In de verdediging

| In defence

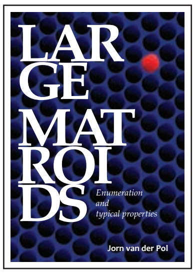

Foto: Elisa van Hout

Large Matroids, Enumeration and Typical Properties
Jorn van der Pol

‘Promovendus blaast nieuw leven in wiskundig vakgebied’, kopte *Cursor*, de nieuwssite van de Technische Universiteit Eindhoven, nadat Jorn van der Pol op 20 september *cum laude* was gepromoveerd. De combinatie van promotor prof. dr. Remco van der Hofstad en co-promotor dr. Rudi Pendavingh had de nieuwsgierigheid al gewekt: een kansrekenaar en een discreet wiskundige. En de titel van zijn proefschrift, *Large Matroids, Enumeration and Typical Properties*, vroeg om uitleg: “Wat zijn matroïden eigenlijk?” De *Cursor*-kop bevestigde de keuze om Van der Pol te interviewen nog eens extra.

Matroïden

Matroïden zijn abstracte meetkundige objecten die bestaan uit punten, lijnen en hoger-dimensionale tegenhangers. Deze objecten zijn sterk verbonden met verschillende gebieden binnen de wiskunde, zoals lineaire algebra en grafentheorie. Veel interessante problemen, zoals het vinden van minimale opspannende bomen, kunnen worden geformuleerd in termen van matroïden, en veel algoritmen om dergelijke problemen op te lossen zijn in feite matroïde-algoritmen.

De term matroïde werd door Hassler Whitney geïntroduceerd in 1935 in zijn artikel ‘On the abstract properties of linear dependence’. Hij kwam als volgt tot zijn definitie: neem een verzameling vectoren E die leeft in een of andere eindig-dimensionale vectorruimte V (je kunt dus zeggen dat E de verzameling kolommen van een of andere matrix is) en veronderstel dat I een onafhankelijke verzameling vectoren in E is. Wat betekent dit? Ten eerste kun je stellen dat als J een deelverzameling van I is, ook J linear onafhankelijk is. Ten tweede geldt een soort uitbreidingsprincipe: als een deelverzameling J van E ook lineair onafhankelijk is en J bevat strikt meer elementen dan I, dan moet J een vector bevatten die niet in het opspansel van I zit.

In zijn artikel bestudeerde Whitney structuren die aan precies deze eigenschappen voldoen: hij definieerde een matroïde als een verzameling E en een collectie $\mathcal{I}$ van deelverzamelingen van E, met de volgende eigenschappen:

1. $\mathcal{I}$ is niet-leeg;
2. Als $I \in \mathcal{I}$ is, en $J \subset I$, dan is ook $J \in \mathcal{I}$;
3. Als $I \in \mathcal{I}$ en $J \in \mathcal{I}$, en $|J| > |I|$, dan is er een element $j \in J$ zodat de verzameling $I \cup \{j\} \in \mathcal{I}$.

Eigenschap 1 sluit een triviaal object uit, terwijl eigenschappen 2 en 3 precies corresponderen met de genoemde eigenschappen van vectorruimten. Whitney noemde elementen in $\mathcal{I}$ ‘onafhankelijk’.

Je kunt laten zien dat de inclusiegewijs-maximale elementen van $\mathcal{I}$ allemaal even groot zijn: dit noem je de *rang* van de matroïde. In het geval van vectoren is dit precies de dimensie van $\langle E \rangle$. Inclusiegewijs-maximale onafhankelijke verzamelingen zelf noem je om voor de hand liggende reden *bases*.

Pas gepromoveerden brengen hun werk onder de aandacht. Heeft u tips voor deze rubriek of bent u zelf pas gepromoveerd? Laat het weten aan onze redacteur.

Redacteur: Geertje Hek
la Voie-du-Coin 7
1218 Grand-Saconnex
Zwitserland

verdediging@nieuwarchief.nl

Voorbeeld: grafische matroïden

Een belangrijk voorbeeld is dat van grafische matroïden, matroiden die afkomstig zijn van een graaf. Zij $G=(V,E)$ een graaf met V de verzameling punten (vertices) en E de verzameling kanten (edges), en bekijk de verzameling I van deelverzamelingen van E die geen cykel in G bevatten (zo'n verzameling kanten heet een 'bos'). Met een beetje moeite kun je laten zien dat (E,I) aan de eigenschappen 1, 2, en 3 hierboven voldoet, en dus een matroïde is. Als G samenhangend is, dan zijn de bases van deze matroïde precies de opspannende bomen van G, en de rang van de matroide is $|V|-1$.

Een matroïde heet lineair als hij kan worden gerealiseerd als de matroïde behorende bij een verzameling vectoren, zoals hierboven beschreven. Je kunt laten zien dat iedere grafische matroïde lineair is. Andersom kun je je afvragen, en Whitney deed dat al, of iedere matroïde lineair is. (Het antwoord is nee.)

Grote matroïden

Met de opkomst van *data science* en grote netwerken van informatie kwamen grote matroïden steeds meer in de belangstelling. In zijn proefschrift behandelt Van der Pol twee vragen over zulke grote matroïden. Ten eerste: Hoe hard groeit het aantal matroïden op grondverzameling $\{1,2,\ldots,n\}$ wanneer $n\to\infty$? Of algemener, hoe hard groeit het aantal matroïden op E als functie van $|E|$? Hierop slaat het woord 'Enumeration' uit de ondertitel. Ten tweede: Wat voor eigenschappen kunnen we verwachten als $n$groot wordt?

Een exact antwoord op de eerste vraag is lastig te vinden, maar je kunt wel onder- en bovengrenzen maken. Van der Pol: "Zowel een onder- als een bovengrens bestond al, maar er zat nog een groot gat tussen. Wij vermoedden dat de bovengrens verbeterd zou kunnen worden, en dit hebben we gedaan."

Korte omschrijvingen en telefoonnummers

Een probleem bij het vinden van goede bovengrenzen is dat de gebruikelijke omschrijvingen van matroïden lastig te analyseren zijn. Het belangrijkste ingrediënt van Van der Pols argument bestaat daarom uit een 'korte' omschrijving van een matroïde. Hiervoor heeft hij het verband tussen enumeratie en complexiteit uitgebuit: als je voor iedere matroïde M op E een unieke 'korte' omschrijving $f(M)$ kunt vinden, en je kunt het aantal omschrijvingen begrenzen, dan kun je ook het aantal mogelijke matroïden begrenzen. Met dit argument heeft Van der Pol de vorige bovengrens, die ruim veertig jaar oud was, verbeterd. Een echte doorbraak.

Van der Pol illustreert zijn idee voor een korte omschrijving als volgt: stel dat je in een land bent, en je wilt weten hoeveel mensen daar wonen. Het is lastig om die vraag exact te beantwoorden. Maar als je weet dat iedereen een 4-cijferig telefoonnummer heeft, dan weet je wel dat er nooit meer dan 10000 mensen in dat land kunnen wonen. Hier zijn twee dingen van belang: (1) iedereen heeft een uniek telefoonnummer (anders tel je sommige mensen niet mee, of juist dubbel), en (2) je telefoonnummer is niet langer dan noodzakelijk (als iedereen een 10-cijferig telefoonnummer heeft, dan weet je wel dat er hooguit 10^{10} mensen in het land wonen, maar dat is een beroerde bovengrens).

Typische eigenschappen

De tweede vraag in Van der Pols proefschrift was de vraag hoe een 'typische' matroïde eruitziet. Ofwel: als je alle matroïden in een zak zou stoppen en er willekeurig één uit zou halen, wat voor eigenschappen kun je dan verwachten?

Bewezen stellingen zijn van de vorm "bijna alle matroïden hebben eigenschap X", wat wil zeggen dat de fractie van matroïden met eigenschap X naar 1 gaat als de grootte van de grondverzameling E naar oneindig gaat. Het lastige bij het bewijzen van dit soort stellingen is dat we niet goed weten hoe we een 'random' matroïde kunnen maken. Van der Pol buitte daarom zijn korte omschrijvingsmethode verder uit. Soms heeft de afwezigheid van een bepaalde eigenschap X tot gevolg dat zijn omschrijving nog veel korter is. Het aantal matroïden zónder die eigenschap is dan veel kleiner dan het totaal aantal matroïden en dus moet het aantal matroïden mét eigenschap X juist groot zijn. De eigenschap X is daarmee typisch.

Van der Pol gebruikt wederom de telefoonanalogie: Veronderstel dat er 5000 mensen in het land wonen, en neem verder aan dat mensen in de hoofdstad allemaal een nummer hebben dat met 0 begint. Er wonen dan hooguit 1000 mensen in de hoofdstad. Je weet dan ook dat minstens 80 procent van de mensen juist buiten de hoofdstad woont.

Een van de belangrijkste resultaten uit Van der Pols proefschrift is een afschatting van het typische aantal bases in een matroïde. Dit resultaat is een gevolg van de bovengenoemde korte omschrijving. Een aantal van zijn beste resultaten over typische eigenschappen van matroïden volgt relatief makkelijk uit dit algemene resultaat.

Afwisseling

Van der Pol vond vooral de afwisseling tijdens zijn promotietijd prettig: de afwisseling tussen onderwijs geven, zelf een college volgen, en het doen van onderzoek, maar ook de afwisseling die het werken aan verschillende projecten bracht. Die afwisseling hielp ook om over de moeilijkere momenten heen te komen. Lang worstelen met een probleem kan demotiverend werken, aldus Van der Pol. Als het onderzoek lastig was hielp het hem om meerdere ijzers in het vuur te hebben. En dan was het uiteindelijk extra motiverend als hij na lang worstelen een probleem toch wist te temmen.

Zijn project was multidisciplinair, met een kansrekeningcomponent en een discrete-wiskundecomponent, wat ook uitdagingen met zich meebracht. Maar zijn begeleiders roemen Van der Pols talent om de verschillende disciplines in de wiskunde te kunnen verbinden.

Met die begeleiders was het heel prettig samenwerken. Het kantoor van Rudi Pendavingh bevond zich naast dat van Van der Pol, en hij kon altijd bij hem binnenlopen. Het dagelijks samen lunchen en af en toe ondernemen van gezamenlijke activiteiten met de andere promovendi en postdocs uit de gang droeg ook bij aan de goede sfeer. Van der Pol was ook betrokken bij de oprichting van de PhD/PDEng Council van de faculteit, een club die workshops en sociale activiteiten organiseert.

Als mooie momenten noemt Van der Pol nog een aantal mooie reizen die hij gemaakt heeft om andere universiteiten te bezoeken en conferenties bij te wonen: vaak uitputtend, maar hij heeft er veel leuke contacten aan overgehouden. En, hoewel hij daar tijdens het schrijven zelf wel eens anders over dacht, vond hij ook het schrijven van zijn proefschrift heel leuk. Een groot aantal van de resultaten was al opgeschreven, maar het schrijven van het proefschrift gaf hem de kans de teksten nog eens extra te polijsten en met elkaar te verbinden. Met succes: bij zijn promotie zei James Oxley, een van de autoriteiten binnen het vakgebied, dat hij in zijn 35 jaar als onderzoeker nog nooit zo'n goed proefschrift had gezien.