Kan een sudoku vastlopen op het laatst in te vullen veld? **Piet Lemmens** dacht intuïtief van niet, maar wilde zeker weten of dat ook echt zo was. Hier volgt een verslag van zijn zoektocht naar het antwoord op die intrigerende vraag.

Onoplosbare sudoku's

Inleiding

Laatst riep een mij dierbaar familielid uit dat haar sudoku-puzzel vastliep op het laatste nog in te vullen veld. In mijn overmoed zei ik dat zulks niet mogelijk was, maar bij nader inzien moest ik er toch even over nadenken. Inderdaad, van een latijns vierkant (een 'sudoku' zonder de eis op de drie-bij-drie vierkanten) wist ik dat het niet kan vastlopen op een laatste in te vullen veld. Een en ander was voor mij aanleiding voor nader onderzoek naar het voorkomen van lege velden die niet meer kunnen worden ingevuld. Dit is een verslag van mijn exploratie.

Definities

Omdat ik het wil hebben over niet volledig invulbare sudoku-diagrammen, kies ik voor definities die op het eerste gezicht afwijken van de meer gebruikelijke. Een *sudoku-stand* is een vierkant diagram van 81 velden in negen rijen en negen kolommen en negen drie-bij-drie vierkanten (zoals gebruikelijk) waarin sommige (eventueel alle of geen) velden zijn ingevuld, elk met een van de cijfers 1 tot en met 9, zo dat in geen rij, in geen kolom, en in geen drie-bij-drie vierkant een cijfer meervoudig voorkomt. Een sudoku-stand heet *maximaal* als er geen lege velden meer ingevuld kunnen worden met behoud van de regels.

Ouder dan de sudoku's zijn de latijnse vierkanten, waarin alleen de rijen en kolommen een rol spelen. Een *LV-stand* is een vierkant diagram van 81 velden in negen rijen en negen kolommen waarin sommige (eventueel alle of geen) velden zijn ingevuld, elk met een van de cijfers 1 tot en met 9, zo dat in geen rij en in geen kolom een cijfer meervoudig voorkomt. Een LV-stand heet *maximaal* als er geen lege velden meer ingevuld kunnen worden met behoud van de regels.

Voorbeelden

In de figuren 1 en 2 staan voorbeelden van maximale sudoku-standen met twee of drie lege velden. Voor de hand liggend is de vraag naar het grootste aantal lege velden in een maximale sudoku-stand. Ik kan geen voorbeeld vinden met meer dan veertig lege velden, maar heb geen bewijs kunnen vinden dat veertig het maximum is.

7	1	6	8	3	4	9	-	5
8	9	-	2	1	5	7	3	4
3	4	5	7	6	9	8	2	1
4	3	8	5	9	6	1	7	2
2	5	9	1	4	7	3	8	6
1	6	7	3	2	8	5	4	9
9	7	4	6	5	3	2	1	8
6	2	3	9	8	1	4	5	7
5	8	1	4	7	2	6	9	3

fig. 1 Maximale sudoku-stand met twee lege velden.

8	1	6	7	3	4	9	-	5
-	9	7	2	1	5	8	3	4
3	4	5	8	6	9	7	2	1
4	3	8	5	9	6	1	7	2
2	5	9	1	4	7	3	8	6
7	6	-	3	2	8	5	4	9
9	7	4	6	5	3	2	1	8
6	2	3	9	8	1	4	5	7
5	8	1	4	7	2	6	9	3

fig. 2 Maximale sudoku-stand met drie lege velden.

Een andere vraag is of elk aantal tussen twee en het grootste aantal lege velden gerealiseerd kan worden met een maximale sudoku-stand. De figuren 3 en 4 laten maximale sudoku-standen met 36 lege velden zien. Merk op dat beide figuren dezelfde configuratie van lege velden hebben. In figuur 3 speelt de sudoku-regel over de vierkanten geen rol: het is tevens een maximale LV-stand met 36 lege velden.

Bijzonderheden van figuur 4 zijn dat elk cijfer vijf keer voorkomt, en dat het geen maximale LV-stand is. Als

LV-stand is het volledig in te vullen zoals bijvoorbeeld in figuur 5.

1	-	-	2	-	-	3	-	-
2	-	-	3	-	-	1	-	-
3	-	-	1	-	-	2	-	-
-	4	7	-	5	8	-	6	9
-	5	8	-	6	9	-	4	7
-	6	9	-	4	7	-	5	8
-	7	4	-	8	5	-	9	6
-	8	5	-	9	6	-	7	4
-	9	6	-	7	4	-	8	5

fig. 3 Maximale LV-stand met 36 lege velden.

1	-	-	4	-	-	7	-	-
2	-	-	5	-	-	8	-	-
3	-	-	6	-	-	9	-	-
-	4	7	-	1	8	-	6	2
-	5	8	-	2	9	-	4	3
-	6	9	-	3	7	-	5	1
-	7	4	-	8	1	-	2	6
-	8	5	-	9	2	-	3	4
-	9	6	-	7	3	-	1	5

fig. 4 Geen maximale LV-stand, elk cijfer komt vijf keer voor.

1	2	3	4	5	6	7	8	9
2	3	1	5	6	4	8	9	7
3	1	2	6	4	5	9	7	8
5	4	7	9	1	8	3	6	2
6	5	8	7	2	9	1	4	3
4	6	9	8	3	7	2	5	1
9	7	4	3	8	1	5	2	6
7	8	5	1	9	2	6	3	4
8	9	6	2	7	3	4	1	5

fig. 5 Figuur 4 volledig ingevuld als LV stand.

Een voorbeeld van een maximale sudoku-stand met veertig lege velden is weergegeven in figuur 6. Dit is tevens een maximale LV-stand!

1	-	2	3	-	-	4	-	-
-	5	-	-	6	7	-	8	9
-	6	-	-	9	8	-	5	7
4	-	3	1	-	-	2	-	-
-	8	-	-	5	9	-	7	6
-	7	-	-	8	6	-	9	5
2	-	1	4	-	-	3	-	-
3	-	4	2	-	-	1	-	-
-	9	-	-	7	5	-	6	8

fig. 6 Maximale sudoku-stand met veertig lege velden.

Met leeg te blijven velden kunnen ook sudoku-puzzels worden gemaakt. Vul bijvoorbeeld de sudoku-stand in figuur 7 verder in, zo dat een maximale sudoku-stand ontstaat waarin de twee van een minteken voorziene velden precies de lege velden zijn.

5	2						-	
4	-	1	6					
	6		3				2	
		8					7	9
6				8	4		3	2
				7				
				4	5	3		
8				1		7	9	6

fig. 7 In te vullen als oefening voor de lezer...

Stellingen

Wanneer we het hebben over de vierkanten van een sudoku-stand, bedoelen we daarmee de drie-bij-drie vierkanten.

Eerst nog een definitie: in een sudoku-stand noemen we het *bereik* van cijfer C de vereniging van alle velden in de vierkanten, de rijen en de kolommen waarin C voorkomt.

Stelling 1

In een maximale sudoku-stand ligt elk vierkant in het bereik van elk van de cijfers 1 tot en met 9.

Bewijs:

Als vierkant V niet in het bereik ligt van cijfer C, dan heeft V een leeg veld (want C komt er niet in voor), en daarin kan C worden ingevuld, in strijd met de maximaliteit. [ ]

Stelling 2

In een maximale sudoku-stand komt elk van de cijfers 1 tot en met 9 minstens drie keer voor.

Bewijs:

Zij C een cijfer dat hoogstens twee keer voorkomt. Die voorkomens zijn dan gesitueerd in hoogstens twee van de negen vierkanten. Er is dus een vierkant waarvan geen veld in het bereik van C zit, in strijd met stelling 1. [ ] Stelling 2 geldt niet voor elke maximale LV-stand, getuige figuur 8.

In een sudoku-stand ligt een veld v alleen in het bereik van de cijfers die voorkomen in het vierkant dat v bevat, in de kolom van v en in de rij van v. Daaruit kunnen we conclusies trekken over configuraties van lege velden die niet kunnen voorkomen in een maximale sudoku-stand.

1	-	-	-	-	-	-	-	-
-	2	3	4	5	6	7	8	9
-	3	4	5	6	7	8	9	2
-	4	5	6	7	8	9	2	3
-	5	6	7	8	9	2	3	4
-	6	7	8	9	2	3	4	5
-	7	8	9	2	3	4	5	6
-	8	9	2	3	4	5	6	7
-	9	2	3	4	5	6	7	8

fig. 8 Een maximale LV-stand.

Stelling 3
In een maximale sudoku-stand kan geen rij zeven of meer lege velden hebben.

Bewijs:
Bij een rij R met zeven of meer lege velden is er een vierkant V dat drie lege velden van R bevat. In V vinden we dus hoogstens zes cijfers, en in R hoogstens twee, samen acht cijfers. Er is dus een cijfer C dat niet voorkomt in R en niet in V. Opdat C niet kan worden ingevuld in een van de velden van R binnen V, moet elk van de drie kolommen die door V gaan het cijfer C bevatten buiten V, hetgeen onmogelijk is omdat deze kolommen slechts door twee andere vierkanten gaan. [ ]

Stelling 4
In een maximale sudoku-stand kan geen kolom zeven of meer lege velden hebben.

Bewijs:
Verwissel rij en kolom in het bewijs van stelling 3. [ ]

Stelling 5
In een maximale sudoku-stand kan geen vierkant zeven of meer lege velden hebben.

Bewijs:
Bij zeven of meer lege velden in een vierkant V zijn er drie lege velden van V die in een rij R liggen. In V vinden we dus hoogstens twee en in R hoogstens zes, samen hoogstens acht cijfers. De argumentatie is nu verder dezelfde als in het bewijs van stelling 3. [ ]

Een gevolg van stelling 2 is dat een maximale sudoku-stand hoogstens $54 (= 81 - 9 \times 3)$ lege velden kan hebben, en dat voor vierenvijftig lege velden noodzakelijk ieder cijfer drie keer moet voorkomen. De volgende redenering laat echter zien dat er geen vierenvijftig lege velden kunnen zijn in een maximale sudoku-stand waarbij de cijfers 1 en 2 elk drie keer voorkomen. Immers, ieder leeg veld moet in het bereik liggen van elk van de cijfers 1 tot en met 9. Na een beetje experimenteren zien we dat er maximaal zevenenvijftig velden in het bereik van cijfer 1 kunnen liggen. Dat doet zich voor als de drie vierkanten die 1 bevatten horizontaal en verticaal vrij van elkaar liggen. Bij die zevenenvijftig velden zijn ook de drie voorkomens van 1 meegeteld. Buiten de van 1 voorziene velden zijn er dus maximaal vierenvijftig velden in het bereik van 1. Dezelfde overweging geldt voor het cijfer 2. Uit de situatieschets volgt dat de twee verzamelingen van elk vierenvijftig velden niet kunnen samenvallen. Daarmee is bewezen:

Stelling 6
Het aantal lege velden in een maximale sudoku-stand is kleiner dan vierenvijftig. [ ]

Ten koste van veel 'gepeuter' kan ik stelling 6 wel scherper krijgen, maar daarbij kom ik niet in de buurt van veertig lege velden zoals in figuur 6. Er valt dus nog wel iets te doen voor de geïnteresseerde lezer! In het kader van het inventariseren van onmogelijke configuraties – zoals de stellingen 3, 4 en 5 – vermeld ik zonder bewijs:

Stelling 7
Er is geen maximale sudoku-stand met vijf lege velden in elke rij, elke kolom en elk vierkant. [ ]
Bij LV-standen gaat het eenvoudiger omdat de vierkanten daar geen rol spelen.

Stelling 8
Zij n een geheel getal met $5 \le n \le 9$. Als n het grootste aantal lege velden per rij of kolom is in een maximale LV-stand, dan zijn er in totaal hoogstens $2n(9 - n)$ lege velden.

Bewijs:
We veronderstellen dat het grootste aantal lege velden optreedt in rij R, en dat 1 tot en met $9\text{-}n$ de in rij R voorkomende cijfers zijn. In elk van de n kolommen van de lege velden in rij R moeten dan de cijfers $9 - n + 1$ tot en met 9 voorkomen en dus hoogstens $9 - n$ lege velden. In elk van de overige $9 - n$ kolommen komen hoogstens n lege velden voor. [ ]
Een gevolg van stelling 8 is:

Stelling 9
Het grootste aantal lege velden in een maximale LV-stand is veertig.

Bewijs:
We hebben met figuur 6 reeds een voorbeeld van een maximale LV-stand met veertig lege velden. Bij meer dan zesendertig lege velden moet er een rij zijn met meer dan vier lege velden, dus stelling 8 is van toepassing. Het maximum van $2n(9 - n)$ voor gehele n met $5 \le n \le 9$ treedt op bij $n = 5$ en is dan 40. [ ]

Een mooi gestructureerde maximale LV-stand met veertig lege velden is te zien in figuur 9. Dan nu de stelling dat een sudoku-stand met slechts één leeg veld niet maximaal kan zijn, dus dat dit veld ook kan worden ingevuld. We behandelen eerst een LV-stand met slechts één leeg veld. Dat dit veld ook kan worden ingevuld is een gevolg van de huwelijksstelling van P. Hall uit 1935 (zie Lemmens & Springer (1992); stelling 12.4), maar een direct bewijs is in dit geval wellicht aardiger.

1	2	3	4	-	-	-	-	-
2	3	4	1	-	-	-	-	-
3	4	1	2	-	-	-	-	-
4	1	2	3	-	-	-	-	-
-	-	-	-	5	6	7	8	9
-	-	-	-	6	7	8	9	5
-	-	-	-	7	8	9	5	6
-	-	-	-	8	9	5	6	7
-	-	-	-	9	5	6	7	8

fig. 9 Maximale LV-stand met veertig lege velden.

Stelling 10
Een LV-stand met slechts één leeg veld kan niet maximaal zijn.

Bewijs:
We duiden de positie van een veld aan met (m, n) waarin m het nummer van de rij en n dat van de kolom is, dus $1 \le m \le 9$ en $1 \le n \le 9$. Door eventuele permutaties van rijen en kolommen mogen we aannemen dat veld $(9, 9)$ leeg is. Voor de laatste kolom is dan een cijfer – zeg C – over, dat zou kunnen worden ingevuld in veld (9, 9). Dat resulteert precies dan in een LV-stand als C niet reeds voorkomt in de laatste rij, dus in een van de velden $(9, n)$ met $n \le 8$.

Om dat te verifiëren, merken we op dat de eerste acht rijen volledig zijn ingevuld, dus in elk van die rijen komt C voor, echter niet in de laatste kolom. In het subdiagram van de velden (m, n) met $m \le 8$ en $n \le 8$ staan dus acht cijfers C, en bijgevolg komt C in elk van de acht kolommen en in elk van de acht rijen van dat subdiagram voor. Maar dan kan C inderdaad niet ook nog voorkomen in een van de velden $(9, n)$ met $n \le 8$. [ ]

Stelling 11
Een sudoku-stand met slechts één leeg veld kan niet maximaal zijn.

Bewijs:
Zij V het vierkant waarin het lege veld v voorkomt. Volgens stelling 10 is er een cijfer C dat niet voorkomt in de rij of de kolom van v. We moeten dus nog aantonen dat C ook niet voorkomt in V. We weten dat alle andere vierkanten reeds volledig zijn ingevuld, dus in elk daarvan komt C voor. Als C ook zou voorkomen in V, dan zouden er een rij en een kolom zijn waarin C tweemaal optreedt. De lezer kan dit zelf gemakkelijk nagaan. Daarom komt C niet reeds voor in V, en kan dus in v worden ingevuld. [ ]

Toegift
De eerder genoemde huwelijksstelling van Hall impliceert dat een LV-stand, waarin een aantal rijen volledig is ingevuld en nog geen velden buiten die rijen, aangevuld kan worden tot een volledig ingevulde LV-stand. Dit geldt niet algemeen voor een sudoku-stand, zoals figuur 10 aantoont. Immers in het eerste veld van de zesde rij moet wegens de eerste kolom een van de cijfers 6 tot en met 9 worden ingevuld, maar die staan reeds in het vierkant van dat veld!

1	8	9	2	3	4	5	6	7
2	6	7	1	5	8	9	4	3
3	4	5	9	6	7	2	1	8
4	9	8	3	2	1	6	7	5
5	7	6	4	8	9	3	2	1

fig. 10 Niet verder in te vullen als sudoku...

Piet Lemmens,
Mathematisch Instituut, Universiteit Utrecht

Literatuur

Lemmens, P.W.H., & Springer, T.A. (1992). *Hoofdstukken uit de Combinatoriek*. Utrecht: Epsilon Uitgaven, 25.