Wahrscheinlichkeitsmaße in unendlichen Gruppen

Karsten Hiddemann

29. April 2003

Zusammenfassung

In dem relativ neuen, als "statistische Gruppentheorie" bezeichneten Teil der Gruppentheorie, sind typische Ergebnisse Aussagen wie "ein zufälliges Element (oder Tupel von Elementen) einer Gruppe G hat Eigenschaft P mit Wahrscheinlichkeit p". Da solche Aussagen stark abhängig davon sind, wie man Wahrscheinlichkeit auf Gruppen definiert bzw. wie man Mengen in Gruppen mißt, sollen hier neue Zugänge zur Definition von Maßen gegeben werden.

1 Ziel

Es werden Ansätze untersucht, zu einer gegebenen unendlichen Gruppe G Maße zu konstruieren, die den folgenden Vorgaben gerecht werden:

(C1) Das Maß μ sollte natürlich sein, d.h. es sollte unseren Erwartungen entsprechen, was die Größe verschiedener Mengen betrifft.

(C2) μ sollte sensibel sein, d.h. intuitiv verschieden große Mengen sollten auch ein unterschiedliches Maß besitzen.

(C3) $\mu(S)$, oder eine ziemlich genaue Schranke, sollte für "natürliche" Mengen einfach berechenbar sein.

(C4) μ sollte einen natürlichen Erzeuger zufälliger Elemente erlauben.

2 Atomare Wahrscheinlichkeitsmaße

Erinnerung 2.1. Ein Maß μ auf einer Menge X heißt *atomar*, falls X abzählbar und S μ-meßbar für alle $S \subseteq X$. Aufgrund der Additivität wird μ in diesem Fall eindeutig durch die Einermengen $\{\omega\}$, $\omega \in X$, bestimmt:

$$\mu(S) = \sum_{\omega \in S} \mu(\{\omega\})$$

Wir schreiben daher im folgenden statt $\mu(\{\omega\})$ auch $\mu(\omega)$ für $\omega \in X$.

Wir betrachten speziell atomare Wahrscheinlichkeitsmaße auf der freien Gruppe $G = F_n$. Dazu fixieren wir eine beliebige Basis $\{x_1, \ldots, x_n\}$ von G, denn es gilt das folgende

Lemma 2.2. *Das singuläre Maß ist das einzige* Aut F_n *invariante Wahrscheinlichkeitsmaß.*

Beweis. Für $g^h := hgh^{-1}$ gelte also $\mu(g^h) = \mu(g)\ \forall h \in G$. Da G unendlich ist, folgt wegen der Additivität und Beschränktheit von μ für $g \neq 1$ damit $\mu(g) = 0$, denn sonst wäre $\mu(\{g^h \mid h \in I\}) = \sum_{h \in I} \mu(g) = |I| \cdot \mu(g) > 1$ für ein genügend groß gewähltes $I \subseteq G$. Für $g = 1$ gilt $1^h = 1$ und mit dem vorigen folgt $\mu(1) = 1$. □

Definition 2.3. Eine Funktion $c : G \to \mathcal{N}$ heißt *Komplexität* oder auch *Komplexitätsfunktion*, falls für alle $n \in \mathcal{N}$ gilt, daß das Urbild $c^{-1}(n) \subseteq G$ endlich ist.

Beispiel 2.4. Sei G eine endlich erzeugte Gruppe und $S \subseteq G$ eine erzeugende Menge von G. Die Längenfunktion

$$l_S(g) = \min\{n \mid g = y_1 \dots y_n, y_i \in S \cup S^{-1}\}$$

ist eine Komplexität auf G.

Bemerkungen 2.5.

- Besitzt G eine Komplexität, so ist G also insbesondere abzählbar. Durch geeignete Verallgemeinerungen, z.B. mit reellwertigen Komplexitäten, kann man diesen Begriff aber geeignet verallgemeinern.
- Ist l_S wie oben und $H < G$, so ist $l_S|_H$ ebenfalls eine Komplexität auf H, dies muß aber nicht notwendigerweise mehr eine Längenfunktion auf H sein.

Definition 2.6. Sei $c : G \mapsto \mathcal{N}$ eine Komplexität und μ ein Maß auf G. Wir nennen μ *c-invariant*, falls $\mu(u) = \mu(v)$ für alle $u, v \in G$ mit $c(u) = c(v)$.

Wir interessieren uns insbesondere für komplexitätsinvariante Maße. Bezeichne C_k die Sphäre $C_k = \{\omega \in G \mid c(\omega) = k\}$ und B_n den Ball $B_n = \{\omega \in G \mid c(\omega) \leq n\}$, so ist der Ansatz zur allgemeinen Konstruktion von Maßen, die unseren Wünschen entsprechen, das folgende

Lemma 2.7. *Sei μ ein atomares Maß und c eine Komplexität auf G.*

a) *Ist μ c-invariant, so ist $d_\mu : \mathcal{N} \to \mathcal{R}$, $d_\mu : k \mapsto \mu(C_k)$ ein Wahrscheinlichkeitsmaß auf $\mathcal{N}$.*

b) *Ist $d : \mathcal{N} \to \mathcal{R}$ ein Wahrscheinlichkeitsmaß auf $\mathcal{N}$, so ist $p_{c,d} : G \to \mathcal{R}$,*

$$p_{c,d} : \omega \mapsto \frac{d(c(\omega))}{|C_n|}, n = c(\omega)$$

eine Funktion, die ein atomares c-invariantes Maß auf G liefert.

Man sucht jetzt also geeignetes c und d, die den Vorgaben (C1) - (C4) genügen.

3 Von Random Walks erzeugte Maße

Wir untersuchen zunächst die Erzeugung zufälliger Elemente. Im Allgemeinen werden zufällige Elemente erzeugt, indem man ein Wort der Länge l generiert und Relationen anwendet, dies entspricht einem Random Walk im Cayleygraphen der Gruppe. Sei M_n der Monoid zu F_n mit den Erzeugenden $x_1, X_1, \dots, x_n, X_n$ und dem Homomorphismus $\pi : M_n \to F_n$, welcher x_i auf x_i und X_i auf x_i^{-1} abbildet. Dann kann man einen Random Walk der Länge k im Cayleygraphen durch ein Wort von M_n beschreiben.

Feststellung 3.1. Sei $d : \mathcal{N} \to \mathcal{R}$ eine Wahrscheinlichkeitsverteilung, c eine Komplexität auf M_n, $\tilde{\mu} = \mu_{c,d}$ ein c-invariantes atomares Maß auf M_n. Die Wahrscheinlichkeit, daß ein Random Walk das Element x erzeugt, ist

$$\mu(x) = \sum_{\pi(\omega)=x} \tilde{\mu}(\omega)$$

Lemma 3.2. *Für den Erwartungswert der geodätischen Länge von x in $\Gamma(F_n)$, welches von einem Random Walk der Länge k erzeugt wird, gilt die Abschätzung*

$$\frac{n-1}{n}k \leq E(l_{geod}(x)) \leq k$$

Beweis. Verfolgt man die schrittweise Änderung der geodätischen Länge beim Random Walk, so entspricht dies einer Bewegung auf $\mathbb{N} \cup \{0\}$ mit Spiegelung in Null. In Null beträgt die Wahrscheinlichkeit für eine Bewegung nach rechts 1, ansonsten hat man Wahrscheinlichkeiten $p = \frac{2n-1}{2n}$ für rechts und $q = \frac{1}{2n}$ für links (in diesem Fall läuft man die zuvor gewählte Kante im Graphen wieder zurück). Statt an Null zu spiegeln betrachten wir nun den Lauf auf $\mathbb{Z}$ ohne Spiegelung an Null, der ursprüngliche Erwartungswert von l ist mindestens so groß wie der Erwartungswert in diesem Fall. Bei k Schritten vollzieht man m Rechtsschritte mit der Wahrscheinlichkeit $\binom{k}{m} \cdot p^m \cdot q^{k-m}$, mit dem Endpunkt $l' = m - (k-m) = 2m - k$. Daher ist die Zufallsvariable l' eine lineare Transformation der Zufallsvariablen m, verteilt anhand der Binomialverteilung. Da gilt

$$\begin{aligned} E(m) &= \sum_{m=0}^{k} m\binom{k}{m}p^m q^{k-m} = pk\sum_{m=1}^{k}\binom{k-1}{m-1}p^{m-1}q^{k-1-(m-1)} \\ &= pk\sum_{m=0}^{k-1}\binom{k-1}{m}p^m q^{k-1-m} = pk(p+q)^{k-1} = pk \end{aligned}$$

erhält man für den Erwartungswert $E(l_{geod}(x))$ die untere Schranke

$$E(l') = 2E(m) - k = 2pk - k = \frac{n-1}{n}k$$

Die obere Schranke ist klar. □

4 Induzierte Maße auf endlichen Faktorgruppen

Wir würden für Mengen gerne natürliche Maße haben, z.B. das Maß $\frac{1}{2}$ für Wörter gerader Länge oder das Maß $\frac{1}{m}$ für Untergruppen $H < F$ mit Index $|F : H| = m$. Ein Problem stellt hierbei der sogenannte Einfluß kurzer Elemente dar: Sei μ ein Wahrscheinlichkeitsmaß und C_k wie oben, dann folgt aus $\sum_k \mu(C_k) = 1$, daß $\mu(C_k) \to 0$ für $k \to \infty$. Da $1 \in H$ gilt für $m \gg \frac{1}{\mu(1)}$ dann aber $\mu(H) \geq \mu(1) \gg \frac{1}{m}$. Bei dem Arbeiten mit Wahrscheinlichkeiten auf endlichen Gruppen werden Maße bevorzugt, welche gut verträglich sind mit "großen" endlichen Faktoren F_n/R, welche z.B. eine sinnvolle Schranke liefern bei dem Gesamtvarianzabstand

$$\frac{1}{2}\sum_{\overline{g}\in F_n/R}\left|\mu(\overline{g}) - \frac{1}{|F_n : R|}\right| \tag{1}$$

für das induzierte Maß auf F_n/R. Aus praktisch dem gleichen Grund kann diese natürliche Bedingung nicht erfüllt werden, denn $\frac{1}{2}|\mu(1) - \frac{1}{m}|$ mit $m = |F_n : R|$ ist eine untere Schranke welche nicht gegen 0 konvergiert für $m \to \infty$. Diese unentschuldbare Abhängigkeit von dem Maß kurzer Elemente ist eine von vielen Gründen für die folgende (womöglich kontroverse) metamathematische These:

Es gibt kein atomares Maß auf unendlichen Gruppen, welches unseren Erwartungen zur Größe bestimmter Mengen in der Gruppe gerecht wird.

Man kann versuchen unnatürlich große Wahrscheinlichkeiten kurzer Elemente auszuschliessen, indem man für ein Element gR in der Faktorgruppe F_n/R mit B_l wie oben das normalisierte Maß "großer" Elemente in gR verwendet:

$$\overline{\mu}_l(gR) = \frac{\mu((F_n \setminus B_l) \cap gR)}{\mu(F_n \setminus B_l)}$$

Ein Maß $\mu = \overline{\mu}_l$ ist "gut", wenn der Gesamtvarianzabstand (1) für $l \ll m$ nach oben beschränkt ist durch $\frac{1}{e}$ oder eine andere sinnvolle Konstante, oder wenn dieser exponentiell abnimmt für wachsendes l, d.h. $o(e^{-cl})$ ist. Wir nähern uns also dem klassischen Konzept eines Random Walk auf einer Gruppe: Bilder von "ausreichend langen" Elementen zu nehmen bedeutet einen ausreichend langen Random Walk bezüglich der charakteristischen Wortlänge l auf der Faktorgruppe F_n/R mit den Erzeugern $x_1R, \ldots, x_nR$ durchzuführen.

Definition 4.1. Die *Mischzeit* ist definiert als das minimale l_0, so daß

$$\|P_{l_0} - U\| = \frac{1}{2} \sum_{\overline{g} \in F_n/R} \left| P_{l_0}(\overline{g}) - \frac{1}{|F_n : R|} \right| < \frac{1}{e}$$

wobei $P_l(\overline{g})$ die Wahrscheinlichkeit ist, daß ein Random Walk von Länge l bei dem Element $\overline{g}$ stoppt und U die Gleichverteilung auf F_n/R ist.

Erinnerung 4.2. Ab der Mischzeit konvergiert das von Random Walks erzeugte Maß exponentiell schnell gegen die Gleichverteilung:

$$\|P_{kl_0} - U\| = \frac{1}{2} \sum_{\overline{g} \in F_n/R} \left| P_{kl_0}(\overline{g}) - \frac{1}{|F_n : R|} \right| = o(e^{-ck})$$

Wir können nun ein alternatives Kriterium für gutes Verhalten von Maßen auf endlichen Faktorgruppen aufnehmen in unseren ursprünglichen Wunschzettel:

(C1*) Für jede endliche Faktorgruppe $G = F_n/R$ existiert ein l_0, so daß

$$\frac{1}{2} \sum_{\overline{g} \in F_n/R} \left| \overline{\mu}_{kl_0}(\overline{g}) - \frac{1}{|F_n : R|} \right| = o(e^{-ck})$$

Es ist recht einfach zu sehen, daß das Kriterium (C1*) von einer großen Klasse von Maßen erfüllt wird:

Satz 4.3. *Erfüllt die Wahrscheinlichkeitsverteilung $d(k)$ die Bedingung $d(k) > 0$ für unendlich viele k, dann erfüllt das induzierte Maß $\overline{\mu}_l$ auf der endlichen Faktorgruppe $G = F_n/R$ die Bedingung (C1*).*

Beweis. Sei $m = |G|$. Identifiziere die Wahrscheinlichkeitsverteilungen auf G mit $(p_1, \ldots, p_m) \in \mathbb{R}^m$, d.h. $p_1 + \ldots + p_m = 1$, $p_i \geq 0 \ \forall i$. Dann entspricht die Menge der Wahrscheinlichkeitsverteilungen der konvexen Hülle Δ von den Verteilungen E_g mit $E_g(h) = 1$, falls $g = h$ und 0 sonst. Der Gesamtvarianzabstand

$$\|P - Q\| = \frac{1}{2} \sum_{g \in G} |P(g) - Q(g)|$$

ist dann eine Metrik auf Δ. Ein Schritt eines Random Walks in dem Cayley-Graphen induziert eine affine Transformation $\tau : \mathbb{R}^m \to \mathbb{R}^m$ auf den Verteilungen durch

$$\tau : E_g \mapsto \frac{1}{2n} \sum_{h \text{ adjazent zu } g} E_h,$$

d.h. τ hat den Fixpunkt $U = (1/m, \ldots, 1/m) \in \Delta$.

Verschieben wir nun den Ursprung des Koordinatensystems von $\mathbb{R}^m$ nach U, so wird der Abstand $\|X-U\|$ zu einer Norm, welche wir nun kurz mit $\|X\|$ bezeichnen und τ wird dann auf dem durch Δ gegebenen Unterraum zu einem linearen Operator. τ ist eine Selbstabbildung von Δ auf sich, genauso für $\tau^k(\Delta)$, daher haben wir

$$\|\tau^{k+1}(E_g)\| \leq \|\tau^k(E_g)\|$$

Nachrechnen ergibt für die euklidische Norm $\|\tau(E_g)\|_E < \|E_g\|_E$, d.h. es gilt

$$\|\tau\|_E = \max_{X\in\Delta, X\neq 0} \frac{\|\tau(X)\|_E}{\|X\|_E} < 1$$

für die natürliche zugehörige Matrixnorm. Es folgt $\|\tau^k(X)\| \to 0$ für $k \to \infty$, die Konvergenz ist sogar exponentiell schnell: $\|\tau^k(X)\|_E = o(c^k)$. Da im endlichdimensionalen Fall jeweils zwei Normen zueinander äquivalent sind, gibt es ein C, so daß

$$\frac{1}{C}\|X\|_E < \|X\| < C\|X\|_E$$

und damit folgt, daß $\|\tau^k(X)\| = o(c^k)$. Wir haben in unserer neuen Notation

$$\begin{aligned}\frac{1}{2}\sum_{g\in G}\left|\overline{\mu}_l(g) - \frac{1}{m}\right| = \|\overline{\mu}_l\| &= \left\|\frac{1}{\sum_{k\geq l} d(k)}\sum_{k\geq l} d(k)\tau^k(E_1)\right\| \\ = \frac{1}{\sum_{k\geq l} d(k)}\left\|\sum_{k\geq l} d(k)\tau^k(E_1)\right\| &\leq \frac{1}{\sum_{k\geq l} d(k)}\left\|\sum_{k\geq l} d(k)\tau^l(E_1)\right\| \\ &= \|\tau^l(E_1)\| \cdot \frac{\sum_{k\geq l} d(k)}{\sum_{k\geq l} d(k)} = \|\tau^l(E_1)\|.\end{aligned}$$

Wir sehen, daß die Wahrscheinlichkeitsverteilung auf G, erzeugt durch einen Random Walk von Länge l oder größer, genauso schnell gegen die Gleichverteilung konvergiert wie die Wahrscheinlichkeitsverteilung nach l Schritten beim ursprünglichen Random Walk, dies zeigt, daß $\overline{\mu}_l$ auf G die Bedingung (C1*) erfüllt. □

5 Wachstumsfunktion und asymptotische Dichte

Zuletzt wollen wir noch das asymptotische Verhalten von Mengen untersuchen.

Definition 5.1. Ein *Pseudomaß* auf G ist eine nichtnegative reellwertige Funktion, die auf bestimmten Teilmengen von G definiert ist. Ein Pseudomaß μ heißt *atomar*, wenn $\mu(S)$ definiert ist für jedes endliche $S \subseteq G$.

Definition 5.2. Wir definieren die *sphärisch-asymptotische Dichte* $\rho^{(s)}$ durch

$$\rho_k^{(s)}(R) = \frac{\mu(R\cap C_k)}{\mu(C_k)}, \rho^{(s)}(R) = \lim_{k\to\infty}\rho_k^{(s)}(R)$$

und die *kugel-asymptotische Dichte* $\rho^{(d)}$ durch

$$\rho_k^{(d)}(R) = \frac{\mu(R\cap B_k)}{\mu(B_k)}, \rho^{(d)}(R) = \lim_{k\to\infty}\rho_k^{(d)}(R)$$

für $R \subseteq G$. Man kann in der obigen Definition auch lim sup verwenden statt lim.

Lemma 5.3. *Ist μ ein atomares Pseudomaß auf G, so auch $\rho^{(s)}$ und $\rho^{(d)}$.*

Lemma 5.4. *Sei μ Pseudomaß auf G mit $\lim_{k\to\infty}\mu(B_k)=\infty$. Dann existiert $\rho^{(s)}(R)$, wenn $\rho^{(d)}(R)$ existiert und in diesem Fall gilt Gleichheit: $\rho^{(s)}(R)=\rho^{(d)}(R)$.*

Beweis. Sei $x_k=\mu(R\cap B_k)$ und $y_k=\mu(B_k)$. Es gilt $y_k<y_{k+1}$ und $\lim y_k=\infty$. Nach Stolz Theorem [3] ist dann

$$\lim_{k\to\infty}\frac{x_k}{y_k}=\lim_{k\to\infty}\frac{x_k-x_{k-1}}{y_k-y_{k-1}}$$

und wegen der Additivität gilt nach Definition von B_k und C_k

$$\rho^{(d)}(R)=\lim_{k\to\infty}\frac{\mu((R\cap B_k)\setminus B_{k-1})}{\mu(B_k\setminus B_{k-1})}=\lim_{k\to\infty}\frac{\mu(R\cap C_k)}{\mu(C_k)}=\rho^{(s)}(R)$$

□

Unter diesem Gesichtspunkt sprechen wir nun allgemein von den Dichten ρ. Asymptotische Dichten sind ein nützliches Werkzeug zum Beschreiben des Verhaltens von Mengen im Unendlichen, allerdings gibt es viele natürliche Mengen, welche hiermit leider nicht meßbar sind. Man sieht z.B. sofort, daß $\rho(E_n)$ für

$$E_n=\{\omega\in F_n\mid |\omega| \text{ gerade }\}$$

nicht definiert ist. Die folgenden Theoreme zeigen, daß ρ nicht sensibel genug ist:

Satz 5.5 (Woess [4]). *Ist N eine normale Untergruppe von F_n, $n\geq 2$, mit unendlichem Index, so ist $\rho(N)=0$.*

Erinnerung 5.6. Wir nennen ein Element $u\in F_n$ *primitiv*, falls es Teil einer freien Basis von F_n ist, bzw. wenn äquivalent gilt $\alpha(u)=x_1$ für ein $\alpha\in\operatorname{Aut}F_n$.

Satz 5.7. *Sei F_n die freie Gruppe von endlichem Rang $n\geq 2$. Dann gilt*

$$\rho(\mathrm{Pr}_n)=0$$

wobei Pr_n die Menge der primitiven Elemente von F_n ist. Genauer, sei $P(n,k)$ die Zahl der primitiven Elemente von Länge k in F_n, $n\geq 3$, dann gibt es Konstanten c_1 und c_2 mit

$$c_1\cdot(2n-3)^k\leq P(n,k)\leq c_2\cdot(2n-2)^k$$

Proposition 5.8. *In F_2 ist die Zahl der primitiven Elemente $P(2,k)$*

(a) *größer als $\frac{4}{\sqrt{3}}\cdot(\sqrt{3})^k$ falls k ungerade ist.*

(b) *größer als $\frac{4}{3}\cdot(\sqrt{3})^k$ falls k gerade ist.*

Dies heißt, daß die "meisten" primitiven Elemente in F_2 Konjugate primitiver Elemente kleinerer Länge sind. Dies ist nicht der Fall in F_n mit $n>2$, wo die "meisten" primitiven Elemente von der Form $u\cdot x_i^{\pm 1}\cdot v$ sind, wobei u und v beliebige Elemente sind, die nicht von x_i abhängen.

Literatur

[1] A. Borovik, A. G. Myasnikov, V. Shpilrain, *Measuring sets in infinite groups*, Contemp. Math., Amer. Math. Soc. **298** (2002), 21–42.

[2] I. Kapovich, A. G. Myasnikov, P. Schupp, and V.Shpilrain, *Generic-case complexity, decision problems in group theory and random walks*, J. Algebra **264** (2003), 665–694.

[3] Guntram Hainke, *Generische Komplexität gruppentheoretischer Entscheidungsprobleme I: Das Wortproblem*, talk on [2] at the University of Dortmund, 2003

[4] W. Woess, *Cogrowth of groups and simple random walks*, Arch. Math. **41** (1983), 363–370.