

Zusammenhang zwischen Fraktaler Dimension und Konvergenzgeschwindigkeit der Spektralreihe bei rationalen Fraktalen

Diplomarbeit

Universität Siegen
Fachbereich 6 – Mathematik

zur Erlangung des akademischen Grades
Diplom-Mathematiker
vorgelegt von

Simon Budig

Siegen, 12. März 2003

Gutachter: Prof. Dr. Nils-Peter Skoruppa, Universität Siegen
Prof. Dr. Wolfgang Merzenich, Universität Siegen

Zusammenhang zwischen Fraktaler Dimension und Konvergenzgeschwindigkeit der Spektralreihe bei rationalen Fraktalen

Inhaltsverzeichnis

1	Einleitung	4
1.1	Das Thema...	4
1.2	Die Arbeit...	5
1.3	Der Dank...	5
2	Formale Sprachen	6
2.1	Sprachen und Grammatiken	6
2.2	Erkennung von regulären Sprachen	9
2.3	Sprachen als Lösungen von Gleichungssystemen	11
3	Matrizen	13
3.1	Matrixnormen	13
3.2	Die Jordansche Normalform einer Matrix	14
3.3	Eigenschaften von speziellen Matrizen	16
4	Fraktale Mengen	18
4.1	Hausdorff Maß und Dimension	19
4.2	Boxcounting-Dimension	21
5	Bilderzeugung mit Sprachen	23
5.1	Transformationen von Bildern	23
5.2	Bilder aus Sprachen	28
6	Die Spektralreihe	31
6.1	Die Elemente des Spektrums	31
6.2	Konvergenzradius der Spektralreihe	32
6.3	Geschlossene Form für das Spektrum	34
7	Fraktale Dimension	35
7.1	Obere Boxcounting-Dimension	35
7.2	Beziehung zur Hausdorff-Dimension nach [Me/St94]	36

8 Der Maßvektor als Eigenvektor von P **39**

8.1 Boxcounting-Volumen 39

8.2 Irreduzible Automaten 41

8.3 Allgemeine Automaten 41

9 Jordansche Normalform für Dimension und Maß **43**

9.1 Konvergenzradius der Spektralreihe 44

9.2 Boxcounting-Volumen 46

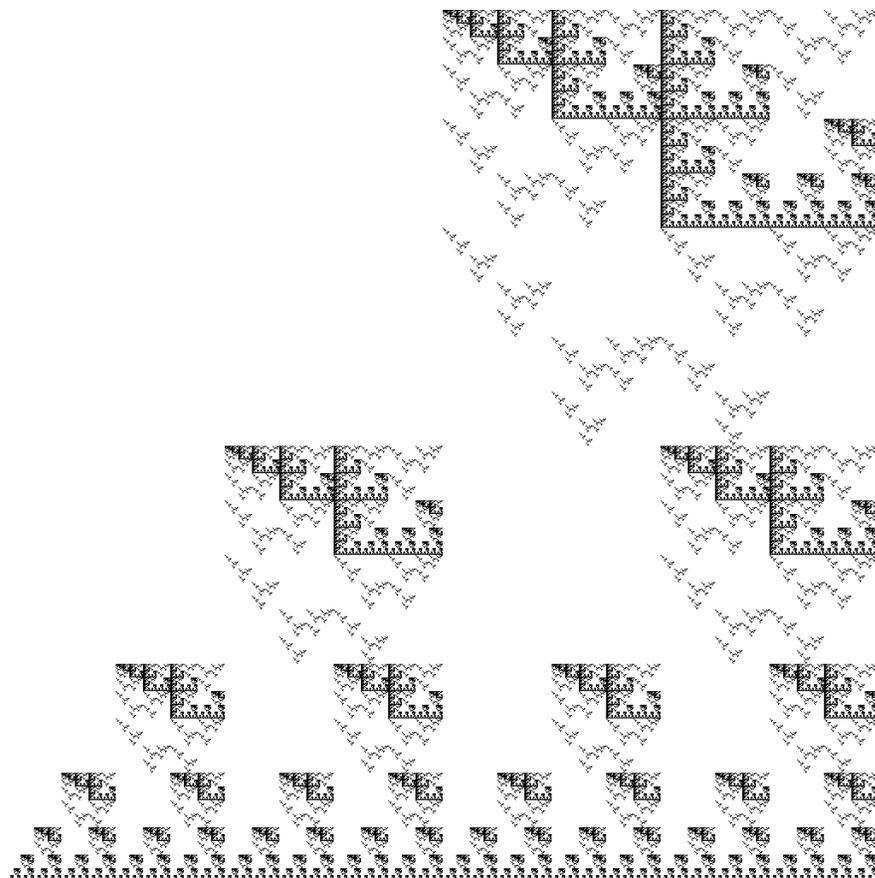
10 Ausblick **48**

A Beispiel **50**

A.1 Berechnung von Boxcounting-Dimension und Maßvektor 50

A.2 Alternativer Ansatz mittels Jordanform 53

A.3 Mit Automaten malen... 56



1 Einleitung

1.1 Das Thema...

In dieser Arbeit betrachten wir rationale Fraktale, die mittels einer sehr simplen und überschaubaren Beschreibung erzeugt werden und eine teilweise überraschende Komplexität zeigen.

Die Beschreibung dient dazu, das Verhalten eines endlichen Automaten – einer „mathematischen Maschine“ – zu erfassen. Automaten werden in der Informatik neben anderen Kriterien dazu genutzt, (nicht natürliche) Sprachen zu erkennen und in verschiedene Komplexitätsklassen einzuteilen. Endliche Automaten beschreiben die Klasse der sogenannten regulären Sprachen, eine Sprachklasse, die für viele Probleme der Informatik nicht komplex genug ist. So gehört zum Beispiel jede in der Praxis verwendete Programmiersprache nicht mehr zu den regulären Sprachen. Dennoch haben reguläre Sprachen eine innere Struktur, die man der relativ einfachen Beschreibung des Automaten nicht ansieht.

In dieser Arbeit werden die Worte einer regulären Sprache dazu verwendet, um nach einer festgelegten Vorschrift eine Bildfolge mit zunehmendem Detailgrad zu erzeugen. Die „Grenzbilder“ dieser Folgen heißen „rationale Fraktale“. Man stellt fest, dass diese Bilder nur bei einer sehr speziellen Sichtweise einfach sind. Neben sehr geometrisch und glatt aussehenden Bildern gibt es Bildfolgen, die mit zunehmendem Detailgrad seltsame und mit Worten schlecht zu beschreibende Strukturen enthüllen. Das sind genau die interessanten und ästhetisch ansprechenden Fälle, die man versucht, mit fraktalen Eigenschaften zu beschreiben.

Ein zentrales Maß für die „Unregelmäßigkeit“ eines derart beschriebenen Bildes ist die „Fraktale Dimension“. Wenn diese einen nicht-ganzzahligen Wert hat, ist dies ein deutlicher Hinweis darauf, dass sich das Bild nicht mit der klassischen Geometrie beschreiben lässt, sich also nicht einfach aus einer endlichen Anzahl von Linien, Kreisen und Polygonen zusammensetzen lässt.

In dieser Arbeit werden wir feststellen, dass es eine enge Beziehung zwischen dieser fraktalen Dimension und der Verteilung der Wörter in der durch den Automaten beschriebenen Sprache gibt. Interessanterweise ist dieser Zusammenhang auch noch in einer Kenngröße des Automaten wiederzufinden. Damit lässt sich zu einem gegebenen Automaten sehr schnell und bequem die fraktale Dimension des erzeugten Bildes errechnen.

1.2 Die Arbeit...

Die Kapitel 2 bis 4 zitieren die notwendigen Grundlagen aus Informatik und Mathematik. Kapitel 5 beschreibt, wie die Bilder aus den Sprachen erzeugt werden und behandelt als erste Ergebnisse wichtige Beziehungen zwischen der Präfixrelation und den erzeugten Bildern.

Die Elemente des Spektrums sowie die zugehörige Spektralreihe werden in Kapitel 6 betrachtet. Zentrales Ergebnis ist die Berechnung des Konvergenzradius. In Kapitel 7 berechnen wir die obere Boxcounting-Dimension eines Bildes und stellen die Beziehung zum Konvergenzradius der Spektralreihe her. Zusätzlich wird diskutiert, warum die obere Boxcounting-Dimension eigentlich genau derselbe Wert ist, wie die über einen ziemlich anderen Ansatz bestimmte Hausdorff-Dimension nach [Me/St94].

In Kapitel 8 betrachten wir einige notwendige Bedingungen für den Maßvektor und geben einen Algorithmus an, wie der Maßvektor für komplexere Automaten konstruiert werden kann. Für Automaten die eine bestimmte Voraussetzung erfüllen, wird schließlich in Kapitel 9 eine Methode vorgestellt, mit der man Dimension und Maßvektor mit Hilfe eines Computeralgebrasystems sehr bequem ausrechnen kann ohne den Automaten speziell sortieren zu müssen.

Der Anhang A führt die in dieser Arbeit vorgestellten Methoden an einem Beispiel ausführlich vor.

1.3 Der Dank...

Die Erstellung dieser Arbeit wäre nicht möglich gewesen ohne die Unterstützung von vielen Leuten. Allen voran möchte ich meinen Eltern danken, die mir während meines Studiums den Rücken freigehalten haben und mich vorbehaltlos unterstützt haben.

Dank geht auch an meine beiden Betreuer, Herrn Merzenich und Herrn Skoruppa. Mit ihnen habe ich viele interessante Gespräche geführt, in denen mein Blick auf das eigentliche Problem geschärft sowie Lösungsansätze entwickelt wurden. Zudem waren die Gespräche sehr unterhaltsam.

Last but not least möchte ich meinen Freunden und Geschwistern danken, ohne die Leben und Studium nicht lebenswert gewesen wäre.

Euch/Ihnen allen Herzlichen Dank!

Siegen, den 12. März 2003, Simon Budig

2 Formale Sprachen

Zunächst ein Überblick über Grundbegriffe der theoretischen Informatik. Wie auch bei natürlichen Sprachen, vereinbart man für die Schriftsprache ein gemeinsames zugrundeliegendes Alphabet. Allein durch die beliebige Aneinanderreihung von Buchstaben entsteht aber nicht unbedingt etwas Sinnvolles. Man muss eine gemeinsame Sprache vereinbaren.

2.1 Sprachen und Grammatiken

Definition 2.1. Ein *Alphabet* ist eine endliche, nichtleere Menge X . Die Menge $X^n := \underbrace{X \times \cdots \times X}_{n \text{ mal}}$ enthält die *möglichen Worte der Länge n* , ein Wort $w \in X^n$ hat die *Länge* $|w| := n$. Die Vereinigung über alle X^n bilden die Menge der *möglichen Worte*: $X^* := \bigcup_{n \in \mathbb{N}_0} X^n$. Die Menge X^* enthält insbesondere auch das leere Wort $\varepsilon \in X^0$.

Eine Teilmenge $L \subseteq X^*$ heißt *Sprache*.

Definition 2.2. Die *Vereinigung* „ \cup “ zweier Sprachen L und L' ist in der offensichtlichen Weise definiert:

$$w \in L \cup L' \iff w \in L \text{ oder } w \in L'.$$

Sind die Sprachen L und L' disjunkt so schreibt man für die Vereinigung auch das Symbol „ $+$ “. In diesem Fall gilt offensichtlich $|L + L'| = |L| + |L'|$.

Das neutrale Element bezüglich der Vereinigung ist die leere Menge \emptyset .

Definition 2.3. Zwei Worte $w, v \in X^*$ können miteinander *konkateniert* (aneinandergehängt) werden. Die Konkatenation „ \cdot “ ist eine Operation auf X^* :

$$\cdot : \begin{cases} X^* \times X^* & \longrightarrow X^* \\ (w, v) & \longmapsto w \cdot v =: wv \end{cases}$$

Das leere Wort ε ist das neutrale Element bezüglich der Konkatenation.

Diese Operation induziert eine entsprechende Operation auf Teilmengen von X^* (bzw. Sprachen): Seien L, L' Sprachen, dann ist

$$L \cdot L' := \{wv \mid w \in L \text{ und } v \in L'\} \subseteq X^*.$$

Konkateniert man eine Sprache mit der leeren Menge (oder umgekehrt), ist das Ergebnis die leere Menge. Konkateniert man hingegen eine Sprache

L mit der Sprache die nur das leere Wort ε enthält (oder umgekehrt), ist das Ergebnis wieder die Sprache L , damit ist $\{\varepsilon\}$ das neutrale Element bezüglich der Konkatenation von Sprachen.

Haben die Sprachen L und L' jeweils nur endlich viele Worte, so enthält auch ihre Konkatenation nur endlich viele Worte. Enthalten L und L' nur Worte der Länge k bzw. k' so enthält die Konkatenation $|L \cdot L'| = |L| \cdot |L'|$ Worte.

Definition 2.4. Seien w, p Worte aus X^* . Das Wort p heißt *Präfix* von w (Notation: $p \sqsubseteq w$) falls ein Wort p' existiert, so dass $p \cdot p' = w$ ist, der Präfix also zu dem Wort w erweitert werden kann.

Definition 2.5. Eine Sprache L heißt *präfixfrei*, wenn für jedes Wort $w \in L$ jeder echte Präfix $p \sqsubseteq w$ mit $p \neq w$ nicht in der Sprache L enthalten ist. Insbesondere ist die Teilsprache $L \cap X^k$ präfixfrei (da jeder echte Präfix eines Wortes $w \in L \cap X^k$ weniger als k Buchstaben hat).

Definition 2.6. Eine Sprache L heißt *präfixabgeschlossen*, wenn zu jedem Wort $w \in L$ alle Präfixe $p \sqsubseteq w$ ebenfalls in L enthalten sind.

Definition 2.7. Das *Spektrum* einer Sprache L beschreibt für alle Wortlängen k das Verhältnis zwischen den gültigen und den möglichen Worten der Länge k

$$g_k = \frac{|L \cap X^k|}{|X^k|} \in [0, 1].$$

Um nun die Sprachen in verschiedene Kategorien aufzuteilen, verwendet die Theorie der Formalen Sprachen sogenannte Grammatiken. Diese stellen ein Werkzeug zur Beschreibung von Sprachen dar.

Definition 2.8. Eine *Grammatik* $G = (N, T, P, \sigma)$ besteht aus

- einer endlichen, nichtleeren Menge N , den *nichtterminalen* Zeichen,
- einer endlichen, nichtleeren Menge T , den *terminalen* Zeichen bzw. das zugrundeliegende *Alphabet*, insbesondere ist $N \cap T = \emptyset$,
- einer endlichen, nichtleeren Menge von *Produktionen* $P \subseteq V^* \times V^*$, wobei $V := N \cup T$,
- es gibt ein ausgezeichnetes *Startsymbol* $\sigma \in N$.

Die Produktionen einer Grammatik G beschreiben mögliche lokale Ersetzungen. Enthält eine Zeichenfolge aus V^* an einer Stelle die linke Seite einer Produktion, kann in der Zeichenfolge die linke Seite durch die rechte Seite der Produktion ersetzt werden. Eine Zeichenfolge aus V^* heißt *Satzform*, wenn man sie über eine Folge derartiger Ersetzungen aus dem Startsymbol σ erzeugen kann. Gelangt man so zu einer Satzform, die nur aus terminalen Zeichen besteht, hat man ein Wort der durch die Grammatik G beschriebenen Sprache $L(G)$ gefunden. Umgekehrt kann man zu jedem Wort aus $L(G)$ eine Folge von Ersetzungen (eine *Ableitung*) angeben, die dieses Wort aus σ erzeugen. Abhängig von der Grammatik ist das jedoch keineswegs eine einfache Aufgabe.

Um eine gewisse Regelmäßigkeit zu bekommen, fordert man, dass die linke Seite der Produktionen nur aus Nichtterminalzeichen besteht. Damit bekommt der Begriff der „Terminalzeichen“ eine echte Bedeutung: erscheint nämlich ein solches Zeichen in einer Satzform, kann man sicher sein, dass es bei allen weiteren Ersetzungen erhalten bleibt.

Definition 2.9. Eine Grammatik $G = (N, T, P, \sigma)$ heißt *Chomsky-Grammatik*, wenn die Menge der Produktionen $P \subseteq N^+ \times V^*$ ist. Dabei ist $N^+ := N \cdot N^*$.

Eine einzelne Produktion hat also auf ihrer linken Seite nur Nichtterminalzeichen und davon mindestens eins. Man kann zeigen, dass dies keine Einschränkung gegenüber einer allgemeinen Grammatik ist.

Bemerkung. Chomsky-Grammatiken werden weiter unterteilt in Grammatiken vom Typ 0, Typ 1 (*kontextsensitiv*), Typ 2 (*kontextfrei*) und Typ 3 (*regulär*) wobei hier zunehmende Anforderungen an die Produktionen gestellt werden.

Die verschiedenen Grammatikklassen definieren verschiedene Sprachklassen \mathcal{L}_i . Eine Sprache L ist in \mathcal{L}_i enthalten, wenn eine Grammatik G vom Typ i existiert, so dass $L(G) = L$. Die Sprachklassen sind echte Teilmengen voneinander: $\mathcal{L}_3 \subsetneq \mathcal{L}_2 \subsetneq \mathcal{L}_1 \subsetneq \mathcal{L}_0$.

Die von uns betrachteten Sprachen sind regulär, also in \mathcal{L}_3 enthalten. Die regulären Sprachen über einem Alphabet X bezeichnet man auch als $\text{Reg}(X)$.

Definition 2.10. Eine Chomsky-Grammatik $G = (N, T, P, \sigma)$ wird *reguläre Grammatik* genannt, wenn die Menge der Produktionen $P \subseteq N \times (N \cup \{\varepsilon\})T^*$ ist (*linkslinär*) oder $P \subseteq N \times T^*(N \cup \{\varepsilon\})$ ist (*rechtslinär*).

2.2 Erkennung von regulären Sprachen

Man kann nun zeigen, dass eine Sprache L genau dann regulär ist, wenn ein endlicher Akzeptor \mathcal{A} (Deterministic Finite Automaton (DFA), endlicher Automat, „mathematische Maschine“) existiert, der Worte der Sprache L akzeptiert:

Definition 2.11. Ein *endlicher Akzeptor* wird durch ein 5-Tupel

$$\mathcal{A} = (X, S, \delta, s, F)$$

spezifiziert, wobei

- X das Eingabealphabet, $0 < |X| < \infty$,
- S die Zustandsmenge, mit $0 < |S| < \infty$,
- $F \subseteq S$ die Menge der akzeptierenden Zustände,
- $s_0 \in S$ der Startzustand und
- $\delta : S \times X \rightarrow S$ die (evtl. nur partiell definierte) Übergangsfunktion

ist.

Ein endlicher Akzeptor wendet ausgehend vom Startzustand s_0 die Übergangsfunktion δ auf die Buchstaben eines Wortes an und wechselt so von Buchstabe zu Buchstabe seinen Zustand. Befindet er sich am Ende des Wortes in einem akzeptierenden Zustand, so ist das Wort in der vom Automaten akzeptierten Sprache $L_{\mathcal{A}}$ enthalten.

Abkürzend schreibt man auch ¹ $sx = \delta(s, x)$ für $s \in S$ und $x \in X$ und setzt die Abbildung δ auf Worte aus X^* fort: $s\varepsilon = \delta(s, \varepsilon) := s$ und mit $w = xv \in X^*$ für ein $x \in X$ sei $\delta(s, w)$ rekursiv definiert als $\delta(s, w) := \delta(\delta(s, x), v)$ bzw. $s(xv) = (sx)v$.

Damit operiert X^* mittels δ auf S und die Akzeptanz von Worten aus $L_{\mathcal{A}}$ lässt sich folgendermaßen notieren:

$$w \in L_{\mathcal{A}} \iff s_0w \in F \quad \text{oder äquivalent:} \quad L_{\mathcal{A}} := \{w \in X^* \mid s_0w \in F\}.$$

Ist die Übergangsfunktion δ nur partiell definiert, existieren gewisse Zustand/Buchstabe-Kombinationen (s', x') , für die $\delta(s', x')$ nicht definiert ist. In diesem Fall gilt für alle $w \in X^*$, dass $(s'x')w$ ebenfalls nicht definiert ist.

¹Dies darf nicht mit der Konkatenation wv von zwei Worten $w, v \in X^*$ verwechselt werden, die in der Schreibweise ähnlich aussieht.

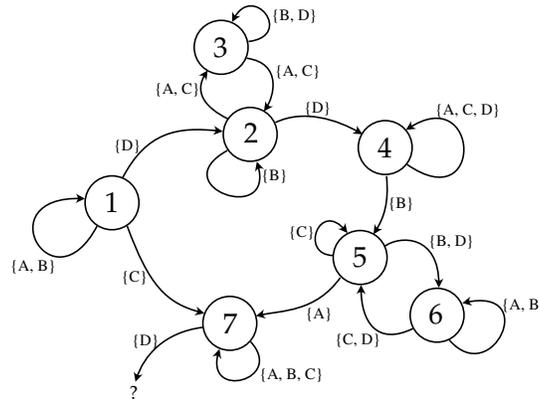


Abbildung 1: Ein möglicher Übergangsgraph eines Automaten, vergleiche Anhang A

Bemerkung. Die oben definierten Operationen „ \cup “ und „ \cdot “ sind bezüglich \mathcal{L}_3 abgeschlossen, d. h. für zwei Sprachen $L', L'' \in \mathcal{L}_3$ ist auch $L' \cup L''$ und $L' \cdot L''$ in der Menge der regulären Sprachen \mathcal{L}_3 . Einen Nachweis findet man zum Beispiel in [Mer96], Korollar 2.3.1.

Satz 2.12. Für jede reguläre, präfixabgeschlossene Sprache $L \neq \emptyset$ existiert ein endlicher Akzeptor $\mathcal{A} = (X, S, \delta, s, F)$ mit $F = S$, er enthält also ausschließlich akzeptierende Zustände.

Beweis. Sei $\mathcal{A}' = (X, S', \delta', s', F')$ ein endlicher Akzeptor der die Sprache L erkennt. Da L präfixabgeschlossen ist, kann man kein Wort $w \notin L$ mit einem Wort $v \in X^*$ derart erweitern, so dass $wv \in L$ ist.

Für den Automaten \mathcal{A}' bedeutet das, dass jeder nicht-akzeptierende Zustand i durch jedes Wort $w \in X^*$ in einen nicht-akzeptierenden Zustand überführt wird:

$$i \in S' \setminus F' \implies iw \in S' \setminus F' \quad \text{für alle } w \in X^*.$$

Da uns nicht interessiert in welcher spezifischen Weise ein bestimmtes Wort nicht erkannt wird, können wir alle nicht akzeptierenden Zustände zu einem Zustand zusammenfassen, der durch jedes Wort w in sich selbst überführt wird (derartige Zustände bezeichnet man auch als Fangzustand). Eine entsprechendes Verhalten können wir aber auch realisieren, indem wir in der Übergangsfunktion die Übergänge in die nicht akzeptierenden Zustände nicht definieren.

Der Automat wird dann für jedes Wort $w \notin L$ auf eine Zustand/Buchstabenkombination stoßen, für die die Übergangsfunktion nicht definiert ist, damit kann es, wie oben erläutert, kein $v \in X^*$ geben, so dass wv wieder erkannt wird.

Da die Sprache $L \neq \emptyset$ präfixabgeschlossen ist, ist auch das leere Wort ε in L enthalten. Damit ist sichergestellt, dass der Startzustand s' ein akzeptierender Zustand ist. Wir setzen also $S := S' \cap F' = F'$, $F := F'$, $s := s' \in S$ und

$$\delta(i, x) := \begin{cases} \delta'(i, x) & \text{falls } ix \in F \\ \text{undefiniert} & \text{sonst.} \end{cases}$$

Der endliche Akzeptor $\mathcal{A} = (X, S, \delta, s, F)$ hat also nur akzeptierende Zustände und erkennt die Sprache L . Damit erfüllt er die Anforderungen des Satzes. \square

2.3 Sprachen als Lösungen von Gleichungssystemen

Die Übergangsfunktion δ eines Automaten $\mathcal{A} = (X, S, \delta, s_0, F)$ ist nicht auf einen bestimmten Startzustand festgelegt. Man kann für jeden Zustand s der Zustandsmenge S eine eigene Sprache definieren:

$$L_s := \{w \in X^* : \delta(s, w) = sw \in F\}.$$

Zwischen diesen Sprachen besteht ein enger Zusammenhang. Die Worte in L_s , die mit dem Buchstaben x anfangen, sind genau die Konkatenation von x mit den Worten aus der Sprache L_{sx} :

$$xw \in L_s \iff w \in L_{sx}.$$

Damit läßt sich für jedes L_s eine Gleichung der folgenden Form aufstellen:

$$L_s = \sum_{x \in X} x \cdot L_{sx} + T_s. \quad (2.1)$$

Die Menge T_s gibt hierbei an, ob das leere Wort ε in der Sprache L_s enthalten ist (s ist ein akzeptierender Zustand) oder nicht:

$$T_s := \begin{cases} \{\varepsilon\} & \text{falls } s \in F \\ \emptyset & \text{sonst.} \end{cases}$$

Dies ist nötig, da die Summe nur eine Aussage über Worte der Mindestlänge 1 macht (jeder Summand beginnt mit einem Buchstaben $x \in X$). Da dieser für

jeden Summanden unterschiedlich ist, ist diese Summation (Vereinigung) von Sprachen disjunkt.

Die Gleichungen (2.1) können wir unter Verwendung von Matrizen und Vektoren über dem Halbring $(\text{Reg}(X), \cup, \cdot)$ zusammenfassen. Dazu definieren wir folgende Matrix:

$$C := (c_{ss'}) \quad \text{mit} \quad c_{ss'} := \{x \in X : \delta(s, x) = s'\} \quad \text{für alle} \quad s, s' \in S.$$

Die Sprachen $c_{ss'}$ in dieser Matrix enthalten die Buchstaben, die den Zustand s in den Zustand s' überführen. Wir können nun schreiben:

$$L_s = \sum_{s' \in S} c_{ss'} L_{s'} + T_s$$

bzw. mit $L := (L_1, \dots, L_{|S|})$ und $T := (T_1, \dots, T_{|S|})$

$$L = CL + T. \tag{2.2}$$

Es sei hier nochmal darauf hingewiesen, dass auch diese Vereinigung disjunkt ist: Für jeden Startzustand s ist jeder Buchstabe x in maximal einem der $c_{ss'}$ enthalten.

Die Sprachfamilie L ist also die Lösung des linearen Gleichungssystems (2.2) und hat die Gestalt

$$L = C^*T := \bigcup_{n \in \mathbb{N}_0} C^n T.$$

Die Matrix C aus diesem Gleichungssystem kann auch dazu verwendet werden, alle Worte einer gewissen Länge k aus der Sprache L zu erzeugen. Da die $c_{ss'}$ nur einbuchstabile Worte enthalten, enthalten die Elemente von C^n (mit $C^n = (c_{ss'}^{(n)})$) nur n -fach konkatenierte einbuchstabile – also n -buchstabile – Worte. Mit $L_s^{(n)} := L_s \cap X^n$ und $L^{(n)} := (L_1^{(n)}, \dots, L_{|S|}^{(n)})$ folgt daraus:

$$L^{(0)} = T, L^{(n)} = CL^{(n-1)} \quad \text{und damit} \quad L^{(n)} = C^n T.$$

Aus der Matrix C errechnen wir nun die sogenannte Adjazenzmatrix A , mit der wir eine Aussage über die Anzahl der Worte in $L^{(n)}$ machen können. Es sei

$$A = (a_{ss'}) := (|c_{ss'}|), \quad \text{also} \quad a_{ss'} := |\{x \in X : \delta(s, x) = s'\}|.$$

Die Einträge von A geben an, wieviele Buchstaben den Zustand s in den Zustand s' überführen. Analog definieren wir den Vektor $\xi := |T|$, dessen Einträge entweder 1 oder 0 sind, je nachdem ob ein Zustand s akzeptierend ist oder nicht.

Einfache kombinatorische Überlegungen liefern sofort, dass die Anzahl der Worte in $L^{(n)} = C^n T$ gegeben ist durch

$$|L^{(n)}| = A^n \xi.$$

Dies werden wir in Kapitel 6.1 nocheinmal aufgreifen.

3 Matrizen

Wie im letzten Kapitel schon angeklungen ist, lassen sich einige Eigenschaften von Sprachen bequem über Matrizen handhaben. Daher hier einige wichtige Definitionen und Sätze aus der Matrixanalysis.

3.1 Matrixnormen

Definition 3.1. Eine Funktion $\| \cdot \| : \mathbb{C}^{n \times n} \rightarrow \mathbb{R}$ heißt *Matrixnorm* sofern sie für alle Matrizen $A, B \in \mathbb{C}^{n \times n}$ folgenden Axiomen genügt:

1. Nichtnegativität: $\|A\| \geq 0$, Positivität: $\|A\| = 0$ genau dann wenn $A = 0$,
2. Homogenität: $\|cA\| = |c| \|A\|$ für alle $c \in \mathbb{C}$,
3. Dreiecksungleichung: $\|A + B\| \leq \|A\| + \|B\|$ und
4. Submultiplikativität: $\|AB\| \leq \|A\| \|B\|$.

Satz 3.2. Die maximale Zeilensummennorm

$$\|A\|_{\infty} := \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}|$$

ist die durch die l_{∞} -Vektornorm induzierte Matrixnorm.

Wichtig ist für uns hier nur die Tatsache, dass es sich bei $\| \cdot \|_{\infty}$ um eine Matrixnorm handelt. Einen Beweis dafür findet man unter anderem in [Ho/Jo85], Satz 5.6.5.

Definition 3.3. Die Menge aller Eigenwerte einer Matrix $A \in \mathbb{C}^{n \times n}$ nennt man auch das *Spektrum* von A . Der *Spektralradius* $\varrho(A)$ einer Matrix A ist

$$\varrho(A) := \max\{|\lambda| : \lambda \text{ ist ein Eigenwert von } A\}.$$

Dies ist der Radius der kleinsten im Nullpunkt zentrierten Kreisscheibe die alle Eigenwerte von A enthält.

Satz 3.4. Sei $\|\cdot\|$ eine Matrixnorm auf $\mathbb{C}^{n \times n}$. Dann ist der Spektralradius

$$\varrho(A) = \lim_{k \rightarrow \infty} \|A^k\|^{\frac{1}{k}}.$$

Ein Beweis dieses Satzes findet man in [Ho/Jo85], Korollar 5.6.14.

3.2 Die Jordansche Normalform einer Matrix

Das Potenzieren von quadratischen Matrizen wird schnell sehr rechenaufwendig. Folgende Sätze bieten einen Ausweg.

Satz 3.5 (Jordansche Normalform). Zu jeder quadratischen Matrix $P \in \mathbb{C}^{n \times n}$ existieren Matrizen $S, J \in \mathbb{C}^{n \times n}$ mit $P = SJS^{-1}$. Die Matrix J hat dabei Jordansche Normalform, das heißt

$$J = \begin{pmatrix} \boxed{J_1} & & & 0 \\ & \boxed{J_2} & & \\ & & \ddots & \\ 0 & & & \boxed{J_k} \end{pmatrix} \quad \text{mit} \quad J_i = \begin{pmatrix} \lambda_i & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & \lambda_i \end{pmatrix}.$$

Matrizen der Form der J_i bezeichnet man als Jordanmatrizen.

Einen Beweis dieses Satzes findet man in Büchern über lineare Algebra, z. B. [Fi86], Anhang B.

Bemerkung. Die λ_i auf der Hauptdiagonalen von J sind die Eigenwerte von P . Sie erscheinen entsprechend ihrer Vielfachheit, jedoch können einem Eigenwert mehrere Jordanmatrizen verschiedener Ordnung zugeordnet sein. Die Summe der Ordnungen der entsprechenden Jordanmatrizen entspricht dann der Vielfachheit des Eigenwerts.

$$j_{kl}^{(n+1)} = \lambda^{(n+1)-(l-k)} \left(\binom{n}{l-k} + \binom{n}{l-1-k} \right) = \lambda^{(n+1)-(l-k)} \binom{n+1}{l-k}.$$

Dies ist von der in der Behauptung geforderten Form.

Um den Beweis abzurunden müssen wir noch nachweisen, dass diese Beziehung auch für $n = 0$ gilt: Auf der Hauptdiagonale ist $l - k = 0$, also ist $j_{kl}^{(0)} = \lambda^0 \binom{0}{0} = 1$. Für alle anderen $l - k$ ist $\binom{0}{l-k}$ per Definition = 0. Damit ist J^0 die Einheitsmatrix.

Damit haben wir eine geschlossene Form für die Potenzen von Jordanmatrizen gefunden. \square

3.3 Eigenschaften von speziellen Matrizen

Die folgenden Definitionen und Sätze werden wir verwenden, um einen Zusammenhang zwischen gewissen Eigenschaften der Zustandsübergangsfunktion δ (als Graph aufgefasst) und der zugehörigen Adjazenzmatrix herzustellen.

Definition 3.7. Eine quadratische Matrix $A \in \mathbb{R}^{n \times n}$ heißt *reduzibel* wenn entweder

1. $n = 1$ und $A = (0)$ oder
2. $n \geq 2$ und es gibt eine Permutationsmatrix $P \in \mathbb{R}^{n \times n}$ sowie ein $r \in \mathbb{N}$ mit $1 \leq r \leq n - 1$, so dass

$$P^T A P = \begin{pmatrix} B & C \\ 0 & D \end{pmatrix},$$

wobei $B \in \mathbb{R}^{r \times r}$ und $D \in \mathbb{R}^{n-r \times n-r}$ quadratische Matrizen sind.

Definition 3.8. Eine Matrix $A \in \mathbb{R}^{n \times n}$ heißt *irreduzibel* wenn sie nicht reduzibel ist.

Definition 3.9. Der *gerichtete Graph* von $A = (a_{ij}) \in \mathbb{R}^{n \times n}$ ist der gerichtete Graph mit n Knoten P_1, P_2, \dots, P_n , der die Kante von P_i nach P_j genau dann enthält, wenn $a_{ij} \neq 0$ ist. Er wird als $\Gamma(A)$ notiert.

Definition 3.10. Ein *gerichteter Pfad* γ in einem Graph Γ ist eine Folge von Kanten $(P_{i_1} P_{i_2}), (P_{i_2} P_{i_3}), (P_{i_3} P_{i_4}), \dots$ in Γ . Die *Länge* eines gerichteten Pfades ist die

Anzahl der in ihm aufeinanderfolgenden Kanten sofern sie endlich ist, ansonsten habe der Pfad die Länge unendlich. Ein *Zykel* ist ein gerichteter Pfad, dessen Start- und Endknoten identisch sind und der jeden anderen Knoten maximal einmal besucht. Ein Zykel der Länge 1 heißt *trivialer Zykel*.

Definition 3.11. Ein gerichteter Graph Γ heißt *stark verbunden* wenn zu jedem Paar verschiedener Knoten P_i, P_j in Γ ein gerichteter Pfad endlicher Länge existiert, der in P_i beginnt und in P_j endet.

Satz 3.12. Die Matrix A ist genau dann irreduzibel, wenn $\Gamma(A)$ stark verbunden ist.

Für einen Beweis siehe z.B. [Ho/Jo85], Theorem 6.2.14 und folgende.

Nichtnegative irreduzible Matrizen haben bezüglich ihrer Eigenwerte interessante Eigenschaften. Die folgenden Sätze werden alle in [Mi88], Kapitel 1.4 bewiesen.

Satz 3.13. Eine irreduzible, nichtnegative Matrix $A \in \mathbb{R}^{n \times n}$ hat einen reellen Eigenwert $r > 0$, so dass $r \geq |\lambda_i|$ für jeden Eigenwert λ_i . Zu dem Eigenwert r existiert ein Eigenvektor mit Einträgen > 0 .

Der Eigenwert $r = \rho(A)$ heißt *maximaler Eigenwert von A* , der zugehörigen Eigenvektor *maximaler Eigenvektor von A* .

Dieses Ergebnis läßt sich fast auf reduzible Matrizen übertragen, nur die Positivität (> 0) des Eigenwerts und des Eigenvektors kann man nicht voll erhalten:

Satz 3.14. Eine nichtnegative, quadratische Matrix A hat einen reellen Eigenwert $r \geq 0$, der mindestens so groß ist, wie der absolute Betrag jedes Eigenwerts von A . Zu dem Eigenwert r existiert ein korrespondierender, nichtnegativer Eigenvektor.

Satz 3.15. Der maximale Eigenwert einer irreduziblen und nichtnegativen Matrix ist eine einfache Nullstelle des zugehörigen charakteristischen Polynoms.

Satz 3.16. Sei A eine nichtnegative, irreduzible Matrix mit maximalem Eigenwert r . Dann existiert ein positiver Eigenvektor zu r und jeder weitere nichtnegative Eigenvektor ist ein skalares Vielfaches dieses Eigenvektors.

Man kann über die Verteilung der Eigenwerte irreduzibler und nichtnegativer Matrizen noch mehr sagen. Dazu zunächst eine Definition:

Definition 3.17. Sei A eine irreduzible, nichtnegative Matrix mit genau h verschiedenen Eigenwerten maximalen Betrags. Die Zahl h heißt *Index der Imprimitivität von A* oder einfach *Index von A* . Ist $h = 1$ nennt man die Matrix A auch *primitiv*, andernfalls *imprimitiv* oder *zyklisch*.

Die folgenden Sätze rechtfertigen den Begriff *zyklisch* für imprimitive Matrizen. Der zugehörige Beweis findet sich in [Mi88], Kapitel 3, Theorem 1.1 und Theorem 1.2.

Satz 3.18. *Sei A eine irreduzible, nichtnegative Matrix mit maximalem Eigenwert r und Index h . Seien $\lambda_1, \lambda_2, \dots, \lambda_h$ die Eigenwerte maximalen Betrages von A . Dann sind $\lambda_1, \lambda_2, \dots, \lambda_h$ die verschiedenen h -ten Wurzeln von r^h , bzw. die mit dem Faktor r skalierten h -ten Einheitswurzeln.*

Aber nicht nur die maximalen Eigenwerte sind regelmäßig auf einem Kreis angeordnet. Jeder Eigenwert ist einer Familie von h Eigenwerten zugeordnet, die regelmäßig auf einem Kreis um den Nullpunkt angeordnet sind:

Satz 3.19. *Das Spektrum einer nichtnegativen irreduziblen Matrix vom Index h ist invariant gegenüber einer Rotation um den Winkel $\frac{2\pi}{h}$, nicht jedoch gegenüber einer Rotation um einen positiven Winkel kleiner als $\frac{2\pi}{h}$.*

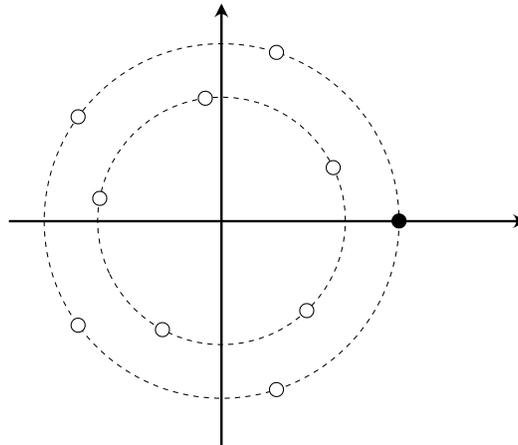


Abbildung 2: Eine mögliche Verteilung von Eigenwerten einer Matrix mit Index der Imprimitivität 5. Der markierte Punkt ist ein Eigenwert maximalen Betrages, es muss einer aus \mathbb{R} existieren.

4 Fraktale Mengen

Der Begriff einer *fraktalen Menge* ist mathematisch nicht eindeutig definiert. Kenneth Falconer beschreibt in [Fal90], dass eine Festlegung auf bestimmte Charakteristika sehr leicht viele interessante Mengen ausschließen würde.

Wenn man von fraktalen Mengen spricht, meint man üblicherweise Mengen F mit mehreren der folgenden (unscharfen) Eigenschaften:

1. F hat eine Feinstruktur, d. h. feine Details in beliebigen Vergrößerungsstufen.
2. F kann weder lokal noch global mit der klassischen Geometrie beschrieben werden.
3. Häufig hat F eine Form der Selbstähnlichkeit.
4. Die „fraktale Dimension“ (ein Sammelbegriff, es gibt mehrere nicht äquivalente Definitionen) ist größer als die topologische Dimension von F .
5. Meistens sind interessante F sehr einfach definiert, z. B. rekursiv.

4.1 Hausdorff Maß und Dimension

Für die eben angesprochene „fraktale Dimension“ gibt es mehrere – keineswegs äquivalente – Definitionen. Allen jedoch ist gemeinsam, dass sie eine Verallgemeinerung des klassischen Dimensionsbegriffs sind.

Der Weg zur Definition eines Dimensionsbegriffs führt über den Begriff des Maßes.

Definition 4.1. Es sei S eine beliebige Menge und Ω ein System von Teilmengen von S . Wir nennen μ ein Maß auf Ω , falls gilt:

1. Ω ist ein σ -Ring, d. h.

$$\emptyset \in \Omega \text{ und es gilt } E \in \Omega \implies S \setminus E \in \Omega,$$

$$E_i \in \Omega \text{ für } i \in \mathbb{N} \implies \bigcup_{i \in \mathbb{N}} E_i \in \Omega.$$

2. $\mu : \Omega \rightarrow [0, \infty]$ mit $\mu(\emptyset) = 0$ ist σ -additiv, d. h. sind $E_i \in \Omega$, $i \in \mathbb{N}$, paarweise disjunkt, so ist

$$\mu \left(\bigcup_{i \in \mathbb{N}} E_i \right) = \sum_{i \in \mathbb{N}} \mu(E_i).$$

3. Ist $N \in \Omega$ mit $\mu(N) = 0$ und $E \subset N$, so ist auch $E \in \Omega$.

$\mu(E)$ heißt auch Maß der Menge E .

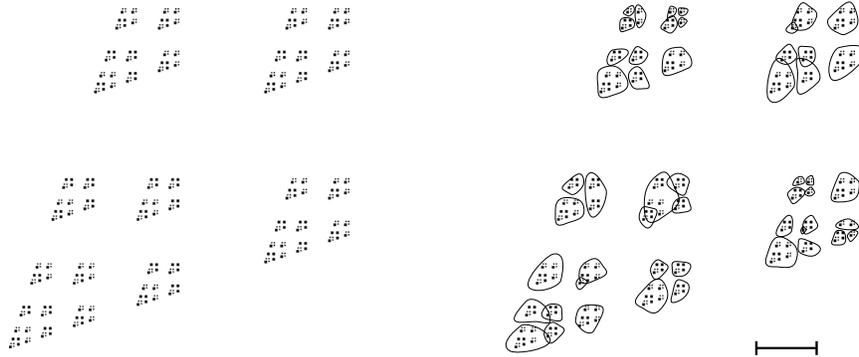


Abbildung 3: Eine Menge F zusammen mit einer δ -Überdeckung von F .

Die Familie der Hausdorff-Maße ordnet einer Teilmenge eines metrischen Raumes über δ -Überdeckungen ein Maß zu. Eine δ -Überdeckung ist eine Überdeckung mit abzählbar vielen Mengen U_i , die den maximalen Durchmesser $|U_i| \leq \delta$ haben. Dabei ist der Durchmesser einer Menge $|U|$ definiert als

$$|U| := \max\{|x - y| : x, y \in U\}.$$

Sei die Menge $F \subseteq \mathbb{R}^n$ und $s \in \mathbb{R}_{\geq 0}$. Für alle $\delta > 0$ sei

$$\mathcal{H}_\delta^s(F) := \inf \left\{ \sum_{i=0}^{\infty} |U_i|^s : \{U_i\} \text{ ist eine } \delta\text{-Überdeckung von } F \right\}.$$

Man überdeckt also F mit Mengen von maximalem Durchmesser δ und summiert die s -te Potenz der einzelnen Durchmesser. Dann bildet man das Infimum über alle derartigen Überdeckungen. Eine Grenzwertbildung führt dann zum Begriff des Hausdorff Maßes.

Definition 4.2. Das s -dimensionale Hausdorff Maß einer Menge F ist definiert als

$$\mathcal{H}^s(F) = \lim_{\delta \rightarrow 0} \mathcal{H}_\delta^s(F).$$

Dieser Grenzwert existiert im uneigentlichen Sinne (d. h. er kann den Wert ∞ annehmen), da die Menge der zulässigen Überdeckungen mit kleineren δ immer weiter eingeschränkt wird.

Nun kann man zeigen, dass es einen kritischen Wert $\dim_{\text{H}} F$ gibt, so dass

$$\mathcal{H}^s(F) = \begin{cases} \infty & \text{für } s < \dim_{\text{H}} F \\ 0 & \text{für } s > \dim_{\text{H}} F. \end{cases}$$

Diesen Wert bezeichnet man auch als *Hausdorff Dimension* der Menge F . Ist $s := \dim_{\text{H}} F$, so kann $\mathcal{H}^s(F)$ sowohl ein endlicher Wert ≥ 0 als auch ∞ sein.

Bemerkung. Veranschaulichen kann man sich diesen Sprung auch an dem klassischen Dimensionsbegriff. Ein zweidimensionales Quadrat hat das dreidimensionale Volumen 0. Wollte man es hingegen mit eindimensionalen Linien messen, so könnte man beliebig viele Linien hineinlegen, deren Länge aufsummiert gegen ∞ wachsen würde. Nur wenn man mit der „passenden Dimension“ misst, hat man eine Chance einen sinnvollen Wert herauszubekommen.

4.2 Boxcounting-Dimension

Die *Boxcounting-Dimension* ist im Vergleich zur Hausdorff Dimension häufig leichter zu berechnen. Es gibt mehrere äquivalente Formulierungen, von denen eine sehr praktisch für unsere Zwecke ist.

Definition 4.3. Sei F eine nichtleere beschränkte Teilmenge des \mathbb{R}^n und sei $N_\delta(F)$ die kleinstmögliche Anzahl von Mengen mit maximalem Durchmesser δ , die zusammen vereinigt eine Überdeckung von F bilden. Damit sind *untere* und *obere* Boxcounting-Dimension definiert als

$$\begin{aligned}\underline{\dim}_{\text{B}} F &= \liminf_{\delta \rightarrow 0} \frac{\log N_\delta(F)}{-\log \delta} \\ \overline{\dim}_{\text{B}} F &= \limsup_{\delta \rightarrow 0} \frac{\log N_\delta(F)}{-\log \delta}.\end{aligned}$$

Sind diese Grenzwerte gleich, so wird der gemeinsame Grenzwert als *Boxcounting-Dimension* von F bezeichnet:

$$\dim_{\text{B}} F = \lim_{\delta \rightarrow 0} \frac{\log N_\delta(F)}{-\log \delta}.$$

In dieser Definition ist das „kleinstmögliche Anzahl von Mengen“ etwas unbequem, da dies nicht trivial zu bestimmen ist. Eine der äquivalenten Formulierungen schränkt die Vielfalt der Überdeckungen ein und erleichtert uns die Berechnung drastisch.

Satz 4.4. Sei \mathbb{R}^n folgendermaßen in ein Gitter aus Würfeln der Kantenlänge δ unterteilt:

$$[m_1\delta, (m_1 + 1)\delta] \times \cdots \times [m_n\delta, (m_n + 1)\delta], \quad m_i \in \mathbb{Z}$$

Sei $N'_\delta(F)$ die Anzahl der Würfel aus diesem Gitter, die einen nichtleeren Schnitt mit der Menge F haben. Dann gelten folgende Beziehungen:

$$\begin{aligned}\underline{\dim}_B F &= \liminf_{\delta \rightarrow 0} \frac{\log N'_\delta(F)}{-\log \delta}, \\ \overline{\dim}_B F &= \limsup_{\delta \rightarrow 0} \frac{\log N'_\delta(F)}{-\log \delta} \quad \text{und} \\ \dim_B F &= \lim_{\delta \rightarrow 0} \frac{\log N'_\delta(F)}{-\log \delta} \quad \text{falls dieser Grenzwert existiert.}\end{aligned}$$

Beweis. Sei eine Überdeckung mit derartigen Würfeln gegeben, die Anzahl der Würfel sei $N'_\delta(F)$. Offensichtlich stellen sie eine $(\delta\sqrt{n})$ -Überdeckung der Menge F dar und $N_{\delta\sqrt{n}}(F) \leq N'_\delta(F)$. Ist $\delta\sqrt{n} < 1$ folgt

$$\frac{\log N_{\delta\sqrt{n}}(F)}{-\log(\delta\sqrt{n})} \leq \frac{\log N'_\delta(F)}{-\log \sqrt{n} - \log \delta}.$$

Grenzwertbildung für $\delta \rightarrow 0$ ergibt

$$\underline{\dim}_B F \leq \liminf_{\delta \rightarrow 0} \frac{\log N'_\delta(F)}{-\log \delta}$$

und

$$\overline{\dim}_B F \leq \limsup_{\delta \rightarrow 0} \frac{\log N'_\delta(F)}{-\log \delta}.$$

Andererseits ist aber jedes Element einer δ -Überdeckung in maximal 3^n aneinander liegenden Würfeln der Gitterzerlegung von \mathbb{R}^n enthalten, also ist $N'_\delta(F) \leq 3^n N_\delta(F)$. Nach Logarithmenbildung ergibt sich (für $\delta < 1$)

$$\frac{\log N'_\delta(F)}{-\log \delta} \leq \frac{\log 3^n}{-\log \delta} + \frac{\log N_\delta(F)}{-\log \delta}$$

und nach Grenzwertbildung ergeben sich die umgekehrten Ungleichungen wie oben (da der mittlere Term für $\delta \rightarrow 0$ gegen 0 konvergiert), weshalb Gleichheit gelten muss. \square

Es ist also nicht notwendig, alle δ -Überdeckungen zu betrachten, es reicht völlig aus, nur die oben beschriebenen gitterartigen Überdeckungen zu betrachten.

Die Boxcounting-Dimension ist weiterhin gegenüber verschiedenen äquivalenten Metriken zur Bestimmung des Durchmessers invariant. Zwei Metriken d_1, d_2 sind äquivalent, wenn Konstanten c und C existieren, so dass für alle $x, y \in \mathbb{R}^n$ gilt

$$c d_1(x, y) \leq d_2(x, y) \leq C d_1(x, y). \quad (4.1)$$

Satz 4.5. Seien d_1 und d_2 zwei äquivalente Metriken. Die Boxcounting-Dimension ist invariant gegenüber der Wahl der Metrik zur Bestimmung des Durchmessers von Mengen. Ist $\dim_{\mathbb{B}}^{(d_i)}(F)$ die Boxcounting-Dimension bezüglich der Metrik d_i , so gilt also:

$$\dim_{\mathbb{B}}^{(d_1)}(F) = \dim_{\mathbb{B}}^{(d_2)}(F).$$

Analoge Aussagen gelten auch für die obere und untere Boxcounting-Dimension.

Beweis. Aufgrund der Beziehung (4.1) zwischen zwei äquivalenten Metriken ist eine δ -Überdeckung bezüglich d_2 auch eine $C\delta$ -Überdeckung bezüglich d_1 und $N_{C\delta}^{(d_1)} \leq N_{\delta}^{(d_2)}$. Division durch $-\log(C\delta)$ (für hinreichend kleine δ) und Grenzwertbildung ergibt:

$$\begin{aligned} \dim_{\mathbb{B}}^{(d_1)}(F) &= \lim_{\delta \rightarrow 0} \frac{N_{C\delta}^{(d_1)}(F)}{-\log(C\delta)} \leq \lim_{\delta \rightarrow 0} \frac{N_{\delta}^{(d_2)}(F)}{-\log(C\delta)} \\ &= \lim_{\delta \rightarrow 0} \frac{N_{\delta}^{(d_2)}(F)}{-\log C - \log \delta} = \dim_{\mathbb{B}}^{(d_2)}(F). \end{aligned}$$

Mit der anderen Ungleichung aus der Äquivalenzbedingung ergibt sich analog die andere Richtung und damit $\dim_{\mathbb{B}}^{(d_1)}(F) = \dim_{\mathbb{B}}^{(d_2)}(F)$. Der Schluß für die obere und untere Boxcounting-Dimension geht analog. \square

5 Bilderzeugung mit Sprachen

Dieses Kapitel beschreibt, wie wir mit „Querschnitten“ durch eine präfixabgeschlossene Sprache L eine Bildfolge erzeugen und diese auf ihre geometrischen Eigenschaften untersuchen.

Mit *Bild* meinen wir hier Punktmengen im \mathbb{R}^d , im Folgenden beschränken wir uns auf Teilmengen des d -dimensionalen halboffenen Einheitswürfels $I := [0, 1)^d$.

5.1 Transformationen von Bildern

Als Teilmenge des Vektorraums \mathbb{R}^d kann ein Bild F wie üblich mit einem Skalar $\sigma \in \mathbb{R}$ multipliziert und durch Addition mit einem Vektor $\tau \in \mathbb{R}^d$ verschoben werden:

$$\sigma F + \tau := \{\sigma x + \tau : x \in F\}.$$

Inspiziert durch die Quadtree/Octree-Adressierung die in der zwei- bzw. dreidimensionalen Computergrafik verwendet wird, ordnen wir jedem Wort

$w \in X^*$ eine Abbildung φ_w zu, die Teilmengen des \mathbb{R}^d in Teilmengen des \mathbb{R}^d abbildet.

Wir beschränken uns auf Alphabete $X = \underbrace{Y \times \dots \times Y}_{d \text{ mal}}$ zu einem gegebenen r -elementigen Alphabet Y . Ohne Einschränkung sei $Y = \{0, 1, 2, \dots, r-1\}$. Zunächst definieren wir φ_x für alle $x = (y_1, \dots, y_d) \in X$:

$$\varphi_x(F) := r^{-1}(F + (y_1, \dots, y_d)^T)$$

Für diese Abbildung wird zunächst der Buchstabe $x \in X = Y^d \subset \mathbb{R}^d$ in den Vektorraum \mathbb{R}^d eingebettet. Ein Bild $F \subseteq I$ wird zunächst um den Vektor x verschoben und anschließend mit Faktor r^{-1} skaliert. Da $y_i \in \{0, 1, 2, \dots, r-1\}$ ist $F + (y_1, \dots, y_d)^T \subset [0, r]^d$ und nach der Skalierung ist $\varphi_x(F) \subset I$ für alle $x \in X$ und für alle $F \subseteq I$.

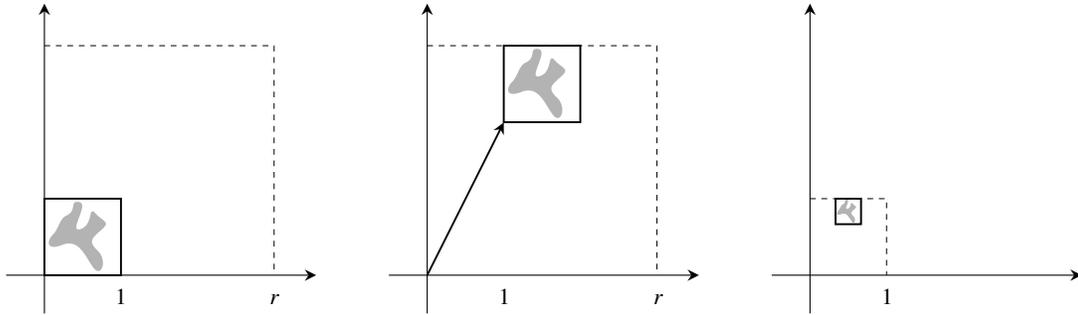


Abbildung 4: Von links nach rechts: Die Menge $F \subseteq I$, Translation um einen Vektor $(y_1, \dots, y_d)^T$, Skalierung mit dem Faktor r^{-1} .

In Abbildung 4 wird das Vorgehen an einem Beispiel mit $r = 3$ und $d = 2$ verdeutlicht. Eine Menge $F \subseteq I$ wird um den Vektor $(1, 2)^T$ verschoben, das resultierende Bild wird mit dem Faktor r^{-1} skaliert. Das Resultat ist wieder in I enthalten.

Wir erweitern nun diese Abbildungen auf Worte $w = x_1 x_2 \dots x_k \in X^*$ indem wir die den Buchstaben x_i zugeordneten Abbildungen miteinander verknüpfen:

$$\varphi_w := \varphi_{x_1} \circ \varphi_{x_2} \circ \dots \circ \varphi_{x_k}.$$

Da sowohl die Konkatination als auch die Operationen eines Vektorraums assoziativ sind folgt aus dieser Definition sofort für zwei beliebige Worte v, w :

$$\varphi_{vw} = \varphi_v \circ \varphi_w$$

Um eine explizite Darstellung der φ_w zu bekommen, definieren wir zunächst die Translationsvektoren τ_x für jeden in \mathbb{R}^d eingebetteten Buchstaben $x = (y_1, \dots, y_d)^T$.

$$\tau_x := \frac{1}{r}(y_1, \dots, y_d)^T.$$

Mit dieser Definition ist $\varphi_x(F) = r^{-1}(F) + \tau_x$. Induktiv können wir nun die Abbildung φ_w für alle $w \in X^*$ bestimmen.

Satz 5.1. *Sei das Wort $w = x_1x_2 \dots x_n \in X^*$. Dann ist*

$$\varphi_w(F) = \frac{1}{r^n}F + \sum_{i=1}^n \frac{1}{r^{i-1}}\tau_{x_i}.$$

Beweis. Diesen Satz beweisen wir per Induktion. Für $n = 1$ ist nichts zu zeigen.

Für $n > 1$ ist

$$\varphi_{wx}(F) = (\varphi_w \circ \varphi_x)(F) = \varphi_w(\varphi_x(F)).$$

Einsetzen der Behauptung für das Wort $w = x_1x_2 \dots x_n$ sowie der Definition von φ_x und $x_{n+1} := x$ ergibt:

$$\begin{aligned} \varphi_{wx}(F) &= \frac{1}{r^n} \left(\frac{1}{r}F + \tau_x \right) + \sum_{i=1}^n \frac{1}{r^{i-1}}\tau_{x_i} \\ &= \frac{1}{r^{n+1}}F + \sum_{i=1}^{n+1} \frac{1}{r^{i-1}}\tau_{x_i} \end{aligned}$$

Daraus folgt die Behauptung für das Wort wx . □

Um allen Worten $w \in X^*$ eine Abbildung φ_w zugeordnet zu haben definieren wir nun noch eine Abbildung für das bisher noch nicht berücksichtigte leere Wort ε :

$$\varphi_\varepsilon(F) := F.$$

Wir vereinbaren für ein Wort $w = x_1x_2 \dots x_n \in X^n$ folgende Bezeichnungen:

$$\begin{aligned} \tau_w &:= \sum_{i=1}^n \frac{1}{r^{i-1}}\tau_{x_i} \text{ und} \\ \sigma_w &:= \frac{1}{r^n} \end{aligned}$$

Damit können wir für alle $w \in X^*$ schreiben

$$\varphi_w(F) = \sigma_w F + \tau_w.$$

Wir zeigen nun, wie die Präfixrelation mit dieser Abbildungsfamilie im Zusammenhang steht. Wir wenden im Folgenden die Abbildung nicht mehr auf eine beliebige Menge F , sondern auf den d -dimensionalen Einheitswürfel $I = [0, 1]^d$ an. Dieser wird auch später zur Bilderzeugung verwendet.

Satz 5.2. Für jeden echten Präfix w eines Wortes $wv \in X^*$ und jede Menge $F \subseteq I$ gilt

$$\varphi_{wv}(F) \subseteq \varphi_{wv}(I) \subset \varphi_w(I).$$

Beweis. Die Inklusion $\varphi_{wv}(F) \subseteq \varphi_{wv}(I)$ folgt sofort aus der Definition von φ_{wv} und $F \subseteq I$. Es bleibt zu zeigen, dass $\varphi_{wv}(I) \subset \varphi_w(I)$.

Zunächst zeigen wir diese Beziehung für $v = x \in X$. Es gilt $\varphi_{wx}(I) = \varphi_w(\varphi_x(I))$ und somit

$$\varphi_{wx}(I) = \sigma_w \left(\frac{1}{r} I + \tau_x \right) + \tau_w \quad \text{und} \quad \varphi_w(I) = \sigma_w I + \tau_w.$$

Sei $x = (y_1, y_2, \dots, y_d) \in X, y_i \in Y$, wie am Anfang des Kapitels beschrieben und der zu x gehörige Translationsvektor sei τ_x . Da die Elemente von τ_x alle $\leq \frac{r-1}{r}$ sind, gilt für alle Vektoren $z \in I$:

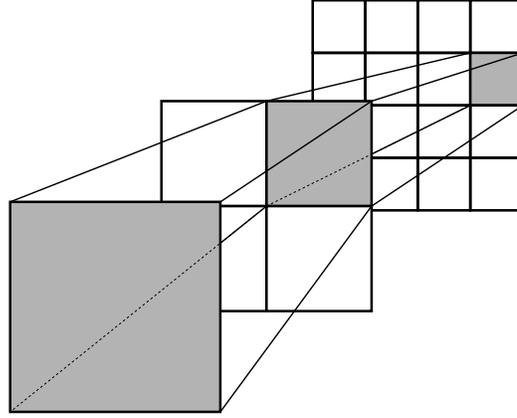
$$\frac{1}{r} z + \tau_x \in I.$$

Damit ist:

$$\begin{aligned} & \frac{1}{r} I + \tau_x \subset I \\ \Leftrightarrow & \sigma_w \left(\frac{1}{r} I + \tau_x \right) \subset \sigma_w I \\ \Leftrightarrow & \sigma_w \left(\frac{1}{r} I + \tau_x \right) + \tau_w \subset \sigma_w I + \tau_w \\ \Leftrightarrow & \varphi_{wx}(I) \subset \varphi_w(I). \end{aligned}$$

Mit dieser Beziehung ergibt sich dann eine Inklusionskette, die die Behauptung des Satzes sofort impliziert. \square

In Abbildung 5 ist veranschaulicht, wie die den verschiedenen Präfixen zugeordneten Bildern einander beinhalten. Von vorne nach hinten sehen wir $I = \varphi_\varepsilon(I)$, $\varphi_x(I)$ und $\varphi_{xy}(I)$. Die Zuordnung der Buchstaben zu einer Abbildung auf ein gewisses Teilquadrat ist hierbei willkürlich gewählt.

Abbildung 5: $I = \varphi_\varepsilon(I) \supset \varphi_x(I) \supset \varphi_{xy}(I)$

Im Folgenden wird ein Wort $w \in X^*$ sehr häufig dem d -dimensionalen Teilwürfel $\varphi_w(I)$ zugeordnet. Wir vereinbaren daher folgende abkürzende Schreibweise:

$$Q_w := \varphi_w(I) \text{ für alle } w \in X^*.$$

Satz 5.3. Seien $w = (w_1, \dots, w_k), v = (v_1, \dots, v_l) \in X^*$ zwei Worte. Falls der Durchschnitt der beiden Teilwürfel Q_w, Q_v nicht leer ist, dann ist ein Wort ein Präfix des anderen.

Beweis. Angenommen, w sei kein Präfix von v und v sei kein Präfix von w . Dann existiert ein Index i so dass $w_j = v_j$ für alle $j < i$ und $w_i \neq v_i$. Der i -te Buchstabe ist also der erste, der sich in den beiden Worten unterscheidet.

Da nach dem eben bewiesenen Satz $Q_w \subset Q_{w_1 \dots w_i}$ und $Q_v \subset Q_{v_1 \dots v_i}$, reicht es zu zeigen, dass bereits diese beiden Teilwürfel keine gemeinsamen Punkte mehr haben. Seien also die Worte w und v gleich lang und nur der letzte Buchstabe sei unterschiedlich. Sei u der maximale gemeinsame Präfix von w und v , also $w = ux$ und $v = uy$ mit geeigneten $x = (x_1, \dots, x_d) \in Y^d = X$ und $y = (y_1, \dots, y_d) \in Y^d = X$.

$$\begin{aligned} Q_w \cap Q_v &= \varphi_{ux}(I) \cap \varphi_{uy}(I) \\ &= \frac{1}{r^n} \left(\frac{1}{r} I + \tau_x \right) + \sum_{i=1}^n \frac{1}{r^{i-1}} \tau_{u_i} \cap \frac{1}{r^n} \left(\frac{1}{r} I + \tau_y \right) + \sum_{i=1}^n \frac{1}{r^{i-1}} \tau_{u_i} \\ &= \frac{1}{r^n} \left(\left(\frac{1}{r} I + \tau_x \right) \cap \left(\frac{1}{r} I + \tau_y \right) \right) + \sum_{i=1}^n \frac{1}{r^{i-1}} \tau_{u_i} \end{aligned}$$

Da die Buchstaben x und y unterschiedlich sind, existiert mindestens ein j , so dass $x_j \neq y_j$ (x_j, y_j sind dabei aus $Y = \{0, \dots, r-1\}$). Nach der Definition von τ_x und τ_y sind daher die j -ten Komponenten der Elemente aus $(r^{-1}I + \tau_x)$ aus dem halboffenen Intervall $[(\tau_x)_j, (\tau_x)_j + r^{-1})$, die entsprechenden Komponenten aus $(r^{-1}I + \tau_y)$ aus dem halboffenen Intervall $[(\tau_y)_j, (\tau_y)_j + r^{-1})$.

Nach der Definition der Translationsvektoren ist $|(\tau_x)_j - (\tau_y)_j| \geq r^{-1}$ und die beiden Intervalle überlappen sich nicht. Die j -te Komponente der Elemente aus Q_w und Q_v ist also immer unterschiedlich, der Schnitt dieser beiden Teilwürfel ist also leer.

Damit ist die Behauptung des Satzes bewiesen. \square

5.2 Bilder aus Sprachen

Wir ordnen nun einer Sprache L ein Bild $\varphi_L(I) \subset I = [0, 1)^d$ zu:

$$\varphi_L(I) := \bigcup_{w \in L} Q_w.$$

Für die leere Sprache $L = \emptyset$ ergibt dies das leere Bild \emptyset .

Satz 5.4. *Sei L eine präfixfreie Sprache. Das Bild $\varphi_L(I)$ ist dann eine disjunkte Vereinigung der den Worten $w \in L$ zugeordneten Bildern:*

$$\varphi_L(I) := \sum_{w \in L} Q_w.$$

Beweis. Angenommen es würden zwei Worte $w, v \in L$ mit $w \neq v$ existieren, so dass $Q_w \cap Q_v \neq \emptyset$. Laut Satz 5.3 ist dann entweder $v \sqsubset w$ oder $w \sqsubset v$. Dies ist ein Widerspruch zur Präfixfreiheit von L . \square

Satz 5.5. *Die Sprache X^k teilt mittels Q_w für alle $w \in X^k$ den Einheitswürfel I vollständig und regelmäßig in Teilwürfel der Seitenlänge r^{-k} auf.*

Beweis. Für $X^0 = \{\varepsilon\}$ ist $Q_\varepsilon = I$ die triviale Aufteilung in (einen) Teilwürfel. Induktiv schließen wir nun von X^k auf X^{k+1} . Es ist

$$Q_{xw} = \varphi_{xw}(I) = \frac{1}{r}(\varphi_w(I)) + \tau_x. \quad (5.1)$$

Da $X^{k+1} = X \times X^k$ ist, teilen die Würfel aus den Gleichungen (5.1) für ein $x \in X$ nach der Induktionsvoraussetzung den Würfel $Q_x = \frac{1}{r}I + \tau_x$ regelmäßig in Würfel der Kantenlänge $\frac{1}{r^{k+1}}$ auf. Wegen $X = Y^d = \{0, 1, \dots, r-1\}^d$ sind

die den Buchstaben $x = (y_1, \dots, y_d)^T$ zugeordneten Translationsvektoren $\tau_x = \frac{1}{r}(y_1, \dots, y_d)^T$ genau folgende Elemente:

$$\{\tau_x : x \in X\} = \frac{1}{r} \mathbb{Z}^d \cap I.$$

An jedem dieser Translationsvektoren wird nun ein Würfel der Kantenlänge $\frac{1}{r}$ positioniert, der wiederum regelmäßig in Würfel der Kantenlänge $\frac{1}{r^{k+1}}$ unterteilt ist. Damit wird I vollständig und regelmäßig in Teilwürfel der Kantenlänge $\frac{1}{r^{k+1}}$ zerlegt. \square

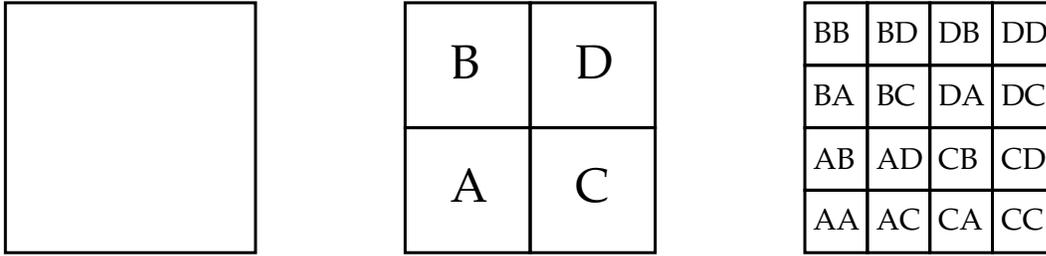


Abbildung 6: Die Zuordnung von Wörtern zu Teilquadraten

Satz 5.6. Sei L eine präfixabgeschlossene Sprache. Die Bildfolge $F_L^{(k)} := \varphi_{L \cap X^k}(I)$ bildet folgende Inklusionskette:

$$I = F_L^{(0)} \supset F_L^{(1)} \supset \dots \supset F_L^{(k)} \supset F_L^{(k+1)} \dots$$

Beweis. Für jeden Punkt $z \in \varphi_{L \cap X^{k+1}}(I) \subset \mathbb{R}^d$ existiert ein Wort $wx \in L \cap X^{k+1}$ mit $x \in X$ so dass $z \in Q_{wx}$. Da L präfixabgeschlossen ist, ist das Wort $w \in L$ und somit $w \in L \cap X^k$. Laut Satz 5.2 ist also $z \in Q_{wx} \subset Q_w$ und damit $z \in \varphi_{L \cap X^k}(I)$. Damit folgt die Behauptung. Die erste Identität gilt wegen $\varepsilon \in L$ und $F_L^{(0)} = \varphi_\varepsilon(I) = I$. \square

Diese Bilder $F_L^{(k)}$ sind quasi „Querschnitte“ durch die Sprache L . Da sie für eine präfixabgeschlossene Sprache eine absteigende Inklusionskette bilden, kann man den Grenzwert bzw. das Grenzbild sinnvoll definieren:

Definition 5.7. Das *Grenzbild* einer nicht-endlichen, präfixabgeschlossenen und regulären Sprache L ist definiert als:

$$F_L := \lim_{k \rightarrow \infty} F_L^{(k)} = \bigcap_{k \in \mathbb{N}} F_L^{(k)}.$$

Diese Grenzbilder F_L werden auch *rationale Fraktale* genannt.

Wir schließen hier Sprachen mit nur endlich vielen Worten aus, da ansonsten $F_L = \emptyset$ wäre (Es würde ein Wort maximaler Länge K existieren und damit wäre $F_L^{(K+1)} = \emptyset$).

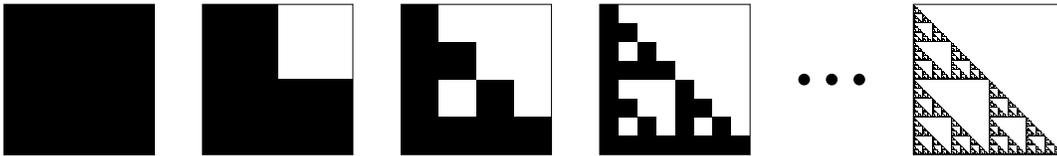


Abbildung 7: Von links nach rechts: $L^{(0)}, L^{(1)}, L^{(2)}, L^{(3)}, \dots, L^{(9)}$

Beispiel. Es sei $r = d = 2$. Wir benennen die Buchstaben eines Alphabetes $X = \{A, B, C, D\}$ und entnehmen der Einfachheit halber die Zuordnung zu Teilquadraten Q_A, \dots, Q_D der Abbildung 6.

Sei $\mathcal{A} = (X = \{A, B, C, D\}, S = \{1\}, \delta, s_0 = 1, F = \{1\})$, mit

$$\delta = \frac{\begin{array}{c|cccc} & A & B & C & D \\ \hline 1 & 1 & 1 & 1 & - \end{array}}$$

Dieser Automat hat nur akzeptierende Zustände, die durch ihn erkannte Sprache L ist also präfixabgeschlossen.

Mit $L^{(n)} := L \cap X^n$ ergibt sich

$$\begin{aligned} L^{(0)} &= \{\varepsilon\} \\ L^{(1)} &= \{A, B, C\} \\ L^{(2)} &= \{AA, AB, AC, BA, BB, BC, CA, CB, CC\} \\ &\text{etc.} \end{aligned}$$

In Abbildung 7 wurde diese Sprache in Bilder umgesetzt. Falls dem Leser die letzte Figur ($L^{(9)}$) bekannt vorkommt: Es handelt sich hierbei um eine Approximation des Sierpinski Dreiecks, einem Standard-Beispiel für fraktale Mengen².

²Eine kompakte Methode, reguläre Sprachen zu beschreiben sind die in vielen Programmiersprachen verwendeten *Regulären Ausdrücke*. Zeigt man jemandem, der mit diesem Werkzeug vertraut ist, den Ausdruck „ $[ABC]^*$ “ so ist derjenige meistens ziemlich verblüfft, wenn man ihm zeigt, wie etwas derartig simples so schöne fraktale Bilder beschreiben kann...

6 Die Spektralreihe

Sei eine präfixabgeschlossene, nicht-endliche Sprache L gegeben, die von einem endlichen Automaten \mathcal{A} erkannt werde. Die Zustandsmenge S des Automaten habe ausschließlich akzeptierende Zustände, der Startzustand sei s_0 . Wir setzen voraus, dass jeder andere Zustand von s_0 aus erreicht werden kann. Nicht erreichbare Zustände können einfach aus dem Automaten entfernt werden, ohne dass die erkannte Sprache L verändert wird.

Wir definieren L_s als die Sprache, die der Automaten \mathcal{A} erkennen würde, wenn er im Zustand s starten würde.

Die Bildfolge $F_{L_s}^{(k)}$ sowie das Grenzbild F_{L_s} sei wie im vorhergehenden Kapitel definiert.

Um diese Bilder quantitativ analysieren zu können müssen wir den zugrundeliegenden Automaten \mathcal{A} auf eine Weise beschreiben, die es erlaubt, bequem die Anzahl der gültigen Worte einer bestimmten Länge zu berechnen. Dazu betrachten wir das Spektrum einer Sprache (vergleiche Definition 2.7).

6.1 Die Elemente des Spektrums

Eine Möglichkeit die Elemente $g_{s,k}$ des Spektrums von L_s zu berechnen führt über die quadratische Adjazenzmatrix $A = (a_{ij})$ des Automaten \mathcal{A} mit Übergangsfunktion δ und Zustandsmenge S :

$$a_{ij} = a_{ij}^{(1)} = |\{x \in X : \delta(i, x) = j\}|, \quad 0 \leq i, j < |S|$$

Wir haben uns diese Matrix bereits in Kapitel 2.3 angesehen, hier kommt eine detaillierte Betrachtung.

Das Element a_{ij} ist die Anzahl der verschiedenen Übergänge von Zustand i nach j in einem Schritt. Per Induktion ergibt sich damit die Anzahl $a_{ij}^{(k)}$ der verschiedenen Übergänge von i nach j in genau k Schritten als

$$a_{ij}^{(k)} = \sum_{v=1}^n a_{iv}^{(k-1)} a_{vj}$$

und damit

$$\left(a_{ij}^{(k)} \right) = A^k.$$

Ist s der Startzustand von \mathcal{A} läßt sich die Anzahl der gültigen Worte der Länge k also beschreiben als

$$|L_s \cap X^k| = e_s A^k \xi,$$

wobei e_s der zu s gehörige Einheitsvektor ist und der Vektor ξ angibt, welche Zustände akzeptierende Zustände sind: $\xi_t = 1$ falls t ein akzeptierender Zustand ist und $\xi_t = 0$ sonst.

Da jedes a_{ij} die Anzahl der Buchstaben angibt, die den Zustand i in den Zustand j überführen und bei einer partiell definierten Übergangsfunktion δ in jedem Zustand maximal $|X|$ Übergänge möglich sind, sind die Zeilensummen von A jeweils maximal $|X|$. Die Matrix

$$P = (p_{ij}) := \frac{1}{|X|} A$$

ist damit eine substochastische Matrix, das heißt jede Zeilensumme ist ≤ 1 . Gleichheit gilt, wenn δ vollständig definiert ist.

Mit der Matrix $P := \frac{1}{|X|} A$ kann man das Spektrum der Sprache L_s also beschreiben als

$$g_{s,k} = e_s P^k \xi. \quad (6.1)$$

Die Matrix $P^n = (p_{ij}^{(n)}) = \frac{1}{|X|^n} A^n$ kann man auch als die Matrix der n -Schritt Übergangswahrscheinlichkeiten von i nach j interpretieren. Startet der Automat im Zustand i und verarbeitet zufällig gebildete Worte der Länge n (bei einer Gleichverteilung der Buchstaben aus X), so befindet er sich am Ende des Wortes mit der Wahrscheinlichkeit $p_{ij} \in [0, 1]$ in Zustand j .

Analog zu Kapitel 2.3 können wir nun ein Gleichungssystem für die Spektren der Sprachen L_s aufstellen. Mit $g^{(k)} := (g_{1,k}, \dots, g_{|S|,k})$ gilt

$$g^{(k)} = P^k \xi \implies g^{(0)} = \xi \quad \text{und} \quad g^{(k+1)} = P g^{(k)}. \quad (6.2)$$

Dies zeigt, dass die Spektren der Sprachfamilie – wie zu erwarten war – nicht unabhängig voneinander sind.

6.2 Konvergenzradius der Spektralreihe

Definition 6.1. Die *Spektralreihe* ist die formale Potenzreihe

$$R_s(z) := \sum_{n=0}^{\infty} g_{s,n} z^n,$$

deren Koeffizienten dem Spektrum einer Sprache L_s entnommen sind.

Der Konvergenzradius r_s dieser Potenzreihe wird auch als *Rate* des Spektrums bezeichnet und beschreibt, wie sich das Verhältnis zwischen gültigen

und möglichen Worten für große Wortlängen verhält, anschaulich gesprochen, wie „schnell“ sich ein ursprünglich schwarzes Bild ausdünn.

Da die von uns betrachtete Sprache L_s präfixabgeschlossen ist, d. h. es existiert ein erkennender Automat \mathcal{A} mit ausschließlich akzeptierenden Zuständen ($\xi = (1, \dots, 1)^T$), können wir ausnutzen, dass die Matrixnorm $\|\cdot\|_\infty$ eine sehr große Ähnlichkeit zu der Beschreibung des Spektrums hat. Für die Übergangsmatrix $P \geq 0$ des Automaten \mathcal{A} gilt:

$$\|P\|_\infty = \max_{1 \leq i \leq n} e_i P \xi.$$

Satz 6.2. *Für einen Automaten \mathcal{A} mit ausschließlich akzeptierenden Zuständen und Übergangswahrscheinlichkeitsmatrix P existiert ein Startzustand s , so dass die zugehörige Spektralreihe R_s den Konvergenzradius $\varrho(P)^{-1}$ hat. Dies ist gleichzeitig eine untere Schranke für die Konvergenzradien der Spektralreihen, die den anderen Startzuständen zugeordnet sind.*

Beweis. Da $P \geq 0$ ist, gilt

$$\|P^k\|_\infty = \max_{1 \leq i \leq n} \sum_{j=1}^n |p_{ij}^{(k)}| = \max_{1 \leq i \leq n} e_i P^k \xi = \max_{1 \leq i \leq n} g_{i,k}$$

Die Folge der $\|P^k\|_\infty$ ist also eine obere Schranke für die Elemente $g_{i,k}$ des Spektrums jeden Startzustandes i . Gleichzeitig existiert für jedes $\|P^k\|_\infty$ ein Startzustand j , so dass $\|P^k\|_\infty = g_{j,k}$.

Damit ist für alle i

$$\limsup_{k \rightarrow \infty} |g_{i,k}|^{1/k} \leq \limsup_{k \rightarrow \infty} \|P^k\|_\infty^{1/k} = \lim_{k \rightarrow \infty} \|P^k\|_\infty^{1/k} = \varrho(P). \quad (6.3)$$

Nach Cauchy-Hadamard ist folglich $\varrho(P)^{-1}$ eine untere Schranke für den Konvergenzradius der Spektralreihe für jeden Startzustand i .

Es bleibt zu zeigen, dass ein Startzustand s existiert, für den in obiger Ungleichung Gleichheit gilt.

Da nur n – also endlich viele – verschiedene Startzustände existieren, muss ein Startzustand s existieren, für den die $(g_{s,k})^{1/k}$ unendlich viele der $\|P^k\|_\infty$ enthalten³. Damit ist $\varrho(P)$ ein Häufungspunkt der $(g_{s,k})^{1/k}$ zu L_s und es kann keinen größeren Häufungspunkt geben. Damit ist für den Startzustand s die Gleichheit in der Ungleichung (6.3) gegeben. \square

³Um diesen Sachverhalt anhand einer Variation von Hilberts Hotel zu beschreiben: Verteilt man unendlich viele Gäste auf endlich viele Zimmer, muss mindestens ein Zimmer existieren, das unendlich viele Gäste enthält. . .

Habe nun die Spektralreihe für einen geeigneten Zustand s den Konvergenzradius $\varrho(P)$. Ist s' ein Zustand, von dem aus in endlich vielen Schritten der Zustand s erreicht werden kann, können wir mittels Gleichung (6.2) schließen, dass der Konvergenzradius zu der Spektralreihe $R_{s'}$ ebenfalls $\varrho(P)$ ist:

Dass der Konvergenzradius von $R_{s'}$ größer oder gleich $\varrho(P)$ sein muss, haben wir in dem vorangehenden Satz bewiesen. Ist der Zustand s in j Schritten von s' aus erreichbar, wissen wir aufgrund von Gleichung (6.2), dass

$$g_{s',k+j} = e_{s'} P^j g^{(k)} = \sum_{i=1}^{|S|} p_{s'i}^{(j)} g_{i,k},$$

wobei insbesondere $p_{s's}^{(j)} > 0$ ist.

Da die Matrix P (und damit auch P^j) nur nichtnegative Einträge enthält, ist $g_{s',k+j} \geq p_{s's}^{(j)} g_{s,k}$ für alle $k \in \mathbb{N}$.

Ist $z \in \mathbb{C}$ ein Punkt, für den $R_s(z)$ divergiert, muss folglich auch $R_{s'}(z)$ divergieren. Der Konvergenzradius von $R_{s'}$ ist also kleiner oder gleich dem Konvergenzradius von R_s sein.

Also haben die Spektralreihen zu s und s' den gleichen Konvergenzradius. Da wir vorausgesetzt haben, dass von dem Startzustand s_0 jeder andere Zustand des Automaten erreicht werden kann, ist insbesondere auch der Konvergenzradius der Spektralreihe zum Startzustand s_0 (und damit der Konvergenzradius der Spektralreihe der Sprache L von der wir ausgegangen waren) gleich $\varrho(P)$.

6.3 Geschlossene Form für das Spektrum

Die Darstellung der $g_{j,k}$ in Gleichung (6.1) ermöglicht es, die Koeffizienten des Spektrums für einen gegebenen Automaten leicht auszurechnen. Das „leicht“ ist jedoch mit Vorsicht zu genießen, da die k -fache Matrizenmultiplikation sehr aufwendig ist. Die Jordansche Normalform (Satz 3.5) hilft uns aus dem Dilemma.

Sei die Matrix P die zu dem Automaten \mathcal{A} gehörende Matrix der Übergangswahrscheinlichkeiten und $J = \text{diag}(J_1, \dots, J_l)$ eine Jordansche Normalform mit zugehöriger Transformationsmatrix S , so dass $P = SJS^{-1}$. Die quadratischen Matrizen J_i seien Jordanmatrizen der Größe m_i mit nicht notwendigerweise paarweise verschiedenen λ_i auf den Hauptdiagonalen.

Wie wir in Satz 3.6 gesehen haben, haben die einzelnen Einträge in den J_i^k die Gestalt

$$\lambda_i^{k-l} \binom{k}{l}$$

Damit kann man das Spektrum mit geeigneten $c_{i,\nu}$ auf folgende Weise darstellen:

$$g_{s,k} = e_s P^k \xi = e_s S J^k S^{-1} \xi = \sum_{i=1}^l \sum_{\nu=0}^{m_i-1} c_{i,\nu} \lambda_i^{k-\nu} \binom{k}{\nu}.$$

Die $c_{i,\nu}$ ergeben sich dabei aus dem Einheitsvektor e_s , der den Startzustand beschreibt, dem Vektor der akzeptierenden Endzustände ξ , sowie der Matrix S , die J in P überführt. Die $c_{i,\nu}$ sind unabhängig von k , da jede Abhängigkeit von k in den Einträgen von J^k – also in den Binomialkoeffizienten und den Potenzen der Eigenwerte – aufgefangen wird. Sie lassen sich ausrechnen, indem man eine aus symbolischen Werten bestehende Matrix J' bestimmt, die beschreibt, wie sich die Einträge $\lambda_i^{k-l} \binom{k}{l}$ in J^k verteilen. Berechnet man dann $e_s S J' S^{-1} \xi$ und faßt die gleichen symbolischen Werte zusammen, so erhält man die $c_{i,\nu}$.

In Anhang A.2 findet sich ein Beispiel wie man die $c_{i,\nu}$ und die Elemente des Spektrums direkt berechnen kann.

7 Fraktale Dimension

Wir haben jetzt die notwendigen Werkzeuge zusammen, um die obere Boxcounting-Dimension von Bildern, die gemäß Kapitel 5 durch Sprachen beschrieben sind, zu bestimmen.

Sei L_s eine nicht-endliche, präfixabgeschlossene Sprache, die von einem endlichen Automaten \mathcal{A} mit ausschließlich akzeptierenden Zuständen und Startzustand s erkannt wird. Die Bildfolge $F_{L_s}^{(k)}$ sowie das Grenzbild F_{L_s} sei wie im Kapitel 5 definiert.

7.1 Obere Boxcounting-Dimension

Wir haben in Kapitel 5.2 gezeigt, dass das Grenzbild F_{L_s} in dem Einheitswürfel I enthalten ist und die Worte aus $L_s^{(k)}$ so auf regelmäßig verteilte Würfel der Kantenlänge r^{-k} in I abgebildet werden, dass die Bilder $F_{L_s}^{(k)}$ das Grenzbild F_{L_s} überdecken. Es liegt nahe, diese Aufteilung von \mathbb{R}^n für die Bestimmung der Boxcounting-Dimension zu verwenden.

Die Anzahl der F überdeckenden Würfel in $F_{L_s}^{(k)}$ ist gleich der Anzahl der Worte in $L_s^{(k)}$. Wir wissen aus dem letzten Kapitel, dass $|L_s^{(k)}| = e_s A^j \xi$ ist.

Die Kantenlänge von d -dimensionalen Würfeln ist bezüglich der Maximummetrik auch der Durchmesser dieser Würfel. Da die Maximummetrik zur gewöhnlichen Euklidischen Metrik äquivalent ist, können sie wir zur Bestimmung der Boxcounting-Dimension heranziehen.

$$\begin{aligned}
\overline{\dim}_B F &= \limsup_{j \rightarrow \infty} \frac{\log N_{r^{-j}}(F)}{-\log(r^{-j})} \\
&= \limsup_{j \rightarrow \infty} \frac{\log_r N_{r^{-j}}(F)}{j} \\
&= \limsup_{j \rightarrow \infty} \frac{\log_r e_k A^j \xi}{j} \\
&= \limsup_{j \rightarrow \infty} \frac{\log_r (r^d)^j e_k P^j \xi}{j} \\
&= \limsup_{j \rightarrow \infty} \frac{dj}{j} + \frac{\log_r e_k P^j \xi}{j} \\
&= d + \limsup_{j \rightarrow \infty} \log_r \sqrt[j]{g_{k,j}} \\
&= d + \log_r \limsup_{j \rightarrow \infty} \sqrt[j]{g_{k,j}} \\
&= d + \log_r \varrho(P) \\
&= d - \log_r \varrho(P)^{-1}
\end{aligned}$$

Hier tritt der enge Zusammenhang zwischen Konvergenzradius $r_s = \varrho(P)^{-1}$ und oberer Boxcounting-Dimension deutlich zutage. Es gilt

$$\overline{\dim}_B F = d - \log_r r_s.$$

Bemerkung. Die obere Boxcounting-Dimension ist gleich dem Spektralradius der Adjazenzmatrix A :

$$\overline{\dim}_B F = d + \log_r \varrho(P) = \log_r r^d \varrho(P) = \log_r \varrho(r^d P) = \log_r \varrho(A).$$

7.2 Beziehung zur Hausdorff-Dimension nach [Me/St94]

In dem Artikel „Fractals, Dimension, and Formal Languages“ von W. Merzenich und L. Staiger ([Me/St94]) wird die Hausdorff-Dimension von Teilmengen F eines metrischen Raumes X^ω berechnet. Hier nun ein kurzer Überblick

über die Beziehungen zwischen dem ersten Kapitel von [Me/St94] und der vorliegenden Arbeit.

[Me/St94] setzt die Punkte des abgeschlossenen d -dimensionalen Einheitswürfels I mit unendlich langen Worten β über einem Alphabet $X = Y^d := \{0, 1, \dots, r-1\}^d$ gleich. Die Menge aller unendlich langen Worte über X wird mit X^ω bezeichnet. Die Komponenten z_i eines Punktes $z = (z_1, \dots, z_d) \in I \subset \mathbb{R}^d$ werden r -adisch entwickelt, d. h. z_i wird dargestellt als

$$z_i = \sum_{j=1}^{\infty} \frac{1}{r^j} y_{ij}, \quad \text{wobei } y_{ij} \in Y.$$

Das zu z zugehörige unendlich lange Wort ist dann $\beta = x_1 x_2 x_3 \dots$ mit $x_j = (y_{1j}, y_{2j}, \dots, y_{dj}) \in X$.

Man beachte, dass diese Zuordnung nicht eindeutig ist. Wie auch im Dezimalsystem ($0.4999999999\dots = 0.5000000000\dots$) kann es für einen Punkt z mehrere r -adische Entwicklungen geben. Da wir in dieser Arbeit nur den halboffenen Einheitswürfel $I = [0, 1)$ betrachten, können wir eine Eindeutigkeit herstellen, indem wir die Worte $\beta = x_1 x_2 x_3 \dots$ mit $x_j = (y_{1j}, y_{2j}, \dots, y_{dj})$ ausschließen, für die ein i existiert, so dass die Folge y_{ij} für $j \rightarrow \infty$ mit einer unendlichen Folge von $r-1 \in Y$ endet.

Diese Beziehung zwischen unendlich langen Worten aus X^ω und Punkten z aus I ist mit unserer Bijektion zwischen endlichen Worten w und den Teilwürfeln Q_w eng verwandt. Ist nämlich w ein endlicher Präfix von β mit $|w| = n$, so ist $z \in Q_w$:

$$\begin{aligned} z &= \sum_{j=1}^{\infty} \frac{1}{r^j} x_j \\ &= \sum_{j=1}^n \frac{1}{r^j} x_j + \frac{1}{r^n} \underbrace{\left(\sum_{j=1}^{\infty} \frac{1}{r^j} x_{j+n} \right)}_{\in I} \\ &\in \sum_{j=1}^n \frac{1}{r^{j-1}} \tau_{x_j} + \frac{1}{r^n} I = \sigma_w I + \tau_w = Q_w. \end{aligned}$$

Dass $\sum_{j=1}^{\infty} \frac{1}{r^j} x_{j+n} \in I$ ist, folgt, indem wir die x_{j+n} nach oben mit $(r-1, r-1, \dots, r-1)^T$ abschätzen und den Grenzwert ($= (1, 1, \dots, 1)^T$) der resultierenden geometrischen Reihe ausrechnen. Wir haben eben Worte ausgeschlossen, die in mindestens einer Komponente mit einer unendlichen Folge des Buch-

stabens $r - 1 \in Y$ enden. Daher ist dieser Grenzwert komponentenweise echt größer als $\sum_{j=1}^{\infty} \frac{1}{r^j} x_{j+n}$, die Summe ist also in $I = [0, 1)^d$ enthalten.

Zusammen mit der Metrik

$$\varrho(\beta, \xi) := \inf\{r^{-|w|} : w \text{ ist gemeinsamer Präfix von } \beta \text{ und } \xi\}$$

bildet X^ω einen metrischen Raum, die offenen Kugeln in X^ω bezüglich ϱ sind Mengen der Form $w \cdot X^\omega$, wobei $w \in X^*$ ein endliches Wort ist, der Durchmesser dieser Kugeln ist $r^{-|w|}$.

Die *Entropie* einer Sprache $L \subseteq X^*$ ist definiert als (vergleiche [Me/St94]):

$$H_L = \begin{cases} 0 & \text{falls } L \text{ endlich} \\ \limsup_{n \rightarrow \infty} \frac{1}{n} \log_r |\{v : v \in L \text{ und } |v| = n\}| & \text{sonst,} \end{cases}$$

das δ -Limit von L als

$$L^\delta := \{\xi : \xi \in X^\omega \text{ und } \xi \text{ hat unendlich viele Präfixe in } L\},$$

und die *Sprache der endlichen Präfixe* einer Menge $F \subseteq X^* \cup X^\omega$ als

$$\mathbf{A}(F) := \{p : p \in X^* \text{ und } \exists w : (w \in F \text{ und } p \sqsubseteq w)\}.$$

Vergleicht man das δ -Limit und das Grenzbild F_L einer präfixabgeschlossenen Sprache L :

$$F_L := \lim_{k \rightarrow \infty} F_L^{(k)} = \bigcap_{k \in \mathbb{N}} F_L^{(k)}$$

folgt daraus zusammen mit der Inklusionskette $F_L^{(k+1)} \subseteq F_L^{(k)}$, dass jeder Punkt $z \in F_L$ in unendlich vielen (genauer gesagt in allen) der $F_L^{(k)}$ enthalten ist, das z zugeordnete Wort ζ also unendlich viele Präfixe in L hat und damit in L^δ enthalten ist. Ist umgekehrt ein Wort $\zeta \in L^\delta$, hat es unendlich viele Präfixe in L und aufgrund der Präfixabgeschlossenheit von L je einen Präfix in $L^{(k)}$ für alle $k \in \mathbb{N}_0$. Der ζ zugeordnete Punkt z ist also in dem Grenzbild F_L enthalten. Das Bild F_L und L^δ beschreiben also dieselbe Punktmenge.

Insbesondere entsprechen die Kugeln $w \cdot X^\omega$ genau unseren Q_w , der Durchmesser $\text{diam}(w \cdot X^\omega) = r^{-|w|}$ ist die Kantenlänge von $Q_w \subset \mathbb{R}^d$.

Eine r^{-n} -Überdeckung einer Menge $F \subset X^\omega$ ist eine Familie von Kugeln $(v \cdot X^\omega)_{v \in V}$, so dass $V \cdot X^\omega \supset F$. Die Kugeln dürfen dabei maximal den Durchmesser r^{-n} haben, dies ist äquivalent dazu, dass die Worte $v \in V$ mindestens die Länge n haben.

Mit anderen Worten: Jedes unendlich lange Wort $\zeta \in F$ hat einen Präfix $v \in V$, der ζ zugeordnete Punkt $z \in I \subset \mathbb{R}^d$ ist also in Q_v enthalten. Die Vereinigung $\bigcup_{v \in V} Q_v$ ist also auch eine δ -Überdeckung von F als Teilmenge des \mathbb{R}^d , wobei hier $\delta = r^{-n} \sqrt{d}$ wenn man die euklidische Metrik bzw. $\delta = r^{-n}$ wenn man die Maximummetrik verwendet um den Durchmesser der Q_v zu bestimmen.

Der Vergleich der Definition der Entropie H_L mit der Definition der oberen Boxcounting-Dimension zeigt, dass sie für unendliche Sprachen identisch definiert sind. In Lemma 8 von [Me/St94] wird gezeigt, dass für reguläre, abgeschlossene Bilder $F \subseteq X^\omega$ die Hausdorff-Dimension $\dim_{\text{H}} F = H_F$ ist.

Sei F_L das Grenzbild einer regulären, präfixabgeschlossenen Sprache L . Damit ist F_L regulär und abgeschlossen ($(\mathbf{A}(F_L))^\delta = F_L$) und es gilt,

$$\dim_{\text{H}} F_L = H_{F_L} = \dim_{\text{B}} F_L.$$

8 Der Maßvektor als Eigenvektor von P

In diesem Abschnitt bestimmen wir einen Maßvektor m als einen „Eigenvektor“ der Matrix der Übergangswahrscheinlichkeiten P . Hierbei steht das Wort „Eigenvektor“ in Anführungszeichen, da wir nicht nur Eigenvektoren im Sinne der Linearen Algebra als Lösung auffassen, sondern auch Vektoren, die in einem oder mehreren Einträgen den Wert ∞ enthalten können.

8.1 Boxcounting-Volumen

Jedes Bild der Bildfolge $F_{L_s}^{(k)}$ ist eine Überdeckung des Grenzbildes F_{L_s} . Jedes derartige Bild ist wiederum zusammengesetzt aus einer endlichen Anzahl disjunkter gleichgroßer Teilwürfel.

Wir nutzen diese Bilder um das Boxcounting-Volumen zur Dimension σ zu definieren.

Definition 8.1. Das *Boxcounting-Volumen* einer Menge $F_{L_s} \subseteq I$ ist

$$\mathcal{B}^\sigma(F_{L_s}) = \lim_{k \rightarrow \infty} \sum_{v \in L_s^{(k)}} |Q_v|^\sigma = \lim_{k \rightarrow \infty} r^{-\sigma k} |L_s^{(k)}|$$

Die überdeckenden Teilwürfel jeder Approximationsstufe werden also gezählt und mit $r^{-\sigma k}$ gewichtet.

Für das Boxcounting-Volumen gilt

$$\mathcal{B}^\sigma(\varphi_w(F)) = \left(\frac{1}{r^{|w|}}\right)^\sigma \mathcal{B}^\sigma(F).$$

Diese Beziehung gilt, da für jedes $L_s^{(k)}$ gilt

$$\varphi_w(\varphi_{L_s^{(k)}}(I)) = \varphi_{w \cdot L_s^{(k)}}(I) \supseteq \varphi_w(F)$$

und damit

$$\sum_{wv \in w \cdot L_s^{(k)}} |Q_{wv}|^\sigma = \sum_{wv \in w \cdot L_s^{(k)}} r^{-\sigma|wv|} = r^{-\sigma|w|} \sum_{v \in L_s^{(k)}} r^{-\sigma|v|} = r^{-\sigma|w|} \sum_{v \in L_s^{(k)}} |Q_v|^\sigma$$

Diese Skalierungseigenschaft überträgt sich auf das Boxcounting-Volumen.

Sei F_s das Bild das entsteht, wenn der betrachtete Automat im Zustand s startet. Dieses Bild ist die disjunkte Vereinigung seiner Teilwürfel, die wiederum um den Faktor $\frac{1}{r}$ skalierte Bilder zu gewissen Startzuständen sind. Die Adjazenzmatrix $A = (a_{ij})$ bestimmt, wie oft welches Bild verwendet wird.

Mit der eben erläuterten Skalierungseigenschaft des Boxcounting-Volumens gilt:

$$\begin{aligned} \mathcal{B}^\sigma(F_s) &= \sum_{x \in X} \mathcal{B}^\sigma(\varphi_x(F_{sx})) = \sum_{t \in S} a_{st} \mathcal{B}^\sigma\left(\frac{1}{r} F_t\right) = \sum_{t \in S} a_{st} r^{-\sigma} \mathcal{B}^\sigma(F_t) \\ &= \sum_{t \in S} p_{st} r^{d-\sigma} \mathcal{B}^\sigma(F_t) = \frac{1}{\varrho(P)} \sum_{t \in S} p_{st} \mathcal{B}^\sigma(F_t) \end{aligned}$$

beziehungsweise mit $m = (\mathcal{B}^\sigma(F_1), \dots, \mathcal{B}^\sigma(F_n))^T$ und $\lambda_{\max} := \varrho(P)$

$$\lambda_{\max} m = Pm \iff \varrho(A)m = Am \quad (8.1)$$

Der Maßvektor m erfüllt also die Eigenwertgleichung für P und λ_{\max} . Damit ist m allerdings nicht eindeutig bestimmt, da die skalaren Vielfachen von m diese Gleichung ebenfalls erfüllen.

In [Me/St94] machen die Autoren analoge Überlegungen für das Hausdorff-Maß. Der Maßvektor bezüglich des Hausdorff-Maßes ist aufgrund der Skalierungseigenschaft ebenfalls ein „Eigenvektor“ von P bezüglich λ_{\max} . Zusätzlich finden die Autoren Kriterien, die für das Hausdorff-Maß Eindeutigkeit herstellen. Leider lassen sich diese nicht auf das Boxcounting-Volumen übertragen.

8.2 Irreduzible Automaten

Sei zunächst die Adjazenzmatrix des Automaten irreduzibel, das heißt jeder Zustand kann von jedem anderen Zustand aus erreicht werden. Offensichtlich sind die Eigenvektoren der Matrix P zu dem Eigenwert λ_{\max} eine Lösung zu Gleichung 8.1. Da die Matrix P nichtnegativ ist und $\lambda_{\max} > 0$ ist, sind die zugehörigen Eigenvektoren skalare Vielfache eines echt positiven Vektors. Der Satz 3.16 stellt zudem sicher, dass außer diesen skalaren Vielfachen keine weiteren nichtnegativen Vektoren existieren können, die die Gleichung 8.1 erfüllen können. Dies rechtfertigt, dass wir hier von einem *Maßvektor* sprechen, obwohl wir nicht nachgewiesen haben, dass das Boxcounting-Volumen ein Maß ist.

In [Me/St94] zeigen die Autoren, dass der Maßvektor zum Hausdorff-Maß mindestens einen Eintrag ≥ 1 hat (Corollary 13) und dass für irreduzible Matrizen alle Einträge nach oben durch 1 beschränkt sind (Theorem 16). Leider läßt sich das nicht auf das Boxcounting-Volumen übertragen. Dennoch läßt sich in vielen Fällen der Maßvektor einfach ausrechnen, Details dazu findet man in Kapitel 9.

8.3 Allgemeine Automaten

Um den Maßvektor für reduzible Automaten bequem bestimmen zu können, müssen wir die Zustände des Automaten nach ihrer Erreichbarkeit sortieren. Zu jedem Zustand s_i bezeichne S_i die Menge der Zustände s_j , für die ein Wort $w \in X^*$ existiert, so dass $s_i w = s_j$. Der Zustand s_j kann also von s_i aus erreicht werden:

$$S_i := \{s \in S \mid \exists w \in X^* : s_i w = s\}.$$

Offensichtlich sind die S_i halb geordnet, denn aus $s_j \in S_i$ folgt sofort $S_j \subseteq S_i$. Zwei Zustände s_i, s_j sind genau dann stark miteinander verbunden (d. h. $s_i \in S_j$ und $s_j \in S_i$), wenn die zugehörigen Mengen der erreichbaren Zustände S_i, S_j gleich sind. Man kann so die Zustände des Automaten in Komponenten unterteilen, deren zugehörigen Zustände untereinander stark verbunden sind. Wir konstruieren so k Teilmengen von S bzw. Teilautomaten, deren Zustandsmengen stark miteinander verbunden sind und deren Adjazenzmatrizen A_i somit irreduzibel sind.

Da für alle Zustände innerhalb einer starken Zusammenhangskomponente die Menge der erreichbaren Zustände gleich ist, bezeichne im Folgenden

S_i die von der Zusammenhangskomponente A_i erreichbare Zustandsmenge. Ohne Einschränkung seien die A_i so angeordnet, dass aus $S_j \subseteq S_i$ (A_j ist also von A_i aus erreichbar) folgt, dass $i < j$. Es sei darauf hingewiesen, dass daraus folgt, dass A_i nicht von A_j aus erreichbar ist, da ansonsten A_i und A_j gemeinsam eine starke Zusammenhangskomponente bilden würden.

Die Adjazenzmatrix A eines Automaten kann daher durch geeignete Permutation der Zustände in folgende Form gebracht werden:

$$A = \begin{pmatrix} \boxed{A_1} & & & * \\ & \boxed{A_2} & & \\ & & \ddots & \\ 0 & & & \boxed{A_k} \end{pmatrix}.$$

Die Einträge unterhalb der A_i können nur 0 sein, da ein Wert ungleich 0 bedeuten würde, dass, im Widerspruch zu unserer Voraussetzung, ein A_i von einem A_j mit $i < j$ aus erreichbar wäre.

Den Maßvektor m bestimmen wir nun sukzessive von unten nach oben. Dazu benennen wir die Elemente von A und m folgendermaßen:

$$A = \begin{pmatrix} \boxed{A_1} & \boxed{R_{1,2}} & \cdots & \boxed{R_{1,k}} \\ & \boxed{A_2} & \cdots & \boxed{R_{2,k}} \\ & & \ddots & \vdots \\ 0 & & & \boxed{A_k} \end{pmatrix}, \quad m = \begin{pmatrix} m^{(1)} \\ \hline m^{(2)} \\ \hline \vdots \\ \hline m^{(k)} \end{pmatrix}.$$

Aus $Am = \varrho(A)m$ folgt nun unmittelbar:

$$\begin{aligned} A_k m^{(k)} &= \varrho(A) m^{(k)} & (8.2) \\ A_{k-1} m^{(k-1)} + R_{k-1,k} m^{(k)} &= \varrho(A) m^{(k-1)} \\ &\vdots \end{aligned}$$

$$A_i m^{(i)} + \sum_{j=i+1}^k R_{i,j} m^{(j)} = \varrho(A) m^{(i)} \quad (8.3)$$

Für sich betrachtet hat der Teilautomat A_k die Dimension $\varrho(A_k) \leq \varrho(A)$. Ist der maximale Eigenwert von A_k kleiner als der maximale Eigenwert von A , also $\varrho(A_k) < \varrho(A)$, gibt es nur die Lösungen $(0, 0, \dots, 0)^T$ und $(\infty, \infty, \dots, \infty)^T$. Da wir zur Dimension $\varrho(A)$ messen, messen wir mit einer „zu großen“ Dimension, das Maß zu jedem der Startzustände muss also 0 sein. Damit ist in diesem Fall dieser Teil des Maßvektors $m^{(k)} = (0, 0, \dots, 0)$.

Ist hingegen $\varrho(A)$ ein Eigenwert der irreduziblen Matrix A_k , können wir Gleichung (8.2) wie im vorhergehenden Abschnitt behandeln. Es existiert dann ein nichttrivialer Maßvektor $m^{(k)}$, der dann der A_k zugeordnete Teil von m ist.

Schrittweise lösen wir nun die Gleichungssysteme (8.3) von unten nach oben. Ist $\sum_{j=i+1}^k R_{i,j}m^{(j)} = 0$ so behandeln wir A_i analog zu A_k . Wenn andernfalls $\sum_{j=i+1}^k R_{i,j}m^{(j)} \neq 0$ ist, wird aus Gleichung (8.3) folgendes Gleichungssystem:

$$(A_i - \varrho(A)I)m^{(i)} = - \sum_{j=i+1}^k R_{i,j}m^{(j)}.$$

Dieses Gleichungssystem ist genau dann eindeutig lösbar, wenn die Determinante $\det(A_i - \varrho(A)I) \neq 0$ ist. Das wiederum ist genau dann der Fall, wenn $\varrho(A_i) < \varrho(A)$. Diese Lösung ist dann unser $m^{(i)}$. Ist $\varrho(A_i) = \lambda_{\max}$, so ist das Gleichungssystem klassisch unlösbar und wir müssen $m^{(i)} = (\infty, \infty, \dots, \infty)$ setzen.

Sobald mit diesem Algorithmus irgendwann der Wert ∞ in den Maßvektor eingetragen werden muss, muss man mit diesem Wert in der üblichen Weise umgehen: $\infty + C = \infty$ und $\infty * \alpha = \infty$ für alle $\alpha > 0$.

In Anhang A.1 ist ein Beispiel zu finden, wie anhand dieser Methode ein Maßvektor berechnet wird.

9 Jordansche Normalform für Dimension und Maß

Eine interessante Alternative zu der Bestimmung von Maß und Dimension eines Bildes F ist der Weg über die Jordansche Normalform der Matrix der Übergangswahrscheinlichkeiten P . Auch wenn dieser Ansatz in der hier dargestellten Form nicht so allgemein ist, stelle ich ihn doch vor, da er eine andere Sichtweise auf z. B. das Spektrum einer Sprache ermöglicht.

Für die folgenden Betrachtungen machen wir die große Einschränkung, dass der maximale Eigenwert $\varrho(A)$ der Adjazenzmatrix des betrachteten Automaten entweder der *einzig*e Eigenwert maximalen Betrags ist (die Adjazenz-

matrix also primitiv ist) oder für die anderen Eigenwerte $\lambda_i \neq \lambda_{\max}$ mit maximalem Betrag die entsprechenden $c_{i,l}$ gleich 0 sind (vergleiche Kapitel 6.3)⁴. Außerdem verlangen wir, dass $\varrho(A) > 0$. Ansonsten wäre A nilpotent, d. h. es gäbe ein N so dass $A^N = 0$. Der Automat beschrieb also eine endliche Sprache, deren zugeordnetes Bild F_L die leere Menge wäre.

Im Folgenden sei wieder $P = \frac{1}{|\Sigma|}A$ die Matrix der Übergangswahrscheinlichkeiten.

9.1 Konvergenzradius der Spektralreihe

Seien $g_{s,n}$ die Koeffizienten der Spektralreihe einer Sprache L_s .

Nach dem Quotientenkriterium für Potenzreihen ist der Konvergenzradius der Spektralreihe folgender Grenzwert, sofern er existiert:

$$r_s := \lim_{n \rightarrow \infty} \left| \frac{g_{s,n}}{g_{s,n+1}} \right| = \lim_{n \rightarrow \infty} \frac{g_{s,n}}{g_{s,n+1}} \quad \text{da } 0 \leq g_{s,k} \leq 1 \text{ für alle } k.$$

Wie in Kapitel 6.3 hergeleitet, können wir das Spektrum mit geeigneten $c_{i,l}$ darstellen als

$$g_{s,n} = \sum_{i=1}^k \sum_{l=0}^{m_i-1} c_{i,l} \lambda_i^{n-l} \binom{n}{l}$$

Sei nun $\lambda_{\max} = \varrho(P) > 0$ der maximale Eigenwert von P bzw. J und m_{\max} die maximale Größe der zu λ_{\max} gehörenden Jordanmatrizen.

Dann lässt sich obiger Grenzwert folgendermaßen berechnen:

$$\begin{aligned} r_s &= \lim_{n \rightarrow \infty} \frac{\sum_{i=1}^k \sum_{l=0}^{m_i-1} c_{i,l} \lambda_i^{n-l} \binom{n}{l}}{\sum_{i=1}^k \sum_{l=0}^{m_i-1} c_{i,l} \lambda_i^{n+1-l} \binom{n+1}{l}} \\ &= \lim_{n \rightarrow \infty} \frac{\frac{1}{\lambda_{\max}^{n+1} \binom{n}{m_{\max}}} \sum_{i=1}^k \sum_{l=0}^{m_i-1} c_{i,l} \lambda_i^{n-l} \binom{n}{l}}{\frac{1}{\lambda_{\max}^{n+1} \binom{n}{m_{\max}}} \sum_{i=1}^k \sum_{l=0}^{m_i-1} c_{i,l} \lambda_i^{n+1-l} \binom{n+1}{l}} \\ &= \frac{1}{\lambda_{\max}} \cdot \frac{\sum_{i=1}^k \sum_{l=0}^{m_i-1} c_{i,l} \lim_{n \rightarrow \infty} \frac{\lambda_i^{n-l} \binom{n}{l}}{\lambda_{\max}^n \binom{n}{m_{\max}}}}{\sum_{i=1}^k \sum_{l=0}^{m_i-1} c_{i,l} \lim_{n \rightarrow \infty} \frac{\lambda_i^{n+1-l} \binom{n+1}{l}}{\lambda_{\max}^{n+1} \binom{n}{m_{\max}}}} \end{aligned}$$

Für die Korrektheit des letzten Schritts müssen wir nachrechnen, dass die einzelnen Grenzwerte existieren und der Nenner des Bruchs ungleich Null ist.

⁴Diese Bedingung ist nicht so weit hergeholt wie es vielleicht klingt. Für präfixabgeschlossene Sprachen bin ich bisher nur auf Matrizen gestoßen, bei denen die Eigenwerte maximalen Betrags $\neq \lambda_{\max}$ in der Summendarstellung des Spektrums durch $c_{i,l} = 0$ neutralisiert wurden.

Für die Existenz der einzelnen Grenzwerte im Zähler müssen wir hier mehrere Fälle unterscheiden.

1. $|\lambda_i| < \lambda_{\max}, l \geq m_{\max}$:

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{\lambda_i^{n-l} \binom{n}{l}}{\lambda_{\max}^n \binom{n}{m_{\max}}} &= \frac{1}{\lambda_i^l} \lim_{n \rightarrow \infty} \left(\frac{\lambda_i}{\lambda_{\max}} \right)^n \cdot \frac{n!}{l!(n-l)!} \cdot \frac{m_{\max}!(n-m_{\max})!}{n!} \\ &= \frac{m_{\max}!}{l! \lambda_i^l} \lim_{n \rightarrow \infty} \left(\frac{\lambda_i}{\lambda_{\max}} \right)^n \cdot (n-m_{\max}) \cdot \dots \cdot (n-l+1) \\ &= \frac{m_{\max}!}{l! \lambda_i^l} \lim_{n \rightarrow \infty} \left(\frac{\lambda_i}{\lambda_{\max}} \right)^{\frac{n}{l-m}} (n-m_{\max}) \cdot \dots \\ &\quad \dots \cdot \left(\frac{\lambda_i}{\lambda_{\max}} \right)^{\frac{n}{l-m}} (n-l+1) \\ &= 0 \end{aligned}$$

2. $|\lambda_i| \leq \lambda_{\max}, l < m_{\max}$:

$$\begin{aligned} \lim_{n \rightarrow \infty} \frac{\lambda_i^{n-l} \binom{n}{l}}{\lambda_{\max}^n \binom{n}{m_{\max}}} &= \frac{1}{\lambda_i^l} \lim_{n \rightarrow \infty} \left(\frac{\lambda_i}{\lambda_{\max}} \right)^n \cdot \frac{n!}{l!(n-l)!} \cdot \frac{m_{\max}!(n-m_{\max})!}{n!} \\ &= \frac{m_{\max}!}{l! \lambda_i^l} \lim_{n \rightarrow \infty} \underbrace{\left(\frac{\lambda_i}{\lambda_{\max}} \right)^n}_{|\cdot| \leq 1} \cdot \frac{1}{(n-l) \cdot \dots \cdot (n-m_{\max}+1)} \\ &= 0 \end{aligned}$$

3. $\lambda_i = \lambda_{\max}, l = m_{\max}$:

$$\lim_{n \rightarrow \infty} \frac{\lambda_{\max}^{n-l} \binom{n}{m_{\max}}}{\lambda_{\max}^n \binom{n}{m_{\max}}} = \frac{1}{\lambda_{\max}^l}$$

Da wir vorausgesetzt haben, dass entweder λ_{\max} der einzige Eigenwert maximalen Betrags ist oder die anderen Eigenwerte maximalen Betrags wegen der $c_{i,l}$ zu dem Grenzwert nichts beitragen können, müssen wir keine weiteren Fälle betrachten.

Die einzelnen Grenzwerte im Nenner berechnen sich fast identisch:

$$\lim_{n \rightarrow \infty} \frac{\lambda_i^{n+1-l} \binom{n+1}{l}}{\lambda_{\max}^{n+1} \binom{n+1}{m_{\max}}} = \lim_{n \rightarrow \infty} \frac{\lambda_i^{n+1-l} \binom{n+1}{l}}{\lambda_{\max}^{n+1} \binom{n+1}{m_{\max}}} \cdot \frac{n+1}{n+1-m_{\max}}$$

Dies konvergiert offensichtlich gegen den entsprechenden Wert aus dem Zähler, da der Korrekturfaktor gegen 1 konvergiert.

Damit konvergieren Zähler und Nenner des Bruchs gegen denselben Wert und wir können das Ergebnis aus Satz 6.2 bestätigen: Für die Rate r_s gilt:

$$r_s = \frac{1}{\lambda_{\max}}.$$

9.2 Boxcounting-Volumen

Nachdem wir eben den zur Bestimmung der Boxcounting-Dimension des Fraktals notwendigen Konvergenzradius bestimmt haben bestimmen wir nun noch das *Boxcounting-Volumen* zu dieser Dimension mittels der Jordanschen Normalform.

Sei $\sigma = \overline{\dim}_B F$. In Definition 8.1 haben wir beschrieben, wie wir jede Box der Seitenlänge r^{-n} mit dem Volumen $(r^{-n})^\sigma$ messen können. Damit können wir den Approximationsschritten leicht ein Volumen zuordnen, indem wir dieses Volumen mit der zur Überdeckung nötigen Anzahl der Boxen multiplizieren. Da

$$F^{(n)} = \varphi_{L \cap X^n}(I) = \sum_{w \in L \cap X^n} Q_w$$

eine Überdeckung von F ist, ist die Anzahl der nötigen Boxen $= |L \cap X^n| = e_s A^n \xi$ und mit $\lambda_{\max} := \varrho(P)$ gilt:

$$r^{-n\sigma} e_s A^n \xi = r^{-n\sigma} r^{nd} g_{s,n} = (r^{d-\sigma})^n g_{s,n}$$

Wegen $\dim_B F = \sigma = d - \log_r \varrho(P)$ ist $d - \sigma = \log_r (\varrho(P))^{-1}$ und $r^{d-\sigma} = \varrho(P)^{-1} = \frac{1}{\lambda_{\max}}$.

Die geschlossene Form für das Spektrum macht es leicht, den Grenzwert für $n \rightarrow \infty$ auch ohne Iteration auszurechnen.

$$\lambda_{\max}^{-n} g_{s,n} = e_s S \left(\frac{1}{\lambda_{\max}} J \right)^n S^{-1} \xi \quad (9.1)$$

Die Einträge von $(\lambda_{\max}^{-1} J)^n$ sind entweder 0 oder in einer Matrix $(\lambda_{\max}^{-1} J_i)^n$ enthalten, wobei J_i eine Jordanmatrix zum Eigenwert λ_i ist. Die Jordanmatrixpotenzen enthalten unterhalb der Hauptdiagonalen nur die 0, die Elemente der Haupt- und Nebendiagonalen haben (siehe Satz 3.6) folgende Gestalt:

$$j_{kl}^{(n)} = \frac{\lambda_i^{n-(l-k)}}{\lambda_{\max}^n} \binom{n}{l-k}$$

Insbesondere sind die Einträge einer Diagonalen von $(\lambda_{\max}^{-1} J_i)^n$ identisch.

Damit können wir ohne weiteres den Wert des Maßvektors $S(\lambda_{\max}^{-1} J)^n S^{-1} \xi$ ausrechnen. Durch die vielen Nullen und identischen Einträge in $(\lambda_{\max}^{-1} J_i)^n$ bleiben typischerweise einfache Terme stehen.

Um nun den Maßvektor zu bestimmen, brauchen wir nur noch die Grenzwerte der Haupt- und Nebendiagonalelemente für $n \rightarrow \infty$ zu bestimmen, dazu ist eine Fallunterscheidung notwendig. Im Folgenden sei $j_{kl}^{(n)}$ ein nichttriviales Element der Matrix $(\lambda_{\max}^{-1} J_i)^n$

1. $|\lambda| < \lambda_{\max}$:

$$\begin{aligned} \lim_{n \rightarrow \infty} j_{kl}^{(n)} &= \lim_{n \rightarrow \infty} \frac{\lambda^{n-(l-k)}}{\lambda_{\max}^n} \binom{n}{l-k} \\ &= \lim_{n \rightarrow \infty} \underbrace{\left(\frac{\lambda}{\lambda_{\max}} \right)^n}_{|\cdot| < 1} \lambda^{k-l} \binom{n}{l-k} \\ &= 0 \end{aligned}$$

da der Binomialkoeffizient nur polynomial wächst, der Restterm jedoch exponentiell fällt.

2. $\lambda = \lambda_{\max}, l = k$:

$$\begin{aligned} \lim_{n \rightarrow \infty} j_{kl}^{(n)} &= \lim_{n \rightarrow \infty} \frac{\lambda_{\max}^n}{\lambda_{\max}^n} \binom{n}{0} \\ &= 1. \end{aligned}$$

da $\lambda_{\max} > 0$.

3. $\lambda = \lambda_{\max}, l > k$:

$$\begin{aligned} \lim_{n \rightarrow \infty} j_{kl}^{(n)} &= \lim_{n \rightarrow \infty} \frac{\lambda_{\max}^{n-(l-k)}}{\lambda_{\max}^n} \binom{n}{l-k} \\ &= \lambda_{\max}^{k-l} \lim_{n \rightarrow \infty} l - k \\ &= +\infty. \end{aligned}$$

Da $l > k$ existiert der Limes nur im uneigentlichen Sinne und ist $+\infty$. Dies ist für die Berechnung des Volumens kein Problem, da das Volumen ∞ nicht sinnlos ist.

Man beachte, dass auch hier eine Betrachtung anderer Eigenwerte maximalen Betrags nicht notwendig ist, da nach Voraussetzung entweder λ_{\max} der einzige derartige Eigenwert ist, oder die $c_{i,l}$ in Gleichung (9.1) die anderen Eigenwerte neutralisieren.

Nach der Bestimmung dieser Grenzwerte kann man den Maßvektor leicht ausrechnen.

Das Vorgehen läßt sich noch etwas vereinfachen, indem man in der Berechnung des Maßvektor die Jordansche Normalform J durch die Matrix J' ersetzt, deren Einträge j_{kl} die Grenzwerte von $j_{kl}^{(n)}$ für $n \rightarrow \infty$ sind. Damit werden alle Jordanmatrizen zu einem Eigenwert mit echt kleinerem Betrag als λ_{\max} durch Nullmatrizen ersetzt und die Jordanmatrizen zu dem maximalen Eigenwert werden durch Matrizen mit Einsen auf der Hauptdiagonalen und ∞ oberhalb der Hauptdiagonalen ersetzt. Damit läßt sich dann der Maßvektor in einem Schritt ausrechnen.

10 Ausblick

Um das Thema über diese Arbeit hinaus weiter zu vertiefen, bieten sich an verschiedenen Stellen Ansatzpunkte. Es wäre interessant, zu wissen, ob das Boxcounting-Volumen vielleicht ein σ -additives Maß ist und ob es zur Bestimmung des Boxcounting-Maßvektors vielleicht ähnlich einfache Kriterien gibt, wie für den Maßvektor des Hausdorff-Maßes.

Ein weiterer interessanter Punkt wäre nachzuweisen, dass die am Anfang von Kapitel 9 gemachte Voraussetzung vielleicht sowieso immer erfüllt ist – wie in einer Fußnote bereits erwähnt habe ich bisher noch kein Gegenbeispiel gefunden. Und (falls es doch Gegenbeispiele geben sollte) kann man vielleicht mehr über die Bilder herausfinden, wenn man die Matrizen P^h betrachtet, wobei h der Index der Matrix P ist?

Auch eine Betrachtung wert wären die Beziehungen zu anderen Sichtweisen auf die hier betrachteten Bilder. Stichworte hierzu wären Iterierte Funktionensysteme (IFS) sowie der Hutchinson-Operator.

Literatur

- [Me/St94] W. Merzenich, L. Staiger: *Fractals, Dimension, and Formal Languages* (Theoretical Informatics and Applications, vol. 28, N° 3-4, 1994, p. 361 bis 386)
- [Mi88] H. Minc: *Nonnegative Matrices* (Wiley, 1988)
- [Ho/Jo85] R. A. Horn, Ch. R. Johnson: *Matrix analysis* (Cambridge University Press, 1985)
- [Fal90] K. J. Falconer: *Fractal Geometry, Mathematical Foundations and Applications* (Wiley Chichester, 1990)
- [Mer96] W. Merzenich: *Algorithmen* (Skript zur Vorlesung, 1996)
- [Ku/Sa86] W. Kuich, A. Salomaa: *Semirings, Automata, Languages* (Springer, 1986)
- [Fi86] G. Fischer: *Lineare Algebra* (Vieweg, 1986)
- [Alt92] H. W. Alt: *Lineare Funktionalanalysis* (Springer, 1992)
- [MuPAD] The MuPAD Research-Group: *MuPAD – The Open Computer Algebra System* (University of Paderborn, <http://www.mupad.de>)
- [Me03] W. Merzenich: *Sammlung von Automatenbeschreibungen* (unveröffentlicht)

A Beispiel

A.1 Berechnung von Boxcounting-Dimension und Maßvektor

Hier nun ein ausführliches Beispiel, dass die in dieser Arbeit vorgestellten Techniken verwendet, um Dimension und Maßvektor eines Automaten zu berechnen. Rechenintensive Arbeiten wurden mit [MuPAD] durchgeführt, jedes andere Computeralgebrasystem kann dies aber auch leisten.

Der zugrundeliegende Automat ist dem letzten Beispiel von [Me/St94] entnommen. Es ist $X = Y^2$ mit $|Y| = 2$. Der Einfachheit halber bezeichnen wir die Buchstaben aus X mit A, B, C, D.

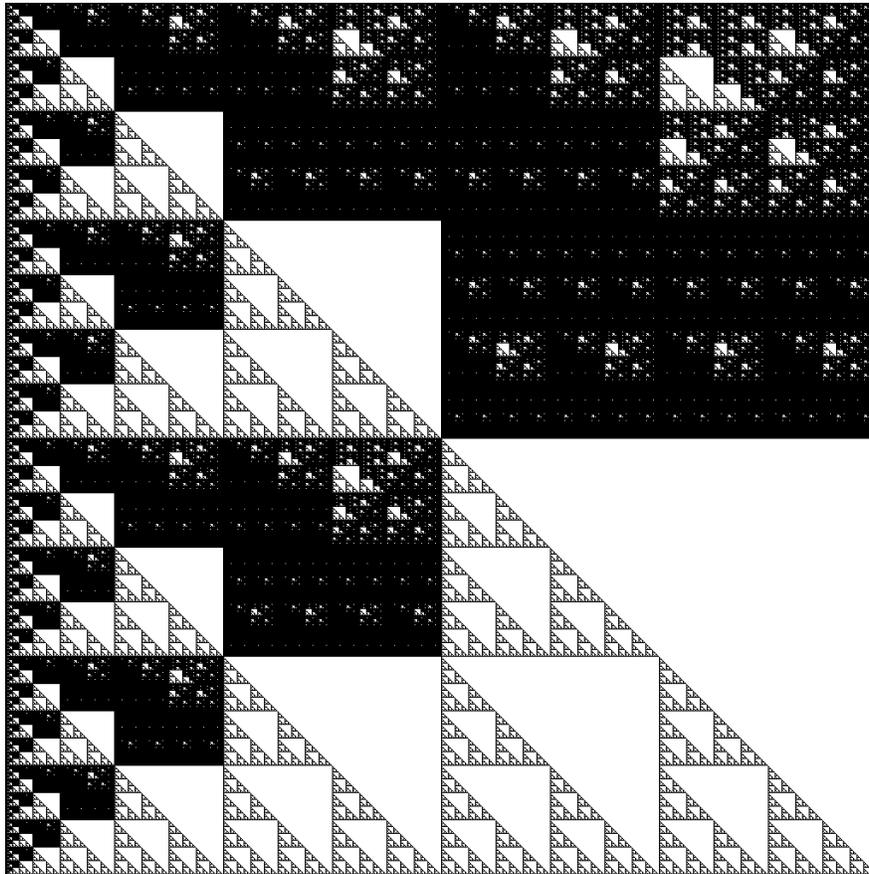


Abbildung 8: Das Bild des hier diskutierten Automaten.

Übergangstabelle und zugehörige Adjazenzmatrix sehen folgendermaßen aus:

$$\delta = \begin{array}{c|cccc} & A & B & C & D \\ \hline 1 & 1 & 1 & 7 & 2 \\ 2 & 3 & 2 & 3 & 4 \\ 3 & 2 & 3 & 2 & 3 \\ 4 & 4 & 5 & 4 & 4 \\ 5 & 7 & 6 & 5 & 6 \\ 6 & 6 & 6 & 5 & 5 \\ 7 & 7 & 7 & 7 & - \end{array}, \quad A = \begin{pmatrix} 2 & 1 & 0 & 0 & 0 & 0 & 1 \\ & 1 & 2 & 1 & 0 & 0 & 0 \\ & & 2 & 2 & 0 & 0 & 0 \\ & & & 3 & 1 & 0 & 0 \\ & & & & 1 & 2 & 1 \\ & & & & & 2 & 2 & 0 \\ 0 & & & & & & & 3 \end{pmatrix}$$

und der Startzustand ist Zustand 1. Die Zustände sind bereits nach ihrer Erreichbarkeit sortiert.

Der maximale Eigenwert von A ist

$$\lambda_{\max} := \varrho(A) = \frac{3 + \sqrt{17}}{2} \approx 3.56155$$

und die obere Boxcounting-Dimension des durch die Sprache L_1 beschriebenen Grenzbildes ist demnach

$$\dim_B(F_{L_1}) = \log_r(\lambda_{\max}) = \log_2\left(\frac{3 + \sqrt{17}}{2}\right) \approx 1.83251.$$

Nach dem Algorithmus aus Kapitel 8.3 bestimmen wir nun sukzessive von unten nach oben den Maßvektor m für den $\lambda_{\max}m = Am$ gilt. Die Zerlegung von A in die A_i und $R_{i,j}$ ist anhand der Linie in der Definition von A zu erkennen.

A_5 hat den einzigen Eigenwert $3 < \lambda_{\max}$. Damit ist die zugehörige Komponente in m gleich 0.

Die Eigenwerte von A_4 sind

$$\lambda_{1/2} = \frac{3}{2} \pm \frac{\sqrt{17}}{2},$$

insbesondere ist λ_{\max} ein Eigenwert von A_4 . Um $m^{(4)}$ zu bestimmen, müssen wir also einen Eigenvektor zu λ_{\max} finden. Für das Boxcounting-Volumen müssten wir an dieser Stelle auf das Verfahren aus Kapitel 9.2 zurückgreifen, um Eindeutigkeit für den Eigenvektor herzustellen. Wir bestimmen an dieser Stelle aber zunächst den Maßvektor für das Hausdorff-Maß gemäß [Me/St94].

Der Maßvektor für das Boxcounting-Volumen wird später bestimmt. Wir wählen also den Eigenvektor, dessen maximaler Eintrag gleich 1 ist, es ist

$$m^{(4)} = \begin{pmatrix} \frac{\sqrt{17}-1}{4} \\ 1 \end{pmatrix} \approx \begin{pmatrix} 0.78078 \\ 1 \end{pmatrix}.$$

Für $m^{(3)}$ ist folgendes Gleichungssystem zu lösen:

$$A_3 m^{(3)} + R_{3,4} m^{(4)} + R_{3,5} m^{(5)} = \lambda_{\max} m^{(3)}.$$

Einsetzen der bisher bekannten Werte ergibt:

$$\begin{aligned} \frac{3 + \sqrt{17}}{2} m^{(3)} &= (3) m^{(3)} + (1 \ 0) \begin{pmatrix} \frac{\sqrt{17}-1}{4} \\ 1 \end{pmatrix} + (3) \cdot 0 \\ \Leftrightarrow m^{(3)} &= \begin{pmatrix} \frac{\sqrt{17}-1}{2(\sqrt{17}-3)} \\ \end{pmatrix} \approx (1.39039). \end{aligned}$$

Die Teilmatrix A_2 ist die gleiche Matrix wie A_4 und hat natürlich ebenfalls λ_{\max} als Eigenwert. Allerdings müssen wir jetzt noch zusätzlich den bereits berechneten Maßvektor berücksichtigen:

$$A_2 m^{(2)} + \sum_{j=3}^5 R_{2,j} m^{(j)} = A_2 m^{(2)} + \begin{pmatrix} 1 \\ 0 \end{pmatrix} \begin{pmatrix} \frac{\sqrt{17}-1}{2(\sqrt{17}-3)} \\ \end{pmatrix} = \lambda_{\max} m^{(2)}$$

Da hier $\sum_{j=3}^5 R_{2,j} m^{(j)}$ ungleich 0 ist, ist dieses Gleichungssystem nicht zu lösen. Wir müssen also setzen

$$m^{(2)} := \begin{pmatrix} \infty \\ \infty \end{pmatrix}.$$

Die letzte fehlende Komponente des Maßvektors muss dann ebenfalls ∞ sein, da $m^{(2)}$ in der Gleichung berücksichtigt wird:

$$\begin{aligned} \lambda_{\max} m^{(1)} &= A_1 m^{(1)} + R_{1,2} m^{(2)} + R_{1,5} m^{(5)} \\ \Leftrightarrow (\lambda_{\max} - 2) m^{(1)} &= \infty + 0 \end{aligned}$$

Damit haben wir den Maßvektor $m = (\infty, \infty, \infty, 1.39039, 0.78078, 1, 0)^T$ bestimmt.

A.2 Alternativer Ansatz mittels Jordanform

Wie in den Kapiteln 6.3 und 9 beschrieben, kann die Jordansche Normalform der Matrix der Übergangswahrscheinlichkeiten P in manchen Fällen die Berechnung der Koeffizienten des Spektrums und des Maßvektors drastisch erleichtern.

Sei $P = \frac{1}{4}A$ die Matrix der Übergangswahrscheinlichkeiten des im vorhergehenden Abschnitt betrachteten Automaten. Es existieren nun Matrizen J und S , so dass $P = SJS^{-1}$. Mittels z. B. [MuPAD] ermitteln wir (numerisch, der Übersichtlichkeit halber die Nullen weggelassen):

$$J = \begin{pmatrix} \boxed{\begin{matrix} 0.75 & 1 \\ & 0.75 \end{matrix}} & & & & & & 0 \\ & \boxed{\begin{matrix} -0.1404 & 1 \\ & -0.1404 \end{matrix}} & & & & & \\ & & \boxed{\begin{matrix} 0.8904 & 1 \\ & 0.8904 \end{matrix}} & & & & \\ 0 & & & & \boxed{0.5} & & \end{pmatrix} \quad \text{und}$$

$$S = \begin{pmatrix} 0.0625 & 0.4219 & -0.0010 & -0.0051 & 0.0243 & -0.4949 & 0.08333 \\ 0.0625 & -0.3281 & 0.0025 & 0.0091 & 0.0379 & -0.6757 & \\ 0.125 & -1.1563 & -0.0020 & -0.0102 & 0.0485 & -0.9898 & \\ -0.125 & 1.9063 & & 0.0163 & & 0.4003 & \\ & -0.5 & & -0.0581 & & 0.2248 & \\ & -1 & & 0.0454 & & 0.2879 & \\ & 1 & & & & & \end{pmatrix}.$$

Wie wir in Kapitel 6.3 gesehen haben, können wir für geeignete $c_{i,v}$ das Spektrum $g_{s,k}$ folgendermaßen darstellen:

$$g_{s,k} = e_s P^k \xi = e_s S J^k S^{-1} \xi = \sum_{i=1}^l \sum_{v=0}^{m_i-1} c_{i,v} \lambda_i^{k-v} \binom{k}{v}.$$

Um die Koeffizienten $c_{i,v}$ zu bestimmen, ersetzen wir das J in der Formel durch eine Matrix J' mit symbolischen Werten in den Einträgen der Jordanmatrizen, die die Struktur der Potenzen $J^k = (j_{mn}^{(k)})$ widerspiegeln. Wir wissen nach Satz 3.6 die Einträge in den Potenzen der Jordanmatrizen nur von dem

zugrundeliegenden Eigenwert, der Potenz k und der Nummer der (Neben-) Diagonale $m - n$ abhängt, genauer:

$$j_{mn}^{(k)} = \lambda^{k-(n-m)} \binom{k}{n-m}.$$

Wir ersetzen nun in J' alle Werte von denen wir wissen, dass sie für alle $k \in \mathbb{N}$ in $J^{(k)}$ gleich bleiben durch die symbolischen Werte $l_{i,\nu}$, wobei $\nu = m - n$.

$$J' := \begin{pmatrix} \boxed{l_{1,0} \ l_{1,1}} & & & & 0 \\ & \boxed{l_{1,0}} & & & \\ & & \boxed{l_{2,0} \ l_{2,1}} & & \\ & & & \boxed{l_{2,0}} & \\ & & & & \boxed{l_{3,0} \ l_{3,1}} \\ & & & & & \boxed{l_{3,0}} \\ 0 & & & & & & \boxed{l_{4,0}} \end{pmatrix}.$$

Um die $c_{i,\nu}$ zu bestimmen können wir nun einfach $e_s S J' S^{-1} \xi$ ausrechnen und die Faktoren vor den $l_{i,\nu}$ ablesen. Mit $\xi = (1, 1, 1, 1, 1, 1, 1)^T$ ist

$$S J' S^{-1} \xi = \begin{pmatrix} 2.25 l_{1,0} + 0.0625 l_{1,1} - 0.0023 l_{2,0} - 0.0007 l_{2,1} - 0.7477 l_{3,0} + 0.1661 l_{3,1} - 0.5 l_{4,0} \\ 1.5 l_{1,0} + 0.0625 l_{1,1} + 0.0032 l_{2,0} + 0.0017 l_{2,1} - 0.5032 l_{3,0} + 0.2594 l_{3,1} \\ 2.5 l_{1,0} + 0.125 l_{1,1} - 0.0046 l_{2,0} - 0.0013 l_{2,1} - 1.4954 l_{3,0} + 0.3322 l_{3,1} \\ -1.75 l_{1,0} - 0.125 l_{1,1} + 0.0107 l_{2,0} + 2.7393 l_{3,0} \\ -0.5 l_{1,0} - 0.03824 l_{2,0} + 1.5382 l_{3,0} \\ -1.0 l_{1,0} + 0.02986 l_{2,0} + 1.9701 l_{3,0} \\ 1.0 l_{1,0} \end{pmatrix}$$

Die Multiplikation mit dem den Startzustand spezifizierenden e_s wählt eine dieser Zeilen aus.

Damit haben wir die $c_{i,\nu}$ bestimmt. Beispielhaft können wir nun ein Element des Spektrums der Sprache L_1 ohne k -fache Matrizenmultiplikation ausrechnen:

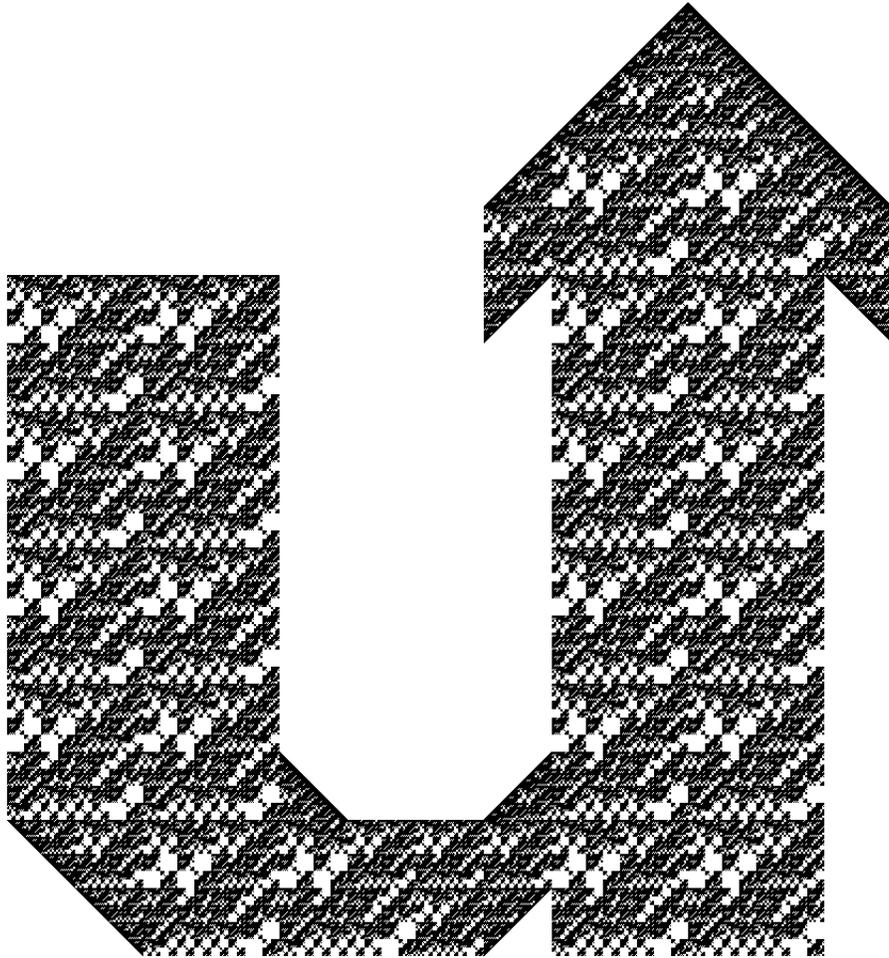
$$\begin{aligned} g_{1,172} &= 2.25 l_{1,0} + 0.0625 l_{1,1} - 0.0023 l_{2,0} - 0.0007 l_{2,1} \\ &\quad - 0.7477 l_{3,0} + 0.1661 l_{3,1} - 0.5 l_{4,0} \\ &= 2.25 \lambda_1^{172} + 0.0625 \lambda_1^{171} \binom{172}{1} - 0.0023 \lambda_2^{172} - 0.0007 \lambda_2^{171} \binom{172}{1} \\ &\quad - 0.7477 \lambda_3^{172} + 0.1661 \lambda_3^{171} \binom{172}{1} - 0.5 \lambda_4^{172} \\ &\approx 0.00000006664 \end{aligned}$$

A.3 Mit Automaten malen...

Mit etwas Erfahrung kann man Automaten so konstruieren, dass sie bestimmte Bilder darstellen. Die ernsthafte Beschäftigung mit diesem Thema (z. B.) im Rahmen der Bildkompression) ist nicht Thema dieser Arbeit, ich möchte aber einen von mir konstruierten Automaten vorstellen:

	A	B	C	D		A	B	C	D
1	2	3	4	5	18	34	18	34	31
2	6	7	8	9	19	25	-	-	-
3	10	-	11	-	20	-	-	26	-
4	12	7	9	9	21	-	-	-	27
5	13	20	15	19	22	29	26	29	29
6	-	-	17	29	23	-	17	-	-
7	-	-	29	29	24	29	29	29	25
8	29	29	29	19	25	29	25	25	-
9	29	29	-	-	26	26	-	29	26
10	-	-	29	-	27	27	29	-	27
11	29	-	-	-	28	29	29	27	29
12	28	20	29	29	29	31	18	30	29
13	21	14	29	29	30	18	33	34	32
14	-	-	29	26	31	30	29	18	32
15	29	29	23	16	32	33	34	30	33
16	29	25	-	-	33	31	31	34	33
17	-	17	17	29	34	-	33	32	-

Ich möchte das resultierende Bild allen Menschen widmen, die mit dazu beigetragen haben, dass ich an der Universität Gesamthochschule Siegen eine Unmenge von hochinteressanten – insbesondere auch fachfremden – Sachen lernen konnte. Herzlichen Dank!



Erklärung

Hiermit erkläre ich, dass ich diese Arbeit unter Verwendung der angegebenen Quellen und Hilfsmittel selbstständig angefertigt habe.

Siegen, den 12. 3. 2003