(f_1, f_2)$ wird durch

$$d(f_1, f_2) := \sum_m \frac{d_m(f_1, f_2)}{2^m}$$

erklärt.

Dann gilt:

Satz 3.5.1. *Bezüglich dieser Metrik ist (\hat{G}, \cdot) eine lokalkompakte abelsche Gruppe.*

Beweis: Man mache sich zunächst klar, dass $\chi_n \rightarrow \chi$ gleichwertig dazu ist, dass die Folge (χ_n) auf jedem K_m gleichmäßig gegen χ konvergiert. Damit ist schnell einzusehen, dass eine topologische Gruppe vorliegt⁸⁾.

Es ist allerdings nicht ganz offensichtlich, dass die Topologie lokalkompakt ist. Wir fixieren eine kompakte Umgebung K von 0 in G und nehmen an, dass ein $\chi \in \hat{G}$ so gegeben ist, dass $|1 - \chi(g)| \leq \eta$ für $g \in K$; dabei ist $0 < \eta \leq 1/4$.

Sei U eine (o.B.d.A.) kompakte Nullumgebung mit $U + U \subset K$. Ist dann $g \in U$, so sind $g, g + g \in K$, d.h. $|1 - \chi(g)|, |1 - \chi(g)^2| \leq \eta$. Das impliziert $|1 - \chi(g)| \leq (2/3)\eta$.

Wenn man das iteriert, folgt: Zu jedem k gibt es eine Nullumgebung U mit $|1 - \chi(g)| \leq (2/3)^k \eta$ für $g \in U$. Und damit ist klar, dass die Menge der χ für die $|1 - \chi(g)| \leq \eta$ auf K gilt, gleichmäßig gleichgradig stetig sein muss.

⁸⁾Zur Übung sollte man sich davon überzeugen, dass in den Fällen $\Gamma, \mathbb{Z}, \mathbb{R}$ die „richtige“ Topologie herauskommt.

Zusammen mit dem Satz von Arzela-Ascoli folgt, dass jede Folge in

$$\Delta := \{\chi \mid |1 - \chi(g)| \leq 1/4 \text{ für } g \in K\}$$

eine in \hat{G} konvergente Teilfolge enthält und folglich kompakt ist. Setzt man speziell $K = K_1$, so heißt das, dass die Kugel um den Charakter 1 mit dem Radius $1/4$ kompakt ist. \square

Es gibt zwei Extremfälle, die den Beispielen \mathbb{Z} und Γ entsprechen:

Satz 3.5.2. (i) Ist G kompakt, so ist \hat{G} diskret.

(ii) Ist G diskret, so ist \hat{G} kompakt.

Beweis: (i) Es sei χ ein Charakter mit $\|\chi - 1\| \leq 1/4$. Notwendig ist dann $\chi = 1$, d.h. es gibt einpunktige Kugeln.

(ii) Die Menge aller Funktionen von G nach Γ , also G^Γ ist in der Produkttopologie nach dem Satz von Tychonoff kompakt und \hat{G} ist darin eine abgeschlossene Menge. In unserem Fall ist die Topologie gerade die Topologie der punktweisen Konvergenz. \square

Die natürliche Einbettung $i_G : G \rightarrow \hat{G}$

In Analogie zur Funktionalanalysis gibt es eine natürliche Abbildung in das biduale Objekt.

Definition 3.5.3. Unter $i_G : G \rightarrow \hat{G}$ verstehen wir die Abbildung

$$g \mapsto (\chi \mapsto \chi(g)).$$

Lemma 3.5.4. (i) i_G ist wohldefiniert, d.h. jedes $i_G(g)$ liegt in \hat{G} .

(ii) i_G ist injektiv.

(iii) i_G ist ein stetiger und offener Gruppenhomomorphismus.

Beweis: (i) Es ist klar, dass $i_G(g)$ ein Gruppenmorphismus ist. Die Stetigkeit folgt daraus, dass $\{g\}$ kompakt ist.

(ii) Die Injektivität gilt deswegen, weil die Charaktere die Punkte trennen.

(iii) Klar ist, dass stets $i_G(g + g') = i_G(g) \cdot i_G(g')$ gilt.

i_G ist stetig. Zu zeigen ist doch: Aus $g_n \rightarrow 0$ folgt $i_G(g_n) \rightarrow 1$, wobei 1 der Eins-Charakter auf \hat{G} ist. Das folgt aus der Stetigkeit von $(\chi, g) \mapsto \chi(g)$ durch ein typisches Kompaktheitsargument. (Ein Beispiel dazu aus der Analysis: $\phi : [0, 1] \times \mathbb{R} \rightarrow \mathbb{R}$ sei stetig. Zu $\varepsilon > 0$ gibt es dann ein $\delta > 0$ mit der folgenden Eigenschaft: Ist $x \in [0, 1]$ beliebig $|y| \leq \delta$, so ist $|\phi(x, y) - \phi(x, 0)| \leq \varepsilon$.)

Dass i_G offen ist, ist etwas komplizierter einzusehen. Man muss zeigen: Ist U eine Nullumgebung in G , so gibt es ein kompaktes $K \subset \hat{G}$ und ein $\varepsilon > 0$, so dass für jedes $g \notin U$ ein $\chi \in K$ mit $|1 - \chi(g)| > \varepsilon$ existiert. Dazu müsste man

eine etwas schärfere Version der Tatsache herleiten, dass für $g \neq 0$ ein χ mit $\chi(g) \neq 1$ existiert. \square

Der Satz von Pontryagin/van Kampen

Im vorigen Abschnitt haben wir gesehen, dass G (bis auf Isomorphie) als Untergruppe von $\hat{\hat{G}}$ aufgefasst werden kann. Die Gruppen sind sogar gleich:

Theorem 3.5.5. (Satz von Pontryagin/van Kampen) i_G ist surjektiv und damit ein Gruppenisomorphismus.

Beweis: Zum Beweis gibt es zwei Beweisstrategien. Die eine beruht auf Funktionalanalysis (vgl. Loomis oder Heyer). Die andere beginnt mit dem Nachweis für Spezialfälle und nutzt dann Strukturaussagen für lokalkompakte abelsche Gruppen: Wenn sie kompakt erzeugt sind, sind sie ein Produkt aus kompakter abelscher Untergruppe, \mathbb{Z}^n und \mathbb{R}^m . Und da jedes $g \in G$ in einer kompakt erzeugten Untergruppe liegt, reicht es, den Satz für solche Gruppen zu beweisen. (Hewitt-Ross I, Pontryagin). Hier zeigen wir nur die ersten Schritte der zweiten Strategie.

Die Aussage stimmt für diskrete Gruppen. Als Baustein zum Beweis verwenden wir den

Satz von Stone-Weierstraß: Es sei K ein kompakter topologischer Raum, und $A \subset C_{\mathbb{C}} K$ sei ein Unterraum mit den folgenden Eigenschaften.

- Mit $f_1, f_2 \in A$ sind auch $\overline{f_1}$ und $f_1 f_2$ in A .
- Die Einsfunktion gehört zu A .
- A trennt Punkte: Zu $x, y \in K$ mit $x \neq y$ gibt es ein $f \in A$ mit $f(x) \neq f(y)$.

(In Kurzfassung: A ist eine punkt-trennende selbstadjungierte Unteralgebra mit Einheit.) Dann liegt A dicht bezüglich der Maximumsnorm.

Hier wenden wir den Satz so an. Wir wissen schon, dass \hat{G} kompakt ist, das ist hier das K . Als A definieren wir die lineare Hülle der Charaktere $i_G(g)$ auf \hat{G} . Das ist eine punkt-trennende selbstadjungierte Unteralgebra mit Einheit, A liegt folglich dicht. Sei nun $\chi' \in \hat{\hat{G}}$. Angenommen, es wäre χ' nicht in $i_G(G)$. Dann wäre χ' orthogonal zu A und folglich gleich Null. Das ist ein Widerspruch.

Die Aussage stimmt für kompakte Gruppen. Als Vorbereitung folgern wir aus der Tatsache, dass es „viele“ Charaktere gibt, die folgende Aussage: Ist H eine echte abgeschlossene Untergruppe der lokalkompakten abelschen Gruppe G , so gibt es einen nichttrivialen Charakter auf G , der auf H gleich Eins ist. (Man muss nur einen nichttrivialen Charakter auf G/H finden und ihn mit der Quotientenabbildung verknüpfen.)

Mal angenommen, es wäre $H := i_G(G)$ eine echte Untergruppe von \hat{G} . Da G kompakt und i_G stetig ist, ist H abgeschlossen. Wähle einen nichttrivialen Charakter χ' auf \hat{G} , der auf H gleich 1 ist. Da \hat{G} diskret ist, wissen wir nach dem ersten Beweisteil, dass χ' die Auswertung bei einem geeigneten $\chi \in \hat{G}$ ist. Anders ausgedrückt: Es gibt ein nichttriviales $\chi \in \hat{G}$, so dass $\chi(g) = 1$ für alle g ist. Das ist offensichtlich ein Widerspruch. \square

Fouriertransformation und inverse Fouriertransformation

Jetzt sollen noch einige Definitionen und Ergebnisse der „klassischen“ Fouriertransformation übertragen werden.

Definition 3.5.6. G sei abelsch und lokalkompakt, und μ sei ein Haarmaß. Für $f \in L^1(G, \mu)$ definieren wir $\hat{f} : \hat{G} \rightarrow \mathbb{C}$ durch

$$\hat{f}(\chi) := \int_G f \bar{\chi} d\mu.$$

\hat{f} heißt die Fouriertransformation von f .

Bemerkungen: 1. Die Definition hängt von der Wahl von μ ab und ist folglich nur bis auf einen positiven Faktor eindeutig.

2. Mit gleichem Recht hätte man $\hat{f}(\chi)$ als $\int_G f \chi d\mu$ definieren können.

3. Es ist klar, dass die üblichen Linearitätseigenschaften gelten und dass es Formeln für die Fouriertransformation der Funktionen $\chi_0 f, f_g, \bar{f}$ gibt.

4. \hat{f} ist eine beschränkte stetige Funktion, die im Unendlichen verschwindet.

5. Im Fall $G = \mathbb{Z}$ ist die Fouriertransformation so definiert: Ist $f = (a_n)_{n \in \mathbb{Z}} \in l^1(\mathbb{Z})$, so ist $\hat{f} : \Gamma \rightarrow \mathbb{C}$ die Abbildung $z \mapsto \sum_n a_n z^n$.

6. Nun sei $G = \Gamma$. Für $f \in L^1(\Gamma)$ besteht $\hat{f} : \mathbb{Z} \rightarrow \mathbb{C}$ aus der Folge der Fourierkoeffizienten von f .

7. Für $G = \mathbb{R}$ ergibt sich (bis auf einen Faktor) die klassische Fouriertransformation.

Wie im klassischen Fall gibt es nun das Problem, dass man einerseits im $L^2(g)$ arbeiten möchte, um die Orthogonalitätsrelationen ausnutzen zu können, dass aber andererseits \hat{f} nur für $f \in L^1$ definiert ist. Schlüssel zur Verallgemeinerung ist wieder ein Planchereltheorem: Man kann ein Haarmaß $\hat{\mu}$ auf \hat{G} so wählen, dass für $f \in L^1 \cap L^2$ die Formel

$$\int_{\hat{G}} |\hat{f}(\chi)|^2 d\hat{\mu}(\chi) = \int_G |f(g)|^2 d\mu(g)$$

gilt. Damit erhält man wieder eine Fouriertransformation für L^2 -Funktionen. Und dann ist $f \mapsto \hat{f}$ ein isometrischer Isomorphismus von $L^2(G)$ nach $L^2(\hat{G})$. Die inverse Abbildung ist durch

$$f(g) = \int_{\hat{G}} \hat{f}(\chi) \chi(g) d\hat{\mu}(\chi)$$

(fast überall) gegeben. Das ist die *inverse Fouriertransformation*.

Bochners Theorem

Mit Hilfe der dualen Gruppe lässt sich ein wichtiger Struktursatz beweisen: Wie genau sehen positiv definite Funktionen aus? Zunächst die Definition.

Definition 3.5.7. G sei abelsch und lokalkompakt. Ein $\phi : G \rightarrow \mathbb{C}$ heißt positiv definit, wenn

$$\sum_{j,k=1,\dots,m} \alpha_k \bar{\alpha}_j \phi(g_k - g_j)$$

für alle $\alpha_1, \dots, \alpha_m \in \mathbb{C}$ und $g_1, \dots, g_m \in G$ reell und ≥ 0 ist. Anders ausgedrückt: Stets ist die Matrix $(\phi(g_k - g_l))_{k,l}$ positiv definit.

Bemerkungen: 1. Charaktere sind positiv definit, da

$$\sum_{j,k=1,\dots,m} \chi_k \bar{\alpha}_j \chi(g_k - g_j) = \left(\sum_k \alpha_k \chi(g_k) \right) \overline{\left(\sum_j \alpha_j \chi(g_j) \right)}.$$

2. Punktweise Limites und positive Linearkombinationen sind es wieder. Es geht sogar etwas allgemeiner: Ist η ein endliches positives reguläres Borelmaß auf \hat{G} , so ist

$$g \mapsto \int_{\hat{G}} \chi(g) d\eta(\chi)$$

stetig und positiv definit.

Bemerkenswerte Weise sind das schon alle solche Funktionen:

Theorem 3.5.8. (Bochners Theorem) $\phi : G \rightarrow \mathbb{C}$ sei stetig und positiv definit. Dann gibt es ein endliches positives reguläres Borelmaß η auf \hat{G} , so dass

$$\phi(g) = \int_{\hat{G}} \chi(g) d\eta(\chi)$$

für alle g gilt.

Beweis: Hier ist die Beweisidee. ϕ sei gegeben. Wir betrachten den Bildraum H der Fouriertransformation, also die Menge aller \hat{f} für $f \in L^1$. Das ist ein Unterraum der stetigen Funktionen auf \hat{G} , die im Unendlichen verschwinden. Erstens liegt dann H dicht (es ist eine punktetrennende Unter algebra, die an keiner Stelle verschwindet), und $Q = H \rightarrow \mathbb{C}$, $\hat{f} \mapsto \int_G f \phi$ ist eine wohldefinierte stetige lineare Abbildung.

Q lässt sich also stetig auf die im Unendlichen verschwindenden Funktionen fortsetzen und damit als Maß schreiben. Es gibt also ein Maß η' , so dass

$$\int_G f(g) \phi(g) d\mu(g) = \int_{\hat{G}} \hat{f}(\chi) d\eta'(\chi)$$

gilt.

Sei η das Maß $\eta(U) := \eta'(-U)$. Definiert man die positiv definite Funktion ϕ_η durch $g \mapsto \int_{\hat{G}} \chi(g) d\eta(\chi)$, so kann man mit dem Satz von Fubini zeigen, dass

$$\int_G f(g)\phi(g) d\mu(g) = \int_G \phi_\eta(g)f(g) d\mu(g)$$

für alle L^1 -Funktionen f gilt. Deswegen muss $\phi = \phi_\eta$ fast überall sein, aus Stetigkeitsgründen stimmen beide Funktionen deswegen überein. \square

Die Struktur unitärer Darstellungen von G (Satz von Stone)

In Kurzfassung besagt der folgende Satz, dass unitäre Darstellungen aus Charakteren aufgebaut sind. Die ausführliche Formulierung beginnt mit einer Definition: Was bedeutet es, dass eine Darstellung „nicht zu groß“ ist?

Definition 3.5.9. *Eine unitäre Darstellung π von G auf H heißt zyklisch, wenn es ein $x_0 \in H$ so gibt, dass die lineare Hülle der Menge $\{\pi(g)x_0 \mid g \in G\}$ dicht liegt.*

Mit dem Zornschen Lemma ist leicht einzusehen, dass jede unitäre Darstellung direkte Summe zyklischer Darstellungen ist. Es reicht also, solche Darstellungen zu charakterisieren.

Theorem 3.5.10. (Stone) (i) *Sei η ein endliches positives reguläres Borelmaß auf \hat{G} . Mit H bezeichnen wir den $L^2(G, \eta)$. Eine Darstellung π_η wird so definiert: Für $g \in G$ und $h \in H$ ist $\pi_\eta(g)h$ die Funktion $\chi \mapsto \chi(g)h(g)$.*

Dann ist π_η eine zyklische unitäre Darstellung von G .

(ii) *Jede zyklische unitäre Darstellung ist äquivalent zu einer des vorstehend beschriebenen Typs.*

Auf den Beweis können wir hier nicht eingehen. Der erste Teil ist unproblematisch⁹⁾, der zweite wird auf Bohners Theorem zurückgeführt: Ist π zyklisch mit zyklischem Vektor x_0 , so betrachte die Abbildung $\phi : G \rightarrow \mathbb{C}$,

$$\phi(g) := \langle \pi(g)x_0, x_0 \rangle.$$

Das nach Bohners Theorem existierende η leistet dann alles.

Und die allgemeine Version von Satz 3.1.5 liest sich wie folgt. Vorher muss man wissen, was ein *projektionswertiges Maß* ist. Sei dazu H ein Hilbertraum und $\Pi(H)$ die Menge der orthogonalen Projektionen auf H . Ist (M, \mathcal{A}) ein Messraum, so ist ein projektionswertiges Maß auf (M, \mathcal{A}) eine Abbildung $\psi : \mathcal{A} \rightarrow \Pi(H)$ mit den folgenden Eigenschaften:

- $\psi(\emptyset) = 0$, $\psi(M) = \text{Id}$.

⁹⁾Ist zum Beispiel η das Punktmaß bei χ , so wird χ erzeugt.

- Sind A_1, A_2, \dots disjunkt in \mathcal{A} , so ist ψ von der Vereinigung das Supremum der $\psi(A_n)$.

Für $x, y \in H$ ist dann $A \mapsto \langle \psi(A)x, y \rangle$ ein signiertes Maß auf M , so dass man für $R : M \rightarrow \mathbb{C}$ das Integral $\int_M R(m) \langle \psi(dm)x, y \rangle$ definieren kann.

Und hier die „konkrete“ Beschreibung von Darstellungen:

Theorem 3.5.11. *Es sei π eine unitäre Darstellung von G auf einem Hilbertraum H . Dann gibt es ein auf den Borelmengen von \hat{G} definiertes projektionswertiges Maß ψ mit Werten in $\Pi(H)$ mit der folgenden Eigenschaft: Für alle $g \in G$ und $x, y \in H$ ist*

$$\langle \pi(g)x, y \rangle = \int_{\hat{G}} \langle \chi(g) d\psi(\chi)x, y \rangle.$$

Beweis: Einen Beweis findet man zum Beispiel im Buch von Loomis, Paragraph 36. \square

3.6 Liegruppen

In vielen Fällen hat die Gruppe, um die es geht, auch noch eine Differenzierbarkeitsstruktur. Zum Kennenlernen empfiehlt es sich, die wichtigsten Begriffe an dem Spezialfall zu studieren, dass die Gruppe als Matrixgruppe realisiert werden kann.

3.6.1 Matrix-Liegruppen

Sei $n \in \mathbb{N}$ und $GL(n, \mathbb{C})$ die Gruppe der invertierbaren $n \times n$ -Matrizen über \mathbb{C} . Unter einer *Matrix-Liegruppe* L verstehen wir eine abgeschlossene Untergruppe von $GL(n, \mathbb{C})$. (Achtung: L muss nicht in den $n \times n$ -Matrizen abgeschlossen sein.)

Hier die wichtigsten Beispiele:

- $GL(n, \mathbb{R})$: die invertierbaren Matrizen mit reellen Einträgen.
- $SL(n, \mathbb{C})$ und $SL(n, \mathbb{R})$: Die Determinante ist jeweils gleich Eins.
- $O(n)$: die reellen orthogonalen Matrizen.
- $SO(n)$: dto, mit Determinante Eins.
- $U(n)$: die komplexen unitären Matrizen.
- $SU(n)$: dto, mit Determinante Eins.

Es gibt noch viele weitere wichtige Beispiele: die symplektischen Gruppen, die Heisenberggruppe, die Poincarégruppe, ...

Alle diese Gruppen sind lokalkompakt, und unter den obigen Beispielen sind nur die $GL(n, \mathbb{C})$, $GL(n, \mathbb{R})$, $SL(n, \mathbb{C})$, $SL(n, \mathbb{R})$ nicht kompakt. Für $n > 1$ sind $GL(n, \mathbb{C})$, $GL(n, \mathbb{R})$, $SL(n, \mathbb{C})$, $SL(n, \mathbb{R})$ nicht kommutativ.

3.6.2 Liegruppen

Für einige fortgeschrittene Strukturuntersuchungen ist der Ansatz zu speziell. Die „richtige“ Definition sieht so aus: Eine *Liegruppe* ist eine differenzierbare Mannigfaltigkeit, die gleichzeitig eine Gruppe ist und für die die Abbildungen $(g, g') \mapsto g \circ g'$ und $g \mapsto g^{-1}$ differenzierbar sind.

Der Vorteil dieses Ansatzes ist, dass man keine speziellen Koordinaten auszeichnen muss. Der schwerwiegende Nachteil ist allerdings, dass man sich in der Theorie der Mannigfaltigkeiten auskennen muss, um Liegruppen untersuchen zu können. Wir bleiben hier bei Matrix-Liegruppen.

3.6.3 Die Exponentialabbildung

Hier wollen wir einige Begriffe und Tatsachen zum Raum der $n \times n$ -Matrizen zusammenstellen, die in der linearen Algebra in der Regel nicht behandelt werden.

Normen

Der \mathbb{R}^n sei mit der euklidischen Norm versehen, es ist also $\|(x_1, \dots, x_n)\| = \sqrt{x_1^2 + \dots + x_n^2}$. Wenn A eine reelle $n \times n$ -Matrix ist, so ist $x \mapsto \|Ax\|$ eine stetige Abbildung von \mathbb{R}^n nach \mathbb{R} , und deswegen existiert die Zahl $\|A\| := \max\{\|Ax\| \mid \|x\| \leq 1\}$ in \mathbb{R} . (Man muss nur beachten, dass $\{x \mid \|x\| \leq 1\}$ als abgeschlossene und beschränkte Menge kompakt ist.)

$\|A\|$ heißt die *Operatornorm* von A , es handelt sich um den Spezialfall einer Norm für lineare Abbildungen zwischen normierten Räumen. Die Abbildung $A \mapsto \|A\|$ hat wirklich die Eigenschaften einer Norm:

- Es ist $\|A\| \geq 0$, und $\|A\| = 0$ gilt genau dann, wenn A die Nullmatrix ist.
- $\|\lambda A\| = |\lambda| \|A\|$ für reelle λ .
- $\|A + B\| \leq \|A\| + \|B\|$.

Diese Aussagen folgen leicht aus den entsprechenden Ergebnissen für die euklidische Norm.

Für konkret gegebenes A ist $\|A\|$ nicht direkt aus den Einträgen zu berechnen, doch das ist hier auch gar nicht erforderlich¹⁰⁾.

Da eine Norm vorliegt, stehen auch alle Begriffe und Ergebnisse zur Verfügung, die man für metrische Räume kennt:

- Durch $d(A, B) := \|A - B\|$ wird eine *Metrik* definiert.
- Man sagt, dass $\lim_{k \rightarrow \infty} A_k = A$ gilt, wenn $\lim \|A_k - A\| = 0$. Die Folge (A_k) heißt dann gegen A *konvergent*.

¹⁰⁾ $\|A\|$ ist das Maximum der Zahlen $\sqrt{\lambda}$, wobei λ die Eigenwerte der symmetrischen positiv definiten Matrix $A^T A$ durchläuft; dabei ist A^T die zu A transponierte Matrix.

- Eine Folge (A_k) von Matrizen heißt eine *Cauchy-Folge*, wenn für jedes $\varepsilon > 0$ ein k_0 existiert, so dass $\|A_k - A_l\| \leq \varepsilon$ für $k, l \geq k_0$.
Konvergente Folgen sind Cauchy-Folgen, bemerkenswerterweise gilt im Raum der $n \times n$ -Matrizen (wie in \mathbb{R} auch) die Umkehrung. Dieser Raum ist also *vollständig*.
- Eine Funktion $\phi : t \mapsto A_t$ heißt *stetig*, wenn $t_k \rightarrow t$ stets $A_{t_k} \rightarrow A_t$ impliziert.
- Eine Funktion $\phi : t \mapsto A_t$ heißt bei t_0 *differenzierbar* mit Ableitung B , wenn aus $t_k \rightarrow t_0$ stets $(A_{t_k} - A_{t_0})/(t_k - t_0) \rightarrow B$ folgt.

Reihen

Es seien A_k $n \times n$ -Matrizen. Man sagt, dass die Reihe $\sum_{k=1}^{\infty} A_k$ konvergiert, wenn die Folge der Partialsummen, also die Folge $(A_1 + \dots + A_k)_k$ konvergent ist.

Satz 1: *Es gebe Zahlen $b_k \geq 0$, so dass $\sum_k b_k < \infty$ und $\|A_k\| \leq b_k$ für jedes k . Dann existiert $\sum_k A_k$.*

Beweis: Die Folge der Partialsummen der A_k ist eine Cauchy-Folge. Für $k < l$ ist nämlich

$$\begin{aligned} \|(A_1 + \dots + A_l) - (A_1 + \dots + A_k)\| &= \|A_{k+1} + \dots + A_l\| \\ &\leq \|A_{k+1}\| + \dots + \|A_l\| \\ &\leq b_{k+1} + \dots + b_l. \end{aligned}$$

Dabei wird die rechte Seite wegen $\sum_k b_k < \infty$ beliebig klein.

Da Cauchy-Folgen von $n \times n$ -Matrizen konvergent sind, ist die Behauptung bewiesen. \square

Die Exponentialfunktion für Matrizen

Ist $x \in \mathbb{R}^n \setminus \{0\}$ so ist die Norm von $x/\|x\|$ gleich Eins, nach Definition von $\|A\|$ gilt also $\|A(x/\|x\|)\| \leq \|A\|$. Das kann zu $\|Ax\| \leq \|A\| \|x\|$ umgeformt werden.

Nun sei $\|x\| \leq 1$. Es ist

$$\|ABx\| \leq \|A\| \|Bx\| \leq \|A\| \|B\| \|x\| \leq \|A\| \|B\|,$$

und das beweist, dass stets $\|AB\| \leq \|A\| \|B\|$ gilt. Durch Induktion kann man das auf mehr als zwei Matrizen verallgemeinern, insbesondere gilt $\|A^k\| \leq \|A\|^k$.

Das hat eine wichtige Konsequenz. Es folgt nämlich, dass $\|A^k/k!\| \leq \|A\|^k/k!$ gilt, und aus Satz 1 können wir schließen, dass die Reihe $Id + A + A^2/2! + A^3/3! + \dots$ konvergent ist. Wir kürzen – in Anlehnung an den entsprechenden reellen Fall – die Reihensumme mit e^A ab. Es gilt dann:

Satz 2: (i) Ist A eine Diagonalmatrix mit Einträgen $\lambda_1, \dots, \lambda_n$, so ist e^A auch diagonal, die Einträge sind $e^{\lambda_1}, \dots, e^{\lambda_n}$.

(ii) Wenn A und B kommutieren, gilt $e^{A+B} = e^A e^B$.

(iii) e^A ist stets invertierbar.

(iv) e^A ist der Limes der $(Id + A/n)^n$.

Beweis: (i) ist offensichtlich, und für den Beweis von (ii) muss man nur den Beweis aus der Analysis für $\exp(x+y) = \exp(x)\exp(y)$ kopieren. Die Kommutativität wird verwendet, um $(A+B)^n$ mit der üblichen binomischen Formel ausrechnen zu können.

(iii) folgt aus (ii) und der Tatsache, dass A mit $-A$ vertauscht. Deswegen ist $Id = e^0 = e^{A-A} = e^A e^{-A}$; hier bezeichnet 0 die Nullmatrix. Und damit ist gezeigt, dass e^{-A} die inverse Matrix zu e^A ist.

Auch für den Beweis von (iv) muss man nur den entsprechenden Beweis aus der Analysis übertragen. \square

In der Vorlesung „Stochastische Differentialgleichungen“ (Kapitel über Markovprozesse in stetiger Zeit) spielte das auch eine wesentliche Rolle. Dort wurde folgendes Ergebnis gebraucht: Wenn $t \mapsto A_t$ (definiert für $t \geq 0$) eine stetige Abbildung ist, die jedem t eine invertierbare Matrix zuordnet, so dass stets $A_{s+t} = A_s A_t$ gilt, so gibt es eine Matrix Q , so dass stets $A_t = e^{tQ}$ gilt. Man spricht von *stetigen Halbgruppen von Matrizen*.

3.6.4 Lie-Algebren

Kennt man eine Liegruppe in der Nähe der Identität, so kann man die zugehörige Lie-Algebra ausrechnen. Das ist ein algebraisches Objekt (keine Topologie mehr!), an dem alle wichtigen Aspekte der Liegruppe mit rein algebraischen Methoden untersucht werden können: Untergruppen, Darstellungen usw.

Hier die wichtige Definition:

Definition 3.6.1. Eine reelle bzw. komplexe Lie-Algebra \mathcal{L} ist ein endlich-dimensionaler reeller bzw. komplexer Vektorraum, auf dem zusätzlich eine bilineare Abbildung $[\cdot, \cdot] : \mathcal{L} \times \mathcal{L} \rightarrow \mathcal{L}$ (die Lieklammer) mit den folgenden Eigenschaften definiert ist:

- Stets ist $[X, Y] = -[Y, X]$.
- Stets gilt die folgende Jacobi-Identität:

$$[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0.$$

Beispiele: 1. \mathcal{L} sei eine Algebra. Mit $[X, Y] := XY - YX$ wird \mathcal{L} zu einer Lie-Algebra.

2. Der \mathbb{R}^3 wird durch das äußere Produkt zu einer Lie-Algebra.

3.6.5 Die Lie-Algebra einer Liegruppe

Zu jeder Liegruppe gehört eine Lie-Algebra:

Definition 3.6.2. Sei L eine Matrix-Liegruppe. Unter der zugehörigen Lie-Algebra \mathcal{L}_L verstehen wir die Menge aller $n \times n$ -Matrizen X , so dass e^{tX} für alle $t \in \mathbb{R}$ zu L gehört. (Das heißt gerade, dass die von X erzeugte ein-Parameter-Untergruppe eine Untergruppe von L ist.)

Die Lie-Klammer wird durch $[X, Y] := XY - YX$ definiert.

Es wird sich zeigen, dass das eine reelle Lie-Algebra ist.

Klar ist eigentlich nur, dass $0 \in \mathcal{L}_L$ und dass mit $X \in \mathcal{L}_L$ auch $tX \in \mathcal{L}_L$ für alle $t \in \mathbb{R}$ gilt.

Satz 3.6.3. (i) \mathcal{L}_L ist ein \mathbb{R} -Vektorraum.

(ii) Für $X \in \mathcal{L}_L$ liegt e^X in der Wegzusammenhangskomponente von Id .

(iii) Mit $X \in \mathcal{L}_L$ und $A \in L$ gilt $AXA^{-1} \in \mathcal{L}_L$.

(iv) Stets ist $[X, Y] \in \mathcal{L}_L$, es liegt also eine Lie-Algebra vor.

Beweis: (i) Für $n \times n$ -Matrizen X, Y gilt die Formel

$$e^{X+Y} = \lim_m (e^{X/m} e^{Y/m})^m.$$

Aus der Abgeschlossenheit von L folgt dann sofort, dass \mathcal{L}_L ein \mathbb{R} -Vektorraum ist

(Es ist $e^{X/m} e^{Y/m} = \text{Id} + (X+Y)/m + O(1/m^2)$. Nun muss man „nur“ noch nachprüfen, dass die obige Formel $(\text{Id} + A/n)^n \rightarrow e^A$ auch bei kleinen Störungen richtig bleibt.)

(ii) $t \mapsto e^{tX}$ ($0 \leq t \leq 1$) ist ein stetiger Weg von Id nach e^X .

(iii) Sei $X \in \mathcal{L}_L$ und $A \in L$. Die Behauptung folgt sofort aus der Formel

$$e^{t(AXA^{-1})} = A e^{tX} A^{-1}.$$

(iv) Als endlichdimensionaler reeller Unterraum der $n \times n$ -Matrizen ist \mathcal{L}_L abgeschlossen.

Seien $X, Y \in \mathcal{L}_L$. Für jedes $t \in \mathbb{R}$ liegt (wegen (iii)) $e^{itX} Y e^{-itX}$ in \mathcal{L}_L , und damit gehört auch die Ableitung dieses Pfades (als Limes der Sekantensteigungen) bei $t = 0$ dazu. Die ist aber gleich $XY - YX$. \square

3.6.6 Untergruppen und Darstellungen

Zum Abschluss stellen wir noch einige Begriffe und Tatsachen über den Zusammenhang zwischen L und \mathcal{L}_L zusammen.

- Die *Exponentialabbildung* von \mathcal{L}_L nach L ist die Abbildung $X \mapsto e^X$.
- Ist L einfach zusammenhängend, so ist L durch \mathcal{L}_L eindeutig bestimmt.

- Gruppenhomomorphismen übersetzen sich in Lie-Algebra-Homomorphismen der zugehörigen Lie-Gruppen.
- Jede reelle Lie-Algebra entsteht als Lie-Algebra einer Lie-Gruppe.
- Die Darstellungen von L lassen sich aus den Darstellungen von \mathcal{L}_L