

Die Probabilistische Methode für zufällige Graphen

Martin Haufschild

3. September 2005

What you need is that your brain is open!

PAUL ERDÖS

1 Einleitung

Dieses Skript entstand im Rahmen eines vom Mathematik-Verein RHO in Zusammenarbeit mit der Universität Rostock organisierten Schüler-Wochenend-Seminars in Bad Doberan, das dort mehrmals jährlich zur mathematischen Förderung veranstaltet wird.

Mit dem Thema soll eine Zusammenführung zweier an für sich unverwandter mathematischer Gebiete verdeutlicht werden - der Graphentheorie und der Wahrscheinlichkeitsrechnung. An Methoden aus der Wahrscheinlichkeitsrechnung wird nicht viel mehr vorausgesetzt als aus der Sekundarstufe II bekannt ist. Aus der Graphentheorie liegen die Kenntnisse im Wesentlichen bereits durch andere Seminare vor.

Das Skript ist so aufgebaut, dass nach einer kurzen Aufgabe mit Lösung die Methoden selbstständig an weiteren Aufgaben trainiert werden sollen, was nicht heißen soll, dass keine weiteren Ideen notwendig wären.

Einen besonderen Dank möchte ich Herrn Prof. Konrad Engel und Herrn Dr. Thomas Kalinowski für ihre Mühe bei der Durchsicht dieses Skripts aussprechen.

Weitere Fehler oder Ideen bitte ich per Email an martinhaufschild@web.de zu senden. Viel Spaß mit den Aufgaben!

Zufällige Graphen

Die Theorie zufälliger Graphen hat sich in jüngerer Zeit als ein äußerst lebendiges Teilgebiet der Graphentheorie herauskristallisiert. Sie verbindet Methoden der Kombinatorik mit denen der Wahrscheinlichkeitstheorie und schlägt mit ihren Resultaten Brücken von der Diskreten Mathematik bis hin zu Teilgebieten der Theoretischen Informatik. So eröffnet die Theorie zufälliger Graphen beispielsweise die Möglichkeit, komplexe real existierende Strukturen wie Bekanntschaftsnetzwerke (insbesondere das *world wide web*) zu analysieren und zu simulieren.

Probabilistische Methode

Lange bevor zufällige Graphen als eigenständiger Forschungsgegenstand auftraten, wurden sie mit der probabilistischen Methode als elegantes Hilfsmittel für Existenz-

beweise genutzt.

Entwickelt wurde die probabilistische Methode von Alfred RÉNYI und Paul ERDŐS, einem der bedeutendsten Mathematiker des 20. Jahrhunderts. Beide etablierten die zufälligen Graphen um 1960 in einer Reihe wegweisender Arbeiten zu einem eigenständigen Forschungsgegenstand, nachdem ERDŐS herausfand, dass die probabilistische Methode nützlich zur Lösung extremaler Probleme in der Graphentheorie ist. Die Bezeichnung *Probabilistische Methode* ist aber insofern nicht korrekt, da sie eigentlich mehrere Methoden umfasst, beispielsweise folgende recht einfache Form

*Wenn auf einer Menge von Objekten die Wahrscheinlichkeit, dass ein Objekt eine bestimmte Eigenschaft nicht hat, kleiner als 1 ist, dann muss es ein Objekt mit der Eigenschaft geben.*¹

Die Grundidee ist also folgende: Um die Existenz eines Objektes mit einer bestimmten Struktur nachzuweisen, generiert man durch ein (geeignet abgestimmtes) Zufallsexperiment ein zufälliges Objekt und zeigt, dass die Wahrscheinlichkeit für die gewünschte Struktur des Objekts positiv ist.

Eine weitere häufig benutzte probabilistische Methode ist:

Es gibt eine Belegung der Zufallsgröße, deren Wert mindestens (höchstens) der Erwartungswert ist.

Die probabilistische Methode ist ein mächtiges Werkzeug in der Graphentheorie und der Kombinatorik, aber auch in der Zahlentheorie und der kombinatorischen Geometrie ist sie von großer Bedeutung. Die Anwendungen erstrecken sich jedoch auch über weite Gebiete der Informatik.

Anstatt des Abzählens wie in der Diskreten Mathematik häufig üblich, nutzt sie die Werkzeuge der Wahrscheinlichkeitsrechnung wie Wahrscheinlichkeit, Komplement und Vereinigung von Ereignissen, unabhängige Ereignisse, Erwartungswert, Varianz und noch viele andere. Erwähnt seien nur Martingale und Korrelationsungleichungen.

Am besten sieht man den eleganten Einsatz der Methoden an ein paar Aufgaben. Doch zuvor noch einige grundlegende Begriffe und Bezeichnungen, auf die wir später immer wieder zurückgreifen werden.

2 Begriffe und Bezeichnungen

2.1 Graphentheorie

Graphen

Ein *Graph* G ist ein geordnetes Paar (V, E) aus einer Menge V von *Knoten* (engl. *vertices*) und einer Menge E von *Kanten* (engl. *edges*) zwischen diesen Knoten, d.h. jede Kante $e \in E$ lässt sich auch durch die beiden von ihr verbundenen Knoten beschreiben.

Wir sprechen von einem *ungerichteten Graphen*, wenn² $E \subseteq \binom{V}{2}$ und damit $\{u, v\}$

¹Man sollte sich vor (den geistigen) Augen halten, dass sich die Wahrscheinlichkeit bei gleichwahrscheinlichen Elementarereignissen (Gleichverteilungsansatz) auch als Verhältnis der günstigen Möglichkeiten zu allen Möglichkeiten deuten lässt.

²Mit $\binom{V}{k}$, wobei $k \in \mathbb{N}$ und V eine Menge, bezeichnen wir alle k -elementigen Teilmengen von V .

für $u, v \in V$ ein ungeordnetes Paar ist, es also auf die Reihenfolge der Knoten nicht ankommt.

Dagegen gilt in einem *gerichteten Graphen* $E \subseteq V \times V$ und damit $(u, v) \neq (v, u)$ für $u, v \in V$, also sind die Elemente von E geordnete Paare und werden als *gerichtete Kanten* oder *Bögen* bezeichnet. Schlingen (u, u) werden meist nicht zugelassen.

Ein Graph heißt *vollständiger Graph*, wenn je zwei verschiedene Knoten u, v durch genau eine Kante verbunden sind. Ein vollständiger Graph mit n Knoten wird mit K_n bezeichnet. Wir werden Knoten der Einfachheit halber oft mit natürlichen Zahlen bezeichnen.

Schlichte Graphen

Ist mehreren Kanten oder Bögen dasselbe ungeordnete bzw. geordnete Paar von Knoten zugeordnet, so spricht man von *Mehrfachkanten* bzw. *Mehrfachbögen*. Eine Kante oder ein Bogen mit identischen Endpunkten heißt *Schlinge*. Graphen ohne Schlingen und Mehrfachkanten bzw. Mehrfachbögen werden *schlicht* genannt.

Im Folgenden werden wir nur schlichte Graphen betrachten ohne dies jedes Mal explizit zu erwähnen.

Untergraph

Sei $G = (V, E)$ ein Graph, dann ist $G' = (V', E')$ mit $V' \subseteq V$ und $E' \subseteq E$ ein *Untergraph von G* .

bipartiter Graph

Ein ungerichteter Graph $G = (V, E)$, dessen Knotenmenge sich in zwei disjunkte Mengen U und W zerlegen lässt und für den jede Kante einen Knoten in U und einen in W hat, wird ein *bipartiter Graph* genannt.

Turniergraph

Ein *Turniergraph* T ist ein vollständiger, gerichteter Graph, so dass für alle $u, v \in V$ entweder $(u, v) \in E$ oder $(v, u) \in E$ gilt. Die Bezeichnung ist insofern einleuchtend, da die Knoten Spieler und die zwischen ihnen liegenden gerichteten Kanten den Ausgang des Spiels darstellen.

Zufälliger Graph

Ein *zufälliger Graph* $G_{n,p}$ ist ein ungerichteter Graph mit n Knoten, in dem jede Kante zwischen zwei Knoten mit Wahrscheinlichkeit p existiert. Für $p = \frac{1}{2}$ ist jeder Graph gleichwahrscheinlich, d.h. wir erhalten die Gleichverteilung auf der Klasse aller Graphen.

Kantenfolgen

Jede Folge $F = ((v_1, v_2), (v_2, v_3), \dots, (v_s, v_{s+1}))$ mit $(v_i, v_{i+1}) \in E$ wird *Kantenfolge* der Länge s genannt. Eine Kantenfolge ist ein *Weg*, wenn v_1, v_2, \dots, v_{s+1} paarweise verschieden sind.

Hamilton-Weg

Ein *Hamilton-Weg* ist ein *Weg*, der jeden Knoten des Graphen genau einmal durchläuft. Dabei sind im Gegensatz zum *Hamilton-Kreis* Start- und Zielpunkt voneinander verschieden.

Färbungen

Wir sprechen von einem *knotengefärbten* Graphen, wenn wir jedem Knoten eine Farbe (natürliche Zahl) zuordnen. Analog ordnen wir bei einem *kantengefärbten Graphen* jeder Kante eine Farbe (natürliche Zahl) zu. Wir sprechen von einer n -Färbung, wenn wir dabei lediglich aus n Farben wählen können.

Meist ist es dabei von Interesse, dass im knotengefärbten Graphen zwei benachbarte Knoten und im kantengefärbten Graphen zwei benachbarte Kanten nicht gleichfarbig sind.

chromatischer Graph

Chromatische Graphen sind knoten- oder kantengefärbte Graphen. In einem monochromatischen Graphen haben alle Knoten bzw. Kanten dieselbe Farbe. Die *chromatische Zahl* bzw. der *chromatische Index* gibt an, wieviele Farben benötigt werden, damit keine benachbarten Knoten bzw. Kanten gleichfarbig sind.

2.2 Wahrscheinlichkeitsrechnung

Ergebnismenge

In einem beliebigen Zufallsexperiment ist die *Ergebnismenge* Ω die Menge aller vorkommenden *Ergebnisse* (Resultate). Beispielsweise erhalten wir bei einem Würfelwurf als Ergebnismenge die Menge der möglichen Seiten.

$$\Omega = \{S_1, S_2, S_3, S_4, S_5, S_6\}$$

Häufig wird ein Ergebnis auch als *Elementarereignis* bezeichnet.

Ereignis

Ein *Ereignis* im Sinne eines zufälligen Ereignisses ist eine Teilmenge von Ω . Im Gegensatz dazu ist das Elementarereignis ein Element von Ω . Ein Ereignis tritt ein, wenn eines der in ihm enthaltenen Elementarereignisse das Resultat des Zufallsexperiments ist.

Ferner bezeichnen wir Ω als *sicheres Ereignis* und \emptyset als *unmögliches Ereignis*. Zum Beispiel ist beim Würfel «gerade Augenzahl» ein Ereignis A mit $A = \{2, 4, 6\}$.

Zufallsgröße

Eine *Zufallsgröße* oder *Zufallsvariable* $X : \Omega \rightarrow \mathbb{R}$ ist entgegen ihrer Bezeichnung keine Variable sondern eine Funktion, die jedem Ergebnis (Elementarereignis) eine reelle Zahl zuordnet, z.B. bei einem Würfelwurf jeder Seite ihre Augenzahl durch $X(S_i) = i$, $i = 1, \dots, 6$.

Durch eine Zufallsgröße werden die einzelnen Ergebnisse erst vergleichbar, wie es beispielsweise für den Erwartungswert nötig ist.

Kann X nur endlich oder abzählbar viele Werte annehmen, so spricht man von einer *diskreten Zufallsgröße* andernfalls von einer *stetigen Zufallsgröße* wie zum Beispiel bei der idealen Messung einer physikalischen Größe. Wir werden uns hier jedoch auf diskrete Zufallsgrößen beschränken.

Wenn wir mit Zufallsgrößen nur beschreiben wollen, ob eine bestimmte Eigenschaft vorliegt oder nicht, wird das Ereignis «Eigenschaft vorhanden» mit 1 und «Ereignis nicht vorhanden» mit 0 belegt. Solche Zufallsgrößen (Zufallsvariablen) werden auch *Indikatorvariablen* genannt.

Manchmal ist es sinnvoll eine Zufallsgröße X in mehrere Zufallsgrößen X_i (meist

Indikatorvariablen) zu zerlegen, so dass jedes Ereignis in genau einem X_i enthalten ist. Es gilt also

$$X = \sum_i X_i.$$

Beispielsweise kann man die Zufallsgröße der Augensumme von zwei Würfeln in zwei Zufallsgrößen für die Augensumme je eines Würfels zerlegen. Umgekehrt ist für alle Zufallsgrößen X_i auch $X = \sum_i c_i X_i$ mit $c_i \in \mathbb{R}$ eine Zufallsgröße.

Wahrscheinlichkeit

Die Wahrscheinlichkeit P eines Ereignisses A gibt unter Verwendung des Gleichverteilungsansatzes (Bedingung ist LAPLACE-Verteilung: alle Elementarereignisse sind gleichwahrscheinlich) das Verhältnis von günstigen zu allen Elementarereignissen an.

$$P(A) = \frac{\text{Anzahl der für } A \text{ günstigen Elementarereignisse}}{\text{Anzahl der möglichen Elementarereignisse}}$$

mit $0 \leq P(A) \leq 1$.

Allgemeiner ist der axiomatische/maktheoretische Ansatz (KOLMOGOROW, 1933), der hier aber nicht benötigt wird, da wir uns auf Gleichverteilungen beschränken werden.

Häufig werden wir in Zusammenhang mit der Wahrscheinlichkeit von beliebigen Ereignissen die 1. Ungleichung von BONFERRONI (auch BOOLEsche Ungleichung genannt) benutzen:

$$P\left(\bigcup_{i=1}^n E_i\right) \leq \sum_{i=1}^n P(E_i),$$

wobei Gleichheit genau dann gilt, wenn die Ereignisse paarweise disjunkt sind.

Erwartungswert

Grob gesprochen ist der Erwartungswert einer Zufallsgröße jener Wert, der sich bei einer oftmaligen Wiederholung des zugrunde liegenden Experiments als Mittelwert der tatsächlichen Ergebnisse (Elementarereignisse) ergibt. Wenn die Zufallsgröße X diskret ist und die Werte x_1, x_2, \dots mit den entsprechenden Wahrscheinlichkeiten p_1, p_2, \dots annehmen kann, errechnet sich der Erwartungswert $E(X)$ als:

$$E(X) = \sum_i x_i p_i.$$

Beispielsweise ist der Erwartungswert beim Werfen eines Würfels

$$E(X) = 1 \cdot \frac{1}{6} + 2 \cdot \frac{1}{6} + 3 \cdot \frac{1}{6} + 4 \cdot \frac{1}{6} + 5 \cdot \frac{1}{6} + 6 \cdot \frac{1}{6} = 3,5.$$

Wenn, wie oben, die Zufallsgröße als Indikatorvariable je nach Vorhandensein einer Eigenschaft nur die Werte 0 oder 1 annehmen kann, so ist der Erwartungswert gerade der Wahrscheinlichkeit des durch den Indikator geprüften Ereignisses, also

$$E(X) = p.$$

Eine besondere Eigenschaft ist die *Linearität* des Erwartungswertes

$$E(kX + d) = kE(X) + d.$$

Für den Erwartungswert von n Zufallsgrößen gilt

$$E\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n E(X_i).$$

Insbesondere gilt $E(X + Y) = E(X) + E(Y)$ für zwei Zufallsgrößen X, Y .

Als Beispiel hierzu betrachten wir Fixpunktpermutationen. Sei X_i mit $1 \leq i \leq n$ die Indikatorvariable dafür, dass i ein Fixpunkt für eine gewählte Permutation σ von n Elementen ist, also $\sigma(i) = i$ gilt.

Es lässt sich leicht zeigen, dass $E(X_i) = \frac{1}{n}$ ist, denn nur für $\sigma(i) = i$ ist die Indikatorvariable $X_i = 1$. Dafür gibt es aber aufgrund der Permutation der anderen $n - 1$ Elemente $(n - 1)!$ Möglichkeiten. Somit ist $E(X_i) = \frac{(n-1)!}{n!} = \frac{1}{n}$.

Wir können dann den Erwartungswert für die Anzahl der Fixpunkte bei einer Permutation von n Elementen durch

$$E\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n E(X_i) = 1$$

bestimmen.

Für unsere Beweise werden für häufig die Tatsache nutzen, dass stets ein Elementarereignis vom Wert x mit $x \leq E(X)$ und stets auch eines vom Wert x mit $x \geq E(X)$ existiert. Andernfalls wäre der Erwartungswert kleiner bzw. größer als alle möglichen Werte der Zufallsgröße und das würde seiner Definition widersprechen.

3 Aufgaben

Im Folgenden werden wir an Sätzen die gängigen Methoden einführen, die dann in den Übungen angewendet werden sollen. Inhaltlich sind die Sätze aber unabhängig von den Übungen.

3.1 Linearität des Erwartungswertes

Der Beweis folgenden kombinatorischen Extremalproblems von SZELE (1943) wird oft als die erste Anwendung der probabilistischen Methode in der Diskreten Mathematik angesehen.

Satz 1. *Es gibt ein Turnier T aus n Spielern mit mindestens $\frac{n!}{2^{n-1}}$ Hamilton-Wegen.*

Beweis. In einem zufälligen Turnier sei X die Zufallsgröße der Anzahl der Hamilton-Wege. Um die möglichen Hamilton-Wege zu untersuchen, betrachten wir einfach alle Permutationen von Knoten und ordnen dann jeder solchen Permutation σ eine Indikatorvariable X_σ zu, die angibt, ob σ einen Hamilton-Weg erzeugt, also ob $(\sigma(i), \sigma(i+1))$ eine Kante für $1 \leq i < n$. Da jede der $n - 1$ Kanten mit Wahrscheinlichkeit $\frac{1}{2}$ in die richtige Richtung zeigt, ist $E(X_\sigma) = 2^{-(n-1)}$. Ferner wissen wir, dass $X = \sum X_\sigma$ ist und folglich gilt

$$E(X) = \sum E(X_\sigma) = \frac{n!}{2^{n-1}}.$$

Somit gibt es ein Turnier mit mindestens $E(X)$ Hamilton-Wegen. □

An diesem Beispiel sieht man, wie mächtig die probabilistische Methode sein kann, wenn man die Eigenschaften und die Zufallsgrößen nur geschickt wählt. Turniergraphen waren die ersten kombinatorischen Strukturen, die probabilistisch studiert wurden, weil sie eng mit der Statistik verbunden sind.

Hier noch zwei machbare Aufgaben ohne Beweis zur Anwendung des Prinzips:

Übung 1. *Es gibt eine 2-Kantenfärbung des K_n mit höchstens*

$$\binom{n}{a} 2^{1-\binom{a}{2}}$$

monochromatischen K_a .

Übung 2. *Sei $G = (V, E)$ ein Graph mit e Kanten. Dann enthält G einen bipartiten Untergraphen mit mindestens $\frac{e}{2}$ Kanten.*

3.2 Komplement und Vereinigung von Wahrscheinlichkeiten

Es geht aber auch ohne den Erwartungswert, was an der folgenden Aufgabe veranschaulicht werden soll.

Ein sehr prominentes Beispiel der probabilistischen Methode stammt aus dem Jahre 1947 und befasst sich mit RAMSEY-Zahlen.

Eine *Clique* (bzw. eine *unabhängige Menge*) ist ein Untergraph, der dadurch definiert ist, dass zwischen jedem Paar von Knoten eine (bzw. keine) Kante existiert. Frank RAMSEY hatte bereits 1930 bewiesen, dass zu je zwei natürlichen Zahlen s und t eine natürliche Zahl $R(s, t)$ existiert, so dass jeder Graph mit $n \geq R(s, t)$ Knoten eine Clique der Größe s oder eine unabhängige Menge der Größe t enthält.³ So ist $R(s, 2) = s$, denn entweder ist der Graph vollständig oder es gibt zwei Knoten, die nicht miteinander verbunden sind. Zwar ist es bisher noch nicht gelungen $R(s, t)$ allgemein zu berechnen, aber es konnten bereits untere und obere Schranken angegeben werden, wie dieses Ergebnis von ERDÖS zeigt:

Satz 2. *Für $k \geq 2$ gilt die folgende untere Schranke für die Ramsey-Zahlen*

$$R(k, k) \geq 2^{\frac{k}{2}}.$$

Beweis. Wir wissen schon, dass $R(2, 2) = 2$. Wenn wir in einem konvexen Fünfeck die Kanten nur auf dem Rand ziehen, haben wir auch sofort $R(3, 3) > 5$. Sei nun $k \geq 4$ und wir nehmen an, die Anzahl der Knoten sei $n < 2^{\frac{k}{2}}$. Für den allgemeinen Fall betrachten wir den zufälligen Graphen $G_{n, 1/2}$. Sei nun A eine feste k -elementige Knotenmenge. Die Wahrscheinlichkeit für das Ereignis A_c , dass A eine Clique bildet, ist $P(A_c) = 2^{-\binom{k}{2}}$. Folglich ist die Wahrscheinlichkeit p_c , dass *irgendeine* k -elementige Clique A existiert, durch

$$p_c = P\left(\bigcup_{|A|=k} A_c\right) \leq \sum_{|A|=k} P(A_c) = \binom{n}{k} 2^{-\binom{k}{2}}$$

beschränkt. Mit $n < 2^{\frac{k}{2}}$ und $k \geq 4$, und unter Verwendung von $\binom{n}{k} \leq \frac{n^k}{2^{k-1}}$ erhalten wir dafür die Abschätzung

$$\binom{n}{k} 2^{-\binom{k}{2}} < 2^{\frac{k^2}{2} - \binom{k}{2} - k + 1} = 2^{-\frac{k}{2} + 1} \leq \frac{1}{2}.$$

³Gelegentlich ist es auch üblich, sich einen vollständigen Graphen vorzustellen, in dem jede Kante eine von zwei Farben annimmt. Daraus ist auch sofort $R(s, t) = R(t, s)$ ersichtlich.

Somit ist $p_c < \frac{1}{2}$. Analog erhalten wir $p_s < \frac{1}{2}$ für die Wahrscheinlichkeit, dass eine k -elementige unabhängige Menge existiert. Daraus folgt dann $p_c + p_s < 1$ und mit der probabilistischen Methode gibt es demnach einen Graphen mit n Knoten, der weder eine Clique noch eine unabhängige Menge enthält. \square

Diese untere Schranke mag vielleicht etwas unscharf erscheinen, denn schon für $k = 3$ ist der Wert der unteren Schranke < 3 und $R(3, 3) = 6$. Doch für kein festes $\varepsilon > 0$ hat man eine untere Schranke der Form $R(k, k) > 2^{(\frac{1}{2} + \varepsilon)k}$ nachweisen können. Erstaunlich an diesem Ergebnis ist auch, dass es trotz energischer Versuche noch keinen konstruktiven Beweis gibt, der auch nur annähernd an die untere Schranke von ERDÖS herankommt. Dies zeigt, dass zufällige Graphen genau die Art von Struktur sind, die dieses und ähnliche Probleme brauchen, und das obwohl Zufall doch häufig mit Chaos gleichgesetzt und als Gegenteil von Ordnung angesehen wird.

Sei eine Grundmenge X von Punkten gegeben. Wir betrachten eine Familie \mathcal{F} von Teilmengen A_i von X mit jeweils genau $d \geq 2$ Punkten. Wir sagen, \mathcal{F} ist *2-färbbar*, wenn es eine Färbung der Punkte in X mit zwei Farben gibt, so dass in jeder Teilmenge A_i beide Farben auftreten.

Es ist ganz offensichtlich, dass man nicht jede Familie 2-färben kann. Beispielsweise ist eine Menge X aus $2k - 1$ Punkten mit $d = k$ und allen(!) d -Mengen nicht 2-färbbar, da es so immer d gleichfarbige Elemente gibt. Andererseits ist aber auch jede Teilfamilie einer 2-färbbaren Familie aus d -Mengen wieder 2-färbbar.

In Abhängigkeit von d interessiert uns nun folglich die *kleinste* Anzahl $m = m(d)$ von d -Mengen einer Familie, die nicht 2-färbbar ist.

Übung 3. *Jede Familie mit höchstens 2^{d-1} d -Mengen ist 2-färbbar, das heißt $m(d) > 2^{d-1}$.*

Beweis. Falls die d -Mengen disjunkt sind, können wir sie beliebig färben; also sind sie auch 2-färbbar. Sei \mathcal{F} eine Familie von d -Mengen, die aus höchstens 2^{d-1} Mengen besteht, von denen nicht alle disjunkt sind. Wir färben X mit zwei Farben, wobei alle Färbungen gleich wahrscheinlich sein sollen. Wir benutzen also feste Mengen und zufällige Färbungen. Für jede Menge $A \in \mathcal{F}$ sei E_A das Ereignis, dass die Elemente von A alle dieselbe Farbe bekommen. Da es allein für A betrachtet genau zwei solche Färbungen gibt, ist die Wahrscheinlichkeit ⁴

$$P(E_A) = \left(\frac{1}{2}\right)^{d-1}.$$

Mit $m = |\mathcal{F}| \leq 2^{d-1}$, wobei die Ereignisse E_A nicht disjunkt sind, gilt dann

$$P\left(\bigcup_{A \in \mathcal{F}} E_A\right) < \sum_{A \in \mathcal{F}} P(E_A) = m \left(\frac{1}{2}\right)^{d-1} \leq 1.$$

Mit der probabilistischen Methode schließen wir daraus, dass es eine 2-Färbung von X ohne einfarbige d -Mengen aus \mathcal{F} geben muss. \square

Anstatt alle Möglichkeiten von einfarbigen d -Mengen bei Färbungen abzählen zu müssen, konnten wir auf diese Weise den Beweis durch Zufallsgrößen viel einfacher erbringen. Während früher in der probabilistischen Methode häufig Abzählen benutzt wurde, ist man später zu Methoden mit Zufallsvariablen übergegangen.

⁴Wir sollten uns immer wieder ins Gedächtnis rufen, dass sich die Wahrscheinlichkeit unter gewissen Bedingungen der Gleichverteilung auch als Verhältnis der günstigen Möglichkeiten zu allen Möglichkeiten deuten lässt. Dann beschreibt diese Formel den Sachverhalt, dass auf je 2^{d-1} Färbungen von X genau eine kommt, bei der A einfarbig ist.

Der Nachteil der probabilistischen Methode ist lediglich der nicht-konstruktive Charakter, d.h. wir können aus unserem Beweis keine solche Familie konkret angeben. Auch im vorherigen Kapitel konnte nur mit Hilfe des angegebenen Beweises kein solches Turnier konstruiert werden.

Eine obere Schranke $m(d) \leq d^2 2^d$ wurde ebenfalls von ERDÖS erzielt, wieder mit Hilfe der probabilistischen Methode, wobei er dieses Mal zufällige Mengen und eine feste Färbung verwendet.

Betrachten wir eine neue Eigenschaft. Wir sagen ein Turniergraph hat die Eigenschaft S_k , wenn für je k Spieler ein anderer Spieler existiert, der sie alle schlägt.

Übung 4. Falls

$$\binom{n}{k} (1 - 2^{-k})^{n-k} < 1,$$

dann gibt es einen Turniergraphen mit n Knoten, der die Eigenschaft S_k besitzt.

Beweis. An der Bedingung ist eigentlich schon fast der Beweis abzulesen. Wir betrachten einen zufälligen Turniergraphen auf der Menge $V = \{1, \dots, n\}$. Für jede feste Teilmenge K von V der Größe k sei A_K das Ereignis, dass kein Knoten existiert, der alle anderen schlägt. Dann ist $P(A_K) = (1 - 2^{-k})^{n-k}$, denn die Wahrscheinlichkeit, dass ein fester Knoten $v \in V \setminus K$ nicht alle Knoten aus K schlägt, ist $(1 - 2^{-k})$. Für alle möglichen $v \in V \setminus K$ sind diese Wahrscheinlichkeiten voneinander unabhängig. Multiplizieren der Einzelwahrscheinlichkeiten ergibt den Wert für $P(A_K)$. Folglich ist dann

$$P\left(\bigcup_{|K|=k} A_K\right) \leq \sum_{|K|=k} P(A_K) = \binom{n}{k} (1 - 2^{-k})^{n-k} < 1.$$

Durch Komplementbildung ist demnach mit positiver Wahrscheinlichkeit kein A_K erfüllt, was soviel bedeutet wie, dass es einen Turniergraphen der Größe n gibt, der die Eigenschaft S_k besitzt. \square

3.3 Abstecher in die Zahlentheorie

Zum Abschluss soll an dieser Stelle noch ein Beispiel abseits der zufälligen Graphen und zwar aus der *Kombinatorischen Zahlentheorie* präsentiert werden. Eine Menge A wird als *summenfrei* bezeichnet, falls für alle $a_1, a_2, a_3 \in A$ gilt $a_1 + a_2 \neq a_3$. ERDÖS bewies 1965 dazu folgenden Satz:

Satz 3. Jede Menge $B = \{b_1, \dots, b_n\}$ aus von Null verschiedenen ganzen Zahlen enthält eine summenfreie Teilmenge A der Größe $|A| > \frac{n}{3}$.

Beweis. Zum Verstehen des Beweises sollte der Leser über die Modulo-Rechnung und den größten gemeinsamen Teiler (ggT) Bescheid wissen.

Sei $p = 3k + 2$ eine Primzahl mit $p > 2 \max_{1 \leq i \leq n} (b_i)$. Wir nehmen uns die Menge $C = \{k + 1, k + 2, \dots, 2k + 1\}$ zur Hilfe, die für die Addition mod p summenfrei ist. Wir erhalten

$$\frac{|C|}{p-1} = \frac{k+1}{3k+1} > \frac{1}{3}.$$

Nun wählen wir eine zufällige ganze Zahl x mit $1 \leq x < p$, wobei eine Gleichverteilung vorausgesetzt sei, d.h. jede Belegung von x ist gleichwahrscheinlich. Seien dann d_1, \dots, d_n definiert durch $d_i = x b_i \pmod{p}$ mit $0 \leq d_i < p$. Für festes i und variables x belegt d_i alle Zahlen $1, 2, \dots, p-1$, ist also wieder gleichverteilt. Dies lässt

sich einfach per indirektem Beweis zeigen: Sei $x_1 b_i \equiv x_2 b_i \pmod{p}$ mit $x_1 \neq x_2$. Äquivalent dazu ist $b_i(x_1 - x_2) \equiv 0 \pmod{p}$. Wegen $\text{ggT}(b_i, p) = 1$ können wir die Abtrennungsregel der Modulo-Rechnung anwenden und es folgt $x_1 - x_2 \equiv 0 \pmod{p}$, also ein Widerspruch zu $x_1 \neq x_2$.

Nach der Gleichverteilung von d_i bei Gleichverteilung von x ist deshalb die Wahrscheinlichkeit, dass bei festem i ein zufälliges d_i zu C gehört

$$P(d_i \in C) = \frac{|C|}{p-1} > \frac{1}{3}.$$

Die erwartete Anzahl der b_i , für die $d_i \in C$ gilt, ist damit größer als $\frac{n}{3}$. Da dies nur der Erwartungswert ist, muss es aber auch ein x und eine Teilmenge A von B der Größe $|A| > \frac{n}{3}$ geben, so dass $xa \pmod{p} \in C$ für alle $a \in A$ gilt. Dieses A ist natürlich summenfrei, denn falls $a_1 + a_2 = a_3$ für beliebige $a_1, a_2, a_3 \in A$, so auch $xa_1 + xa_2 \equiv xa_3 \pmod{p}$. Dies ist aber gerade ein Widerspruch dazu, dass C summenfrei \pmod{p} ist. \square

Es bleibt zu überlegen, ob es noch eine größere Konstante als $\frac{1}{3}$ gibt. ALON und KLEITMAN (1990) zeigten, dass sie nicht durch $\frac{12}{29}$ ersetzt werden kann. Die beste Konstante ist jedoch unbekannt.

4 Literatur

Besonders hervorzuheben sind das Standardwerk der zufälligen Graphen

[1] Bollobas, B.: Random Graphs. Cambridge University Press, 2. Aufl. 2001.

und vor allem das der probabilistischen Methode

[2] Alon, N., Spencer, J., Erdős, P.: The Probabilistic Method. Wiley Sons New York, 2. Aufl. 2000.

Einige schöne Beweise finden sich auch in folgendem Buch, das auch sonst mit äußerst schöner Mathematik gespickt ist

[3] Aigner, M., Ziegler, G.: Das Buch der Beweise. Springer, 2. Aufl. 2000.

Eine gut verständliche, prägnante und trotzdem recht tiefgehende Beschreibung der Thematik, die auch moderne Entwicklungen berücksichtigt und über die hier erwähnten Methoden hinausgeht – leider aber ohne Beweise – findet sich in

[4] Walz, G. (Hrsg.): Faszination Mathematik. Spektrum akademischer Verlag, 2003.