Het kleuringsgetal van Willekeurige Grafen

Laurens van der Beek

Inhoudsopgave

1 Inleiding

Willekeurige grafen zijn een relatief nieuw onderwerp binnen de Wiskunde. Het eerste moment waar zij ter sprake kwamen was in verschillende papers in 1959, toen er door twee afzonderlijke partijen een model voor dit type grafen geïntroduceerd werd. Als eerste werd door Paul Erdös en Alfréd Rényi het Erdös-Rényi model [6] opgesteld, wat inhield dat elke mogelijke willekeurige graaf op een bepaalde verzameling vertices met een bepaald aantal kanten evenveel kans heeft om voor te komen. Niet veel later werd door Edgar Gilbert ook een model opgesteld voor dit type grafen [7], maar dan in een andere vorm. In zijn model had namelijk elke kant een even grote, vaststaande kans om aanwezig te zijn. Later bleek dat dit model ongeveer op hetzelfde neerkwam als het Erdös-Rényi model, als de parameters juist gekozen worden, dit zullen we nu nader toelichten.

Het verwachte aantal kanten in een graaf met n vertices en kans p dat een kant tussen twee vertices aanwezig is (notatie: $G(n,p)$), is $\binom{n}{2}p$. Uit de Wet van Grote Getallen volgt dat elke graaf in $G(n,p)$ ongeveer zoveel kanten heeft, wanneer het verwachte aantal kanten tot oneindig nadert. Oftewel, als $pn^2 \to \infty$ dan gedraagt $G(n,p)$ (Gilbert‘s model) zich ongeveer hetzelfde als $G(n,M)$ (Erdös-Rényi‘s model), wanneer $M = \binom{n}{2}p$ (M staat hier voor het aantal kanten van de graaf). In deze scriptie zullen we (zoals tegenwoordig gebruikelijk is) gebruikmaken van het $G(n,p)$ model.

Naar het kleuringsgetal van een graaf wordt al langer gekeken dan naar willekeurige grafen. Het kleuringsgetal van een graaf is het minimaal aantal kleuren benodigd om een graaf te kleuren op zo‘n manier dat geen twee aan elkaar verbonden punten een zelfde kleur hebben. De eerste personen die keken naar dit probleem keken vrijwel alleen naar planaire grafen, en dan in het bijzonder naar het zogenaamde “vier kleuren probleem”. Dit houdt de vraag in of, gegeven een tweedimensionaal vlak en een verdeling daarvan in stukken (bijvoorbeeld landen met grenzen zonder enclaves), je niet meer dan 4 kleuren nodig hebt om alle stukken op zo‘n manier te kleuren dat geen 2 aangrenzende stukken dezelfde kleur hebben.

De vraag hoeveel kleuren minimaal nodig zijn om een graaf te kleuren op zo’n manier dat geen twee aan elkaar verbonden knopen een zelfde kleur hebben werd voor het eerst genoemd door August Ferdinand Möbius in 1840, in zijn college. Het eerste vermoeden over een antwoord op deze vraag werd echter pas voorgesteld in 1852, toen Francis Guthrie probeerde de kaart van provincies van Engeland te kleuren, en zag dat hij hiervoor maar 4 kleuren nodig had [10]. Hij besprak dit met zijn broer Frederick, die op dat moment student was bij Augustus De Morgan, en zijn broer besprak dit op zijn beurt met De Morgan. Vervolgens publiceerden zowel één van de Guthrie's (het werd geplubiceerd onder de initialen F.G., en dit kan op elk van de broers slaan)(1854) [9] als De Morgan (1860) [3] dezelfde vraag die Möbius stelde, inclusief vermoeden. Zo'n 20 jaar later, in 1879, gaf Alfred Kempe een bewijs van het vier kleuren probleem, en een jaar later, in 1880, werd een ander bewijs gegeven door Peter Guthrie Tait. Echter, 11 jaar later, in 1890, werd het bewijs van Kempe ontkracht door Percy Heawood, en in 1891 werd Tait's bewijs ontkracht door Julius Petersen [12]. Tot de dag van vandaag is men er nog niet in geslaagd zonder de hulp van computers een bewijs of ontkrachting van dit probleem te geven, maar in 1976 werd door Kenneth Appel en Wolfgang Haken bij de Universiteit van Illinois bekend gemaakt dat ze een bewijs hadden gevonden van de voorgestelde oplossing door Guthrie, met behulp van de computer [5]. Niet iedereen accepteert dit bewijs omdat het niet met de hand verifiëerbaar is, maar de meerderheid van de mensen ziet het gegeven bewijs als voldoende, en dus wordt over het algemeen het vier kleuren probleem als opgelost beschouwd.

Ik zal nu de opbouw van mijn scriptie wat toelichten. In het eerste hoofdstuk zullen we een aantal noodzakelijke definities bekijken samen met hun toelichting, zodat deze overzichtelijk in een hoofdstuk bij elkaar staan (en niet kris kras door de hoofdstukken verspreid). Vervolgens zullen we in het tweede hoofdstuk een aantal Lemma's bekijken ter voorbereiding op de hoofdstelling, en zullen we deze Lemma's ook bewijzen. Hierna komen we toe aan het bewijs van de Hoofdstelling van deze scriptie. Tenslotte zullen we het resultaat van de hoofdstelling nader bekijken en proberen toe te lichten wat voor gevolgen dit heeft.

2 Definities

In deze sectie zullen we een aantal belangrijke begrippen met betrekking tot willekeurige grafen en hun kleuringsgetal, die waarschijnlijk nog niet iedere lezer kent, definiëren. Als U zich helemaal thuisvoelt binnen de theorie van willekeurige grafen, kunt U deze sectie overslaan.

Allereerst willen we toewerken naar een wat formelere definitie van het begrip "willekeurige graaf" dan we tot nu toe gebruikt hebben. Om dit te bereiken definiëren we eerst een aantal andere begrippen, om te beginnen het begrip "graaf":

Definitie 2.1. Een **graaf** is een representatie van een verzameling objecten waarin sommige paren objecten verbonden zijn door een lijnstuk. Deze lijnstukken heten **kanten**, en de objecten heten **vertices**.

In figuur 1 staat een voorbeeld van een graaf.

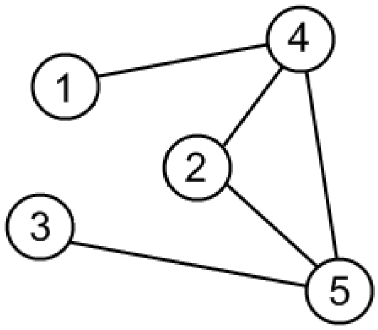

Figure 1: Een graaf.

Om tot een formele definitie van het begrip "willekeurige graaf" te komen, definiëren we nu ook het begrip "kansruimte":

Definitie 2.2. ([11]) Een **kansruimte** is een 3-tupel $(\Omega, \mathcal{F}, \mathbb{P})$ met Ω de **uitkomstenruimte**, die alle mogelijke uitkomsten bevat, $\mathcal{F}$ de **gebeurtenisruimte**, een verzameling van elementen die voorkomen in Ω die alleen de uitkomsten van het huidige experiment bevat, en $\mathbb{P}$ een functie $\mathbb{P} : \mathcal{F} \to \mathbb{R}$ die kansen toekent aan de uitkomsten in $\mathcal{F}$.

Gewapend met deze definities kunnnen we nu een formele definitie geven van het begrip "willekeurige graaf":

Definitie 2.3. ([2]) Laat n een positief geheel getal zijn, en $0 < p < 1$. Dan is de **willekeurige graaf**, $G(n,p)$ een kansruimte over de verzameling van grafen op de verzameling vertices $\{1,...,n\}$, bepaald door onafhankelijke kanten die elk met kans p optreden (dus $P[\{i,j\} \in G] = p$, met de vertices $i,j \in \{1,...,n\}$).

In figuur 2 staat een voorbeeld van een willekeurige graaf, met de kans op iedere mogelijke configuratie van kanten aangegeven, afgezien van labelling. Met labelling heb je 3 verschillende grafen met 1 kant en 3 verschillende grafen met 2 kanten, die elk met een kans van $\frac{4}{125}$ respectievelijk $\frac{16}{125}$ optreden.

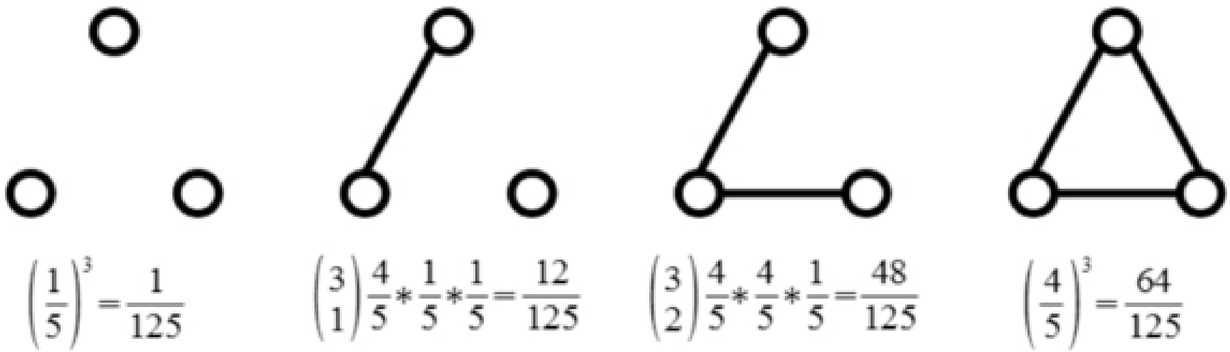

Figure 2: Een willekeurige graaf met $n = 3$ en $p = 4/5$.

We zullen nu ook nog een aantal eigenschappen van grafen definiëren die we later nodig zullen hebben:

Definitie 2.4. Een **kliek** van een graaf G is een complete deelgraaf van G, dit betekent dat het een deelgraaf van G is waarin elke vertex direct verbonden is met elke andere vertex binnen de deelgraaf. Een kliek van maximale mogelijke grootte van een graaf heet een **maximale kliek** van die graaf. Het **kliek nummer** van een graaf is het aantal vertices in een maximale kliek van de graaf. Dit nummer wordt over het algemeen aangeduid met $\omega(G)$.

In figuur 3 staat een voorbeeld van een graaf en een maximale kliek, deze kliek is maximaal omdat er geen enkele andere vertex is die verbonden is met alledrie de vertices 1,2 en 5, en omdat er verder ook geen groep is van meer dan drie vertices die onderling met elkaar verbonden zijn op iedere mogelijke wijze.

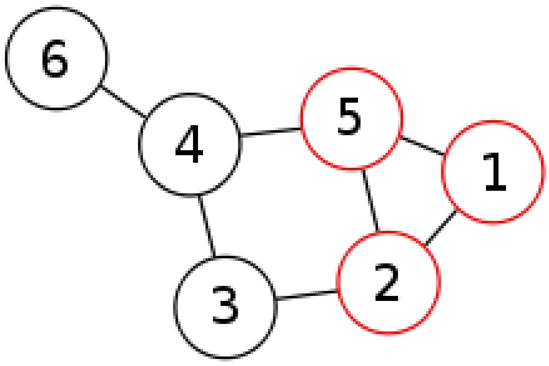

Figure 3: Een graaf met daarin een maximale kliek aangegeven door de rood omcirkelde bolletjes.

Een andere belangrijke eigenschap van een graaf is de volgende:

Definitie 2.5. Het **onafhankelijkheidsgetal** van een graaf is de kardinaliteit van de grootste verzameling onafhankelijke vertices van de graaf, dat wil zeggen een verzameling van vertices van de graaf waartussen geen kant bestaat (dus tussen geen 2 vertices in de verzameling bestaat een kant). Dit wordt over het algemeen aangeduid met $\alpha(G)$, met G de graaf.

In figuur 3 zien we bijvoorbeeld dat $\{1,3,6\}$ een verzameling onafhankelijke vertices is. Aangezien binnen deze graaf elke vertex aan minstens één andere vertex verbonden is door een kant, en je vanuit elke vertex via de kanten naar elke andere vertex kan komen, zien we dat een grotere verzameling onafhankelijke vertices niet kan bestaan, dus is het onafhankelijkheidsgetal van de graaf 3.

Als laatste eigenschap definiëren we het kleuringsgetal van een graaf:

Definitie 2.6. Het **kleuringsgetal** van een graaf is het kleinste aantal kleuren dat nodig is om de vertices van de graaf zó te kleuren, dat geen twee vertices die direct verbonden zijn door een kant dezelfde kleur hebben. Dit getal wordt over het algemeen aangeduid met $\chi(G)$, waarbij G de graaf is.

In figuur 4 zien we een minimaal gekleurde graaf: geen twee aangrenzende vertices hebben dezelfde kleur, en wanneer we proberen een kleur in zijn geheel weg te halen en te veranderen naar een al bestaande kleur, zien we dat niet meer voldaan wordt aan de eis dat elke twee aangrenzende vertices

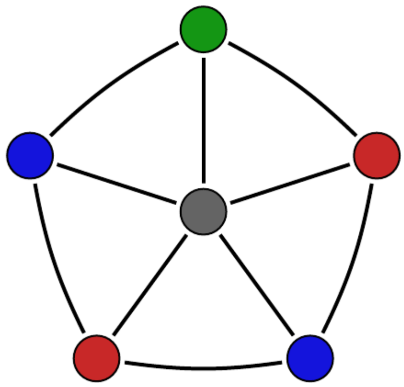

Figure 4: Een minimaal gekleurde graaf, geen twee aangrenzende vertices hebben dezelfde kleur.

verschillende kleuren hebben. Hieruit volgt dat het kleuringsgetal van deze graaf gelijk is aan 4.

Een begrip wat te maken heeft met het kleuringsgetal is de k-kleurbaarheid van een graaf:

Definitie 2.7. Een graaf is ***k*-kleurbaar** als aan die graaf een kleuring van k verschillende kleuren gegeven kan worden die de vertices van de graaf zó kleurt, dat geen twee vertices die direct verbonden zijn door een kant dezelfde kleur hebben.

We zien in dat een graaf met kleuringsgetal m, k-kleurbaar is voor elke $k \geq m$, aangezien het kleuringsgetal van een graaf aangeeft wat het kleinste aantal kleuren is dat nodig is om een graaf juist te kleuren.

We willen nu nog een belangrijk begrip uit de kanstheorie introduceren, de martingaal. Om dit echter te kunnen definiëren moeten we eerst het begrip σ-algebra bekijken:

Definitie 2.8. Een **σ-algebra** op een verzameling X is een collectie C van deelverzamelingen van X die op zijn minst de lege deelverzameling bevat, gesloten is onder complement en gesloten is onder vereniging of doorsnede van een aftelbaar oneindig aantal deelverzamelingen.

Verder introduceren we “conditionele verwachting”:

Definitie 2.9. De **conditionele verwachting** van een willekeurige variabele X gegeven de “condities” F, $E[X|F]$, is een andere willekeurige variabele gelijk aan het gemiddelde van X bij elke mogelijke "conditie" uit F. Wanneer X gedefiniëerd is op een discrete kansruimte, zijn de "condities" een deelverzameling van deze kansruimte.

Voor een definitie van "integreerbaarheid" en "meetbaarheid" verwijs ik naar het boek "Calculus" [1], als ook gebruikt bij Infinitesimaalrekening A en B.

Nu we een definitie van deze begrippen kennen kunnen we een definitie geven van het begrip martingaal:

Definitie 2.10. ([13]) Laat $(\Omega, \mathcal{F}, \mathbb{P})$ een kansruimte zijn. Een **martingaal** van lengte n is een opeenvolging van n willekeurige variabelen $X_1, ...X_n$ met bijbehorende σ-algebra's $\mathcal{F}_1, ..., \mathcal{F}_n$ die voldoen aan de volgende relaties:

1. Elke X_i is een integreerbare willekeurige variabele die meetbaar is ten opzichte van de bijbehorende σ-algebra $\mathcal{F}_i$.
2. Voor de σ-algebra's $\mathcal{F}_i$ geldt dat $\mathcal{F}_i \subset \mathcal{F}_{i+1}$.
3. Voor elke $i \in [2, 3, ..., n-1]$ geldt de relatie $X_i = E[X_{i+1}|\mathcal{F}_i]$ is bijna overal $\mathbb{P}$.

Een voorbeeld van een martingaal is het kapitaal van een gokker, wanneer hij eerlijke gokspellen speelt (dus spellen waarbij de verwachte winst gelijk aan 0 is). We definiëren ook nog een paar speciale martingalen:

Definitie 2.11. ([2]) Laat H een willekeurige graaf zijn van $G(n,p)$, en $f(H)$ een graaf theoretische functie. Label vervolgens de $m = \binom{n}{2}$ mogelijke kanten arbitrair met de rij $1, ..., m$. Definiëer voor $1 \leq j \leq m$ de willekeurige variabele indicator als $I_j = \begin{cases} 1 & \text{als } e_j \in H \\ 0 & anders \end{cases}$, met e_j de kant met label j. Een **kant-blootleggings martingaal** is de rij van willekeurige variabelen $X_0, ..., X_m$ zodat $X_k = E[f(G)|I_1, ..., I_k]$ (overigens is $I_1, ..., I_k$ geen σ-algebra) . Dit geeft dat $X_0 = E[f(G)]$ (want dan heb je geen informatie over welke kanten aanwezig zijn) en $X_m = f(G)$ (want dan heb je alle informatie over welke kanten aanwezig zijn).

Definitie 2.12. ([2]) Laat H een willekeurige graaf zijn van $G(n,p)$, en $f(H)$ een graaf theoretische functie. Label vervolgens de $m = \binom{n}{2}$ mogelijke kanten arbitrair met de rij $1, ..., m$. Definiëer de verzameling $E_i, 1 \leq i \leq n$ als de verzameling van alle mogelijke kanten met vertices in $1, ..., i$. Definiëer ook $\forall j \in E_i$ de willekeurige variabele indicator als $I_j = \begin{cases} 1 & \text{als } e_j \in H \\ 0 & anders \end{cases}$, met

e_j de kant met label j. Definiëer verder de vector $K_i = [I_1, ..., I_j, ...], \forall j \in E_i$ (dit geeft aan welke kanten aanwezig zijn van de mogelijke kanten die in E_i aanwezig zijn). Een **vertex-blootleggings martingaal** is de rij van willekeurige variabelen $Y_0, ..., Y_n$ zodat $Y_k = E[f(G)|K_1, ..., K_i]$ (overigens is $K_1, ..., K_i$ geen σ-algebra). Dit geeft dat $Y_0 = E[f(G)]$ (want dan heb je geen informatie over welke kanten aanwezig zijn) en $Y_n = f(G)$ (want dan heb je alle informatie over welke kanten aanwezig zijn).

Een paar eigenschappen van functies die te maken hebben met kant- en vertex-blootleggings martingalen zijn de Lipschitz condities:

Definitie 2.13. Een graaf theoretische functie f voldoet aan de **Lipschitz kant conditie** wanneer geldt dat als een graaf H en H' maar in één kant verschillen, dan $|f(H) - f(H')| \leq 1$. Diezelfde functie voldoet aan de **Lipschitz vertex conditie** wanneer geldt dat als een graaf H en H' maar in één vertex verschillen, dan $|f(H) - f(H')| \leq 1$.

Een voorbeeld van een toepassing van de Lipschitz condities is het volgende lemma:

Lemma 2.14. Als f voldoet aan de Lipschitz kant conditie, dan voldoet de bijbehorende kant-blootleggings martingaal X (behorend bij de functie f en opgesteld volgens definitie 2.11) aan $|X_{i+1} - X_i| \leq 1$. Als f voldoet aan de Lipschitz vertex conditie, dan voldoet de bijbehorende vertex-blootleggings martingaal Y (behorend bij de functie f en opgesteld volgens definitie 2.12) aan $|Y_{i+1} - Y_i| \leq 1$. Dit eerste volgt uit het feit dat $X_i = E[f(G)|I_1, ..., I_i]$, dus $|X_{i+1} - X_i| = |E[f(G)|I_1, ..., I_{i+1}] - E[f(G)|I_1, ..., I_i]|$, omdat je bij de eerste term informatie hebt over 1 kant meer dan bij de tweede term, verschillen de verwachte grafen behorend bij deze 2 termen maximaal in één kant, en dus volgt uit de Lipschitz kant conditie dat $|X_{i+1} - X_i| \leq 1$. Gebruik makend van $Y_k = E[f(G)|K_1, ..., K_i]$ volgt op dezelfde manier dat $|Y_{i+1} - Y_i| \leq 1$.

Tenslotte zullen we nog een paar definities geven van termen die we later zullen gebruiken: allereerst "bijna altijd".

Definitie 2.15. Iets gebeurt **bijna altijd** wanneer datgene met voldoende grote kans gebeurt. Oftewel, er is een $\alpha \approx 1$ zodat de kans dat hetgeen waar we naar kijken optreedt gelijk is aan α. Hoe groot deze α precies is hangt af van de situatie.

Vervolgens definiëren we nog wat de constante orde van een reeks, $o(1)$, inhoudt.

Definitie 2.16. Als een reeks $f(n) = o(1)$, dan betekent dit dat $\frac{f(n)}{c} \to 0$ voor elke constante c. Oftewel, $o(x)$ geeft een strikte bovengrens x aan voor hetgeen waar de reeks naartoe convergeert.

3 Lemma's, en hun bewijzen

In deze sectie zullen we een aantal lemma's introduceren en bewijzen die we nodig gaan hebben bij het bewijzen van ons hoofdresultaat, en een aantal die interessant zijn en passen bij dit onderwerp, aan de hand van hoofdstuk 7 en hoofdstuk 10 uit [2]. Dit doen we apart zodat we bij het bewijzen van het hoofdresultaat niet telkens van context moeten wisselen, en we dus de focus geheel op het bewijs van het hoofdresultaat kunnen richten in de desbetreffende sectie.

Vanaf hier stellen we dat $G \sim G(n, 1/2)$, met $n \geq 2$. Met het teken $\sim$ bedoelen we bij gebruik bij variabelen dat deze dezelfde orde van grootte hebben, en bij gebruik bij willekeurige grafen dat deze een soortgelijke distributie (n, p) hebben. Ook stellen we dat $f(k) = \binom{n}{k} 2^{-\binom{k}{2}}$, met k_0 de waarde zodat $f(k_0 - 1) > 1 > f(k_0)$. We weten zeker dat zo'n k_0 bestaat omdat met $n \geq 2$ zeker geldt dat $f(1) = n * 1 > 1$, en omdat $f(n) = 1 * 2^{-\binom{n}{2}} < 1$. Stel namelijk dat er niet zo'n k_0 bestaat, dan moet $\forall i \in \{1, .., n\}$ gelden dat óf zowel $f(i)$ als $f(i-1)$ groter is dan 1, óf zowel $f(i)$ als $f(i-1)$ kleiner is dan 1. $f(i)$ kan niet gelijk zijn aan 1 omdat $\binom{n}{k}$ alleen een positieve macht van 2 is (2^1, 2^2, etc.) wanneer $k = 1$ of $k = n - 1$, met n een positieve macht van 2. Als $k = 1$ dan geldt $2^{-\binom{k}{2}} = 1$, en $\binom{n}{k} = n > 1$. Als $k = n - 1$ dan geldt $2^{-\binom{k}{2}} = 2^{-(n-1)(n-2)/2}$ en $\binom{n}{k} = n - 1 < 2^{(n-1)(n-2)/2}$ $\forall n \geq 4$ (voor $n = 2$ geldt $n - 1 = 1$). Dus $f(i)$ is nooit gelijk aan 1. Omdat $f(1) > 1$, kan er alleen geen k_0 bestaan als $\forall i \in \{1, .., n\}$ geldt dat $f(i) > 1$. Echter, $f(n) < 1$, tegenspraak, dus bestaat er zo'n k_0. Uit $f(k_0 - 1) > 1 > f(k_0)$ volgt dat $n = \sqrt{2}^{k(1+o(1))}$, dus voor $k \sim k_0$ geldt dat $f(k+1)/f(k) = \frac{n}{k} 2^{-k}(1 + o(1)) = n^{-1+o(1)}$. Nu stellen we $k = k(n) = k_0(n) - 4$, zodat $k \sim 2\log_2(n)$ (dit volgt uit $n = \sqrt{2}^{k(1+o(1))}$) en $f(k) > n^{3+o(1)}$. Laat verder $Y = Y(H)$ de maximale grootte zijn van een familie van kant-disjuncte klieken, van grootte k in H, met H een willekeurige graaf van G. Dan:

Lemma 3.1. $E[Y] \geq \frac{n^2}{2k^4}(1 + o(1))$, oftewel de verwachtingswaarde van Y is groter of gelijk aan het aantal vertices in G in het kwadraat, gedeeld door twee maal de grootte van de klieken tot de vierde macht vermenigvuldigd met een term die naar 1 convergeert.

Bewijs: Laat K de familie van k-klieken (klieken van grootte k) van G zijn zodat $f(k) = \mu = E[|K|]$. Laat W het aantal ongeordende paren $\{A_i, A_j\}$ zijn van k-klieken van G met $2 \leq |A_i \cap A_j| < k$. Dan geldt dat $E[W] = \Delta/2$, met Δ gelijk aan de som over geordende paren k-klieken, oftewel $\Delta = \sum\limits_{i\sim j} P[A_i \wedge A_j]$, met $i \sim j$ wanneer A_i en A_j niet onafhankelijk zijn en $i \neq j$, dit is dus de som over alle i en j waarvoor dit geldt. We schrijven dit nu om: $\Delta = \sum\limits_{i\sim j} P[A_i \wedge A_j] = \sum\limits_{i} P[A_i] \sum\limits_{j\sim i} P[A_j|A_i]$. We stellen nu $\Delta^* = \sum\limits_{j\sim i} P[A_j|A_i]$, en we zien dat $\sum\limits_{i} P[A_i] = E[|K|] = \mu$. Dus $\Delta = \Delta^*\mu$. Omdat we weten dat $A_i \sim A_j$ dan en slechts dan als $2 \leq |A_i \cap A_j| < k$ zien we dat we Δ^* ook als volgt kunnen schrijven: $\Delta^* = \sum\limits_{i=2}^{k-1} \binom{k}{i}\binom{n-k}{k-i} 2^{\binom{i}{2}-\binom{k}{2}}$, en dus geldt dat $\frac{\Delta^*}{\mu} = \sum\limits_{i=2}^{k-1} g(i)$, met $g(i) = \frac{\binom{k}{i}\binom{n-k}{k-i}2^{\binom{i}{2}}}{\binom{n}{k}}$. We kunnen $g(i)$ zien als de kans dat een willekeurig gekozen A_j een vastgelegde A_m in precies i punten snijdt. Als we nu $i = 2$ laten, vinden we dat $g(2) = 2\frac{\binom{k}{2}\binom{n-k}{k-2}}{\binom{n}{k}} = 2\frac{\frac{(n-k)!k!}{2(k-2)!(k-2)!(n-2k+2)!}}{\frac{n!}{k!(n-k!)}} = \frac{(k!)^2((n-k)!)^2}{((k-2)!)^2(n-2k+2)!n!} = \frac{(k(k-1))^2((n-k)...(n-2k+3))}{(n...(n-k+1))} \sim \frac{k^4}{n^2}$, en uit $\frac{\Delta^*}{\mu} = \sum\limits_{i=2}^{k-1} g(i)$ volgt nu dat $\Delta = \sum\limits_{i=2}^{k-1} g(i) * \mu^2$, en voor $i = 2$ geeft dit dat $\Delta \sim \frac{\mu^2 k^4}{n^2}$. Dit is ook onze schatting van Δ, aangezien $g(2)$ de dominerende term is van $\sum\limits_{i=2}^{k-1} g(i)$ (dus de orde van $g(2)$ is groter dan de orde van de som van alle andere termen in $\sum\limits_{i=2}^{k-1} g(i)$). Om dit te laten zien bekijken we eerst de term waar $i = k-1$, dit geeft $g(k-1) = \frac{k(n-k)2^{-(k-1)}}{\binom{n}{k}2^{-\binom{k}{2}}} \sim \frac{2kn2^{-k}}{\mu}$. Wanneer $k \sim 2\log_2(n)$ dan is de teller $n^{-1+o(1)}$. De noemer is dan groter dan $n^{3+o(1)}$, dus geldt zeker dat $g(k-1) \leq o(n^{-1})$. Op eenzelfde wijze vinden we voor $i \in \{3, ..., k-2\}$ dat $g(i) \leq o(n^{-1})$. Hieruit volgt dat $\sum\limits_{i=3}^{k-1} g(i) \leq o(n^{-1})$. Als we nu kijken naar $g(2)$, dan zien we dat $\frac{k^4}{n^2} > o(n^{-1})$, en dus $o(g(2)) > o(\sum\limits_{i=3}^{k-1} g(i))$, en dus is $g(2)$ de dominerende term van $\sum\limits_{i=2}^{k-1} g(i)$. Laat nu C een willekeurige deelfamilie van K zijn, gedefinieerd door $\forall A_i \in K$ te zetten dat $P[A_i \in C] = q$, met q nader te bepalen. Laat W' het aantal ongeordende paren $\{D, E\}$, $D, E \in C$ zijn, met $2 \leq |D \cap E| < k$. Dan geldt: $E[W'] = E[W]q^2 = \frac{\Delta q^2}{2}$. Verwijder nu uit C één verzameling uit elk gevonden paar $\{D, E\}$. Dit levert een verzameling C^- van kant-disjuncte k-klieken van G en we vinden dat $E[Y] \geq E[|C^-|] \geq E[|C|] - E[W'] = \mu q - \frac{\Delta q^2}{2} = \frac{\mu^2}{2\Delta} \sim \frac{n^2}{2k^4}$, waar we $q = \frac{\mu}{\Delta}$ kiezen om het kwadraat te minimaliseren (want dan is q namelijk kleiner dan

1, $\Delta \sim \frac{\mu^2 k^4}{n^2}$ geeft dat $q = \frac{\mu}{\Delta} = \frac{n^2}{\mu k^4} = \frac{n^2}{f(k)k^4} < \frac{n^2}{n^{3+o(1)}k^4} = \frac{1}{n^{1+o(1)}k^4} < 1$). $\square$

We zullen nu, alvorens een lemma te kijken waarin wij een andere ongelijkheid aantonen, eerst Azuma's ongelijkheid bewijzen, omdat we deze nodig zullen hebben voor het bewijs van lemma 3.3.

Lemma 3.2. (Azuma's ongelijkheid) Laat $0 = X_0, X_1, ..., X_m$ een martingaal zijn met $|X_{i+1} - X_i| \leq 1$ voor alle $0 \leq i < m$. Laat $\lambda > 0$ willekeurig zijn. Dan geldt dat $P[X_m > \sqrt{m}\lambda] < e^{-\lambda^2/2}$

Bewijs: Laat, vooruitkijkend, $\alpha = \lambda/\sqrt{m}$. Laat $Y_i = X_i - X_{i-1}$, zodat $|Y_i| \leq 1$ en $E[Y_i|X_{i-1}, X_{i-2}, X_{i-3}, ..., X_0] = 0$. We weten dat $\cosh(x) = (e^{\lambda} + e^{-\lambda})/2$, we willen nu laten zien dat $\cosh(x) \leq e^{x^2/2}$. Dit doen we door de Taylor-reeksen termsgewijs te vergelijken rondom $a = 0$: de Taylor-reeks van $\cosh(x)$ is $1 + \frac{x^2}{2} + \frac{x^4}{24} + \frac{x^6}{720} + O(x^7)$, en de Taylor-reeks van $e^{x^2/2}$ is $1 + \frac{x^2}{2} + \frac{x^4}{8} + \frac{x^6}{48} + O(X^7)$. We zien duidelijk dat vanaf de derde term de termen van de Taylor-reeks van $e^{x^2/2}$ groter zijn dan de termen van de Taylor-reeks van $\cosh(x)$, en dat de eerste twee termen gelijk zijn. Hieruit concluderen we dat inderdaad $\cosh(x) \leq e^{x^2/2}$. Laat nu $l(x) = \frac{e^{\alpha}+e^{-\alpha}}{2} + \frac{e^{\alpha}-e^{\alpha}}{2}x$. Voor $x \in [-1, 1]$ geldt dat $e^{\alpha x} \leq l(x)$, want $l(x)$ is de koorde door de punten $x = 1$ en $x = -1$ van de convexe curve $e^{\alpha x}$, en hieruit volgt dat $E[e^{\alpha Y_i}|X_{i-1}, X_{i-2}, X_{i-3}, ..., X_0] \leq E[l(Y_i)|X_{i-1}, X_{i-2}, X_{i-3}, ..., X_0] = l(E[Y_i|X_{i-1}, X_{i-2}, X_{i-3}, ..., X_0]) = l(0) = \cosh(\alpha) \leq e^{\alpha^2/2}$.

Dit geeft: $E[e^{\alpha X_m}] = E[\prod_{i=1}^{m} e^{\alpha Y_i}] = E[(\prod_{i=1}^{m-1} e^{\alpha Y_i})E[e^{\alpha Y_m}|X_{m-1}, X_{m-2}, X_{m-3}, ..., X_0] \leq$

$E[\prod_{i=1}^{m-1} e^{\alpha Y_i}]e^{\alpha^2/2} \leq e^{\alpha^2 m/2}$ (Dit vinden we door m keer een term van het product af te splitsen en de zojuist gebruikte ongelijkheid m keer toe te passen). Dit leidt tot de volgende conclusie: $P[X_m > \lambda\sqrt{m}] = P[e^{\alpha X_m} > e^{\alpha\lambda\sqrt{m}}] < E[e^{\alpha X_m}]e^{-\alpha\lambda\sqrt{m}} \leq e^{\alpha^2 m/2 - \alpha\lambda\sqrt{m}} = e^{-\lambda^2/2}$, en dit is precies wat we wilden bewijzen. $\square$

We zullen nu van dit resultaat gebruik maken samen met lemma 3.1 om aan te tonen dat $P[\omega(G) < k] < e^{-(c+o(1))\frac{n^2}{\ln^8(n)}}$, waar $\omega(G)$ de grootte is van een maximale kliek van G, en c een positieve constante is.

Lemma 3.3. $P[\omega(G) < k] < e^{-(c+o(1))\frac{n^2}{\ln^8(n)}}$, met c een positieve constante.

Bewijs: Laat $Y_0, Y_1, ..., Y_m$, $m = \binom{n}{2}$, de kant-blootleggings martingaal op $G(n, 1/2)$ zijn, met de bovenaan deze sectie gedefiniëerde functie Y. De functie Y voldoet aan de Lipschitz kant conditie, omdat het toevoegen van een enkele kant maximaal één extra kliek toe kan voegen aan een familie van kant-disjuncte klieken (de Lipschitz conditie zou niet gelden voor het aantal k-klieken, omdat een enkele kant vele nieuwe klieken toe zou kunnen voegen in dat geval). G heeft geen k-kliek dan en slechts dan als $Y = 0$. We passen nu Azuma's ongelijkheid (zie lemma 3.2) toe met $m = \binom{n}{2} \sim n^2/2$ en $E[Y] \geq \frac{n^2}{2k^4}(1 + o(1))$. Dit geeft dat $\lambda = \frac{E[Y]}{\sqrt{\binom{n}{2}}}$. Dan: $P[\omega(G) < k] = P[Y = 0] \leq P[Y - E[Y] \leq -E[Y]] = P[E[Y] - Y \geq E[Y]]$ **(Azuma)** $\leq e^{-E[Y]^2/2\binom{n}{2}} \leq e^{-(c'+o(1))n^2/k^8} = e^{-(c+o(1))\frac{n^2}{\ln^8(n)}}$, en dit is wat we zochten. □

We zullen nu kijken naar een stelling, die ons leert dat het aantal verschillende waarden voor het kleuringsgetal bij een willekeurige graaf $G(n, p)$ voor waarden p van een nader te specifiëren vorm, bijna altijd slechts vier is (hier geldt dus niet dat per se $G \sim G(n, 1/2)$).

Stelling 3.4. *Laat $p = n^{-\alpha}$, met $\alpha > \frac{5}{6}$, en α vastliggend. Laat $G = G(n, p)$. Dan zijn er $u = u(n, p)$, zodat bijna altijd geldt dat $u \leq \chi(G) \leq u + 3$. Oftewel, de mogelijke waarden van $\chi(G)$ zijn bijna altijd verdeeld over 4 waarden.*

Voordat we deze stelling zullen bewijzen, hebben we eerst een technisch lemma nodig:

Lemma 3.5. Laat α, c vastliggen, met $\alpha > \frac{5}{6}$. Laat $p = n^{-\alpha}$. Dan kunnen bijna altijd elke $c\sqrt{n}$ vertices van $G = G(n, p)$ 3-kleurbaar zijn.

Bewijs: Stel dat dit niet het geval is. Laat T een minimale verzameling vertices zijn die niet 3-kleurbaar is (zie definities voor een definitie van k-kleurbaarheid). Hieruit volgt dus dat $T - \{z\}$ 3-kleurbaar is, en dat z minstens graad 3 moet hebben binnen T (oftewel z zit aan minstens 3 andere vertices direct vast met een kant), dit geldt voor alle $z \in T$, met z een vertex. Hieruit volgt dat als T t vertices heeft, dan moet T minstens $\frac{3t}{2}$ kanten hebben. De kans dat dit voorkomt voor een T met maximaal $c\sqrt{n}$ vertices wordt van bovenaf begrensd door $\sum_{t=4}^{c\sqrt{n}} \binom{n}{t}\binom{\binom{t}{2}}{\lceil\frac{3t}{2}\rceil}p^{\frac{3t}{2}}$. We begrenzen vervolgens $\binom{n}{t} \leq (\frac{ne}{t})^t$, want $\binom{n}{t} = \frac{n!}{t!(n-t)!} = \frac{(n...(n-t+1))}{t!} = (1 - \frac{1}{n})...(1 - \frac{t-1}{n})\frac{n^t}{t!}$, omdat alle factoren binnen de haakjes kleiner dan 1 zijn, geldt dat $\binom{n}{t} \leq \frac{n^t}{t!}$, en omdat alle termen van de reeks $e^t = \sum_{n=0}^{\infty} \frac{t^n}{n!}$ positief zijn als t een positief reëel getal is, moet elke individuele term kleiner zijn dan de hele som, in het bijzonder impliceert dit dat, voor elk niet negatief geheel getal t, geldt: $e^t > \frac{t^t}{t!}$. Als we dit omschrijven krijgen we $1 < \frac{t!e^t}{t^t}$. Wanneer we dit vermenigvuldigen met de eerder gevonden ongelijkheid voor $\binom{n}{t}$, dan vinden we: $\binom{n}{t} \leq \frac{n^t}{t!} * \frac{t!e^t}{t^t} = \frac{n^te^t}{t^t} = (\frac{ne}{t})^t$, zoals gewenst. Verder begrenzen we $\binom{\binom{t}{2}}{\lceil\frac{3t}{2}\rceil} \leq (\frac{te}{3})^{\lceil 3t/2\rceil}$, dit volgt uit de zojuist bewezen ongelijkheid: als we daar $n = \binom{t}{2}$ en $t = \lceil\frac{3t}{2}\rceil$ invullen, dan geeft dit dat $\binom{n}{t} \leq (\frac{ne}{t})^t = (\frac{\binom{t}{2}e}{\lceil\frac{3t}{2}\rceil})^{\lceil\frac{3t}{2}\rceil} \leq (\frac{t(t-1)e}{3t})^{\lceil\frac{3t}{2}\rceil} \leq (\frac{te}{3})^{\lceil\frac{3t}{2}\rceil}$. Uit deze begrenzingen volgt dat elke term maximaal gelijk is aan $[\frac{ne}{t}\frac{t^{3/2}e^{3/2}}{3^{3/2}}n^{-3\alpha/2}]^t \leq [c_1n^{1-\frac{3\alpha}{2}}t^{1/2}]^t \leq [c_2n^{1-\frac{3\alpha}{2}}n^{1/4}]^t = [c_2n^{-\epsilon}]^t$ met $\epsilon = \frac{3\alpha}{2} - \frac{5}{4} > 0$ en c_1, c_2 constanten (we verwaarlozen de ceiling uit de tweede begrenzing hier omdat deze afronding alleen voor de constante term uitmaakt). Hierdoor is de som $o(1)$, en dus zal bijna altijd een verzameling van $c\sqrt{n}$ vertices van $G = G(n,p)$ 3-kleurbaar zijn. $\square$

Nu kunnen we stelling 3.4 bewijzen:

Bewijs [stelling 3.4]: Laat $\epsilon > 0$, met ϵ arbitrair klein, en laat $u = u(n,p,\epsilon)$ het kleinste gehele getal zijn zodat $P[\chi(G) \leq u] > \epsilon$. Definiëer nu $Y(G)$ als de minimale grootte van een verzameling vertices S, waarvoor $G - S$ u-kleurbaar kan zijn. Deze Y voldoet aan de Lipschitz vertex conditie, omdat je in het slechtste geval één vertex toe kan voegen aan S. Pas de vertex-blootleggings martingaal op $G(n,p)$ toe op Y. Laat $\mu = E[Y]$. Uit Azuma's ongelijkheid (lemma 3.2) volgt nu: $P[Y \leq \mu - \lambda\sqrt{n-1}] < e^{-\lambda^2/2}$ en $P[Y \geq \mu + \lambda\sqrt{n-1}] < e^{-\lambda^2/2}$. Dit komt omdat een direct gevolg van lemma 3.2 is dat als $c = Y_0, ..., Y_m$ een martingaal is met $|Y_{i+1} - Y_i| \leq 1$ $\forall$ $0 \leq i < m$, oftewel als c voldoet aan de Lipschitz vertex conditie, dan geldt dat $P[|Y_m - c| > \lambda\sqrt{m}] < 2e^{-\lambda/2}$. Omdat E[Y] een martingaal is die voldoet aan de Lipschitz vertex conditie, volgen de 2 gegeven ongelijkheden uit het zojuist gegeven gevolg van lemma 3.2 (je laat de absolute waarde weg en haalt c naar de andere kant, en je krijgt dus een begrenzing aan 2 kanten). Laat nu λ voldoen aan $e^{-\lambda^2/2} = \epsilon$ zodat deze randgevallen elk optreden met kans $< \epsilon$. We hebben u zó gedefiniëerd, dat met kans minstens ϵ, G u-kleurbaar zou zijn en dus $Y = 0$. Dit geeft dat $P[Y = 0] > \epsilon$. De eerste ongelijkheid levert hierom dat $\mu \leq \lambda\sqrt{n-1}$. Als we nu ook de tweede ongelijkheid gebruiken vinden we dat $P[Y \geq 2\lambda\sqrt{n-1}] \leq P[Y \geq \mu + \lambda\sqrt{n-1}] < \epsilon$. Met kans minstens $1 - \epsilon$ is er een u-kleuring van alle vertices op hoogstens $c'\sqrt{n}$ vertices na. Volgens lemma 3.5 kunnen deze punten bijna altijd, en dus met kans minstens $1-\epsilon$, met 3 meer kleuren kleurbaar zijn, wat een $u+3$-kleuring van G oplevert. De minimaliteit van u garandeert dat met kans minstens $1 - \epsilon$, er minstens u kleuren nodig zijn voor het kleuren van G. Samen levert dit

alles dat $P[u \leq \chi(G) \leq u+3] \geq 1-3\epsilon$, en ϵ was arbitrair klein. $\square$
Nu we deze lemma's behandeld hebben, kunnen we toe naar de hoofdstelling: in de volgende sectie zullen we deze formuleren en bewijzen, met behulp van een aantal van de zojuist bewezen lemma's.

4 Bewijs Hoofdstelling

In deze sectie zullen we enkel kijken naar de hoofdstelling: het bespreken van het resultaat zullen we in de volgende sectie doen. We bewijzen de hoofdstelling aan de hand van paragraaf 10.3 uit [2].

Hoofdstelling 4.1. [Bollobás (1988)]: *Bijna altijd geldt:* $\chi(G) \sim \frac{n}{2\log_2(n)}$, *met* $G \sim G(n, 1/2)$.

Bewijs: Laat $\alpha(G) = \omega(\overline{G})$, met $\overline{G}$ het complement van G, het onafhankelijkheidsgetal van G zijn. Het complement van G heeft dezelfde distributie $G(n, 1/2)$. Dus geldt bijna altijd dat $\alpha(G) \leq (2 + o(1))\log_2(n)$. Dit geeft dat bijna altijd geldt dat $\chi(G) \geq \frac{n}{\alpha(G)} \geq \frac{n}{2\log_2(n)}(1 + o(1))$, want het minimale aantal onafhankelijke verzamelingen dat samen elke vertex bevat is $\lceil \frac{n}{\alpha(G)} \rceil$, omdat $\alpha(G)$ de grootte van de grootste onafhankelijke verzameling is, dus geldt $\chi(G) \geq \frac{n}{\alpha(G)}$, want je kunt binnen iedere onafhankelijke verzameling elk element één kleur geven om zo een distincte kleuring te krijgen wanneer je onafhankelijke verzamelingen hebt die samen elke vertex bevatten uit G. We willen nu ook dat de omgekeerde ongelijkheid geldt. Laat $m = \lfloor n/\ln^2 n \rfloor$. Voor elke verzameling S van m vertices heeft de restrictie $G|_S$ de distributie van $G(m, 1/2)$. Laat $k = k(m) = k_0(m) - 4$ zoals bovenaan hoofdstuk 3. Merk op dat $k \sim 2\log_2(m) \sim 2\log_2(n)$, dit eerste hebben we wederom bovenaan hoofdstuk 3 gezien, en dit tweede volgt uit $2\log_2(m) = 2\log_2(\lfloor n/\ln^2 n \rfloor) \sim 2\log_2(n) - 2\log_2(\ln^2 n) \sim 2\log_2(n)$. Dan, wegens lemma 3.3: $P[\alpha[G|_S] < k] < e^{-m^{2+o(1)}}$. Er zijn $\binom{n}{m} < 2^n = 2^{m^{1+o(1)}}$ van zulke verzamelingen S. Dus $P[\alpha[G|_S] < k$ voor een zekere m-verzameling $S] < 2^{m^{1+o(1)}} e^{-m^{2+o(1)}} = o(1)$ (je vermenigvuldigt de bovengrens voor het aantal verzamelingen met de bovengrens van de kans dat het gebeurt voor een willekeurige verzameling). Oftewel, bijna altijd bevat elke m vertices een onafhankelijke verzameling van k elementen. Neem nu aan dat G deze eigenschap bezit. We halen onafhankelijke verzamelingen van k elementen uit G en geven elk een andere kleur tot er minder dan m vertices over zijn. Vervolgens geven we elk punt een andere kleur. Door het volgen van deze procedure vinden we dat $\chi(G) \leq \lceil \frac{n-m}{k} \rceil + m \leq \frac{n}{k} + m = \frac{n}{2\log_2(n)}(1 + o(1)) + o(\frac{n}{\log_2(n)}) = \frac{n}{2\log_2(n)}(1 + o(1))$, en dit gebeurt voor bijna elke G. $\square$

In de volgende sectie zullen we dit resultaat bespreken.

5 Discussie Resultaat

In deze sectie zullen we kijken naar welke eerdere bevindingen er waren over $\chi(G)$ met $G \sim G(n, 1/2)$, en naar de gevolgen van het feit dat deze stelling bewezen is.

In 1988 heeft Béla Bollabás de stelling bewezen die wij in de vorige sectie behandeld hebben, maar dan in iets algemenere vorm [4]. Dit was echter niet de eerste keer dat iemand keek naar dit probleem, eerder in 1975 hadden Grimmet en McDiarmid [8] namelijk al een interval gevonden waarop $\chi(G)$ bijna altijd moet liggen. Grimmet en McDiarmid vonden dat $\chi(G)$ bijna altijd ligt in het volgende interval: $\frac{n}{2\log_2(n)}(1+o(1)) \leq \chi(G) \leq \frac{n}{\log_2(n)}(1+o(1))$, met $G \sim G(n, 1/2)$. Voor een aantal jaar was dit de nauwkeurigste grens die men op $\chi(G)$ kon plaatsen, totdat in 1987 Matula een nauwkeurigere grens wist te vinden waarbinnen $\chi(G)$ bijna altijd ligt: $\frac{n}{2\log_2(n)}(1 + o(1)) \leq \chi(G) \leq \frac{2n}{3\log_2(n)}(1 + o(1))$. Vervolgens wist Bollobas dus in 1988 te bewijzen dat bijna altijd geldt dat $\chi(G) \sim \frac{n}{2\log_2(n)}$, met $G \sim G(n, 1/2)$. Zoals te zien in de eerder gegeven grenzen was al wel bekend dat $\frac{n}{2\log_2(n)}(1 + o(1))$ een ondergrens was van $\chi(G)$, maar dat dit ook een bovengrens was bewees Bollobás als eerste.

Wat betreft de gevolgen van deze stelling: men heeft nu een zeer strakke grens waar $\chi(G)$ bijna altijd binnen ligt wanneer $G \sim G(n, 1/2)$ (en men heeft dit ook kunnen vinden voor algemenere G). Dit wordt sporadisch gebruikt, wanneer men iets binnen het gebied van de willekeurige grafen wil bewijzen, wat te maken heeft met of gebruikmaakt van het kleuringsgetal van een graaf. Verder kan men hiermee voorspellen, gegeven het aantal vertices van een graaf en de kans waarmee een bepaalde kant optreedt, wat het kleuringsgetal van een graaf bijna altijd is en dus hoeveel kleuren je bijna altijd nodig zal hebben om hem te kleuren op zó'n manier dat geen twee vertices die direct verbonden zijn door een kant dezelfde kleur hebben.

Verwijzingen

[1] Robert A Adams. A complete course calculus, 1995.

[2] Noga Alon and Joel H. Spencer. *The probabilistic method*. Wiley-Interscience Series in Discrete Mathematics and Optimization. Wiley-Interscience [John Wiley & Sons], New York, second edition, 2000. With an appendix on the life and work of Paul Erdős.

[3] De Morgan (anonymous). The philosophy of discovery, chapters historical and critical. by w. whewell. *The Athenaeum*, 1680:501–503, 1860.

[4] B. Bollobás. The chromatic number of random graphs. *Combinatorica*, 8(1):49–55, 1988.

[5] G. Chartrand and L. Lesniak. *Graphs & digraphs*. Chapman & Hall/CRC, Boca Raton, FL, fourth edition, 2005.

[6] P. Erdős and A. Rényi. On random graphs. I. *Publ. Math. Debrecen*, 6:290–297, 1959.

[7] E. N. Gilbert. Random graphs. *Ann. Math. Statist.*, 30:1141–1144, 1959.

[8] Geoffrey R Grimmett and Colin JH McDiarmid. On colouring random graphs. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 77, pages 313–324. Cambridge Univ Press, 1975.

[9] F. Guthrie. Tinting maps. *The Athenaeum*, 1389:726, 1854.

[10] Donald MacKenzie. *Mechanizing proof*. Inside Technology. MIT Press, Cambridge, MA, 2001. Computing, risk, and trust.

[11] Daniel W. Stroock. *Probability theory, an analytic view*. Cambridge University Press, Cambridge, 1993.

[12] Robin Thomas. An update on the four-color theorem. *Notices Amer. Math. Soc.*, 45(7):848–859, 1998.

[13] Ivan Matveevič Vinogradov and Michiel Hazewinkel. *Encyclopaedia of mathematics*. Kluwer Academic Publ, 2001.