

Arjen Doelman

Korteweg-de Vries Instituut
Universiteit van Amsterdam
Plantage Muidergracht 24
1018 TV Amsterdam
doelman@science.uva.nl

Vakantiecursus 2002

Dimensie in theorie en experiment

Jaarlijks organiseert het Centrum voor Wiskunde en Informatica (CWI) onder auspiciën van de Nederlandse Vereniging van Wiskundeleraars een vakantiecursus voor wiskundeleraars en andere belangstellenden. Bij deze gelegenheid verschijnt steeds een syllabus met teksten bij de voordrachten. Het NAW heeft een serie gestart waarin geselecteerde teksten uit recente syllabi worden geplaatst. Het vierde artikel is afkomstig uit de syllabus 'Experimentele wiskunde' bij de Vakantiecursus 2001. Het onderwerp is het begrip dimensie, met name zoals het optreedt in dynamische systemen. Arjen Doelman is hoogleraar in de toegepaste analyse en de dynamische systemen aan het Korteweg-de Vries Instituut van de Universiteit van Amsterdam. Hij werkt onder andere op het onderzoeksgebied dat patroonvorming in systemen met een fysische, chemische of biologische achtergrond tracht te verklaren.

Wij leven in een driedimensionale wereld (voor zover wij weten). Wat betekent dat eigenlijk? In een driedimensionale ruimte zijn er drie onafhankelijke 'vrijheidsgraden' waarmee je een positie in die ruimte kan vastleggen. In de wiskunde maak je die drie vrijheidsgraden vaak expliciet, bijvoorbeeld met Cartesische coördinaten: de positie van een punt leg je vast met drie coördinaten ten opzichte van de drie (x -, y - en z -) assen van het Cartesische coördinatensysteem. Op eenzelfde manier heeft een driedimensionaal voorwerp ook drie 'onafhankelijke richtingen': een lengte, een breedte en een hoogte. Zo'n voorwerp heeft een volume, dat wordt gemeten in kubieke meters, m^3 , en dat bij eenvoudige, blokvormige voorwerpen verkregen wordt uit het product van de lengte, breedte en hoogte.

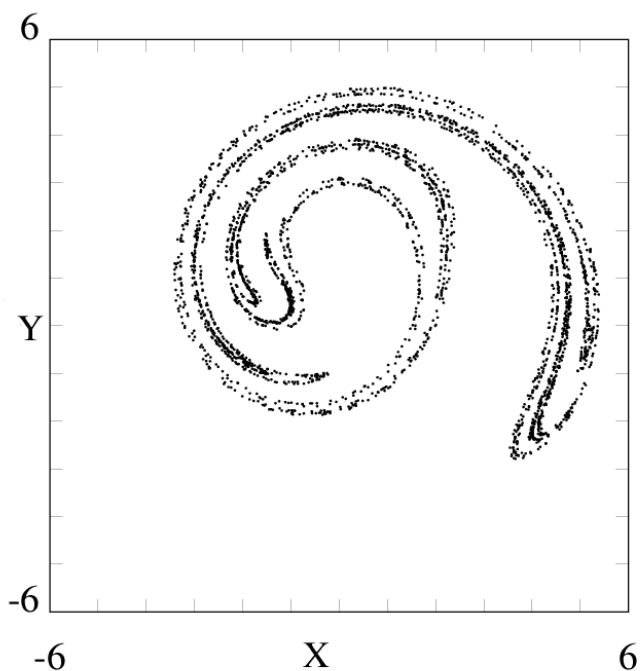
In een tweedimensionale ruimte, een oppervlak, zijn er maar twee van die vrijheidsgraden, die in Cartesische coördinaten worden vastgelegd met de x -as en de y -as. Op eenzelfde manier heeft

een tweedimensionaal object ook bijvoorbeeld wel een lengte en een breedte, maar geen hoogte. Dit heeft natuurlijk niets te maken met het al dan niet 'plat' zijn van het oppervlak: ook op een gekromd oppervlak zijn er volgens onze intuïtie twee, en niet meer dan twee, coördinaten nodig om een positie vast te leggen. Een tweedimensionaal voorwerp heeft een oppervlak dat wordt gemeten in vierkante meters, m^2 . Voor de meest eenvoudige tweedimensionale objecten wordt dit bepaald door het product lengte \times breedte.

Neem je nog zo'n vrijheid weg, dan houd je een ééndimensionale ruimte, ofwel een kromme, over. Een kromme heeft alleen een lengte, en die wordt gemeten in meters, $m = m^1$. Wiskundig kan je ook die laatste vrijheid wegnemen, wat overblijft is een nuldimensionaal object, een mathematisch punt, dat geen lengte, breedte of hoogte heeft. Het enige wat je met zo'n nuldimensionaal object kan doen is het tellen, ofwel bepalen uit hoeveel (losse) punten het bestaat.

Tot zover lijkt ons intuïtieve idee van het begrip dimensie goed te werken. Onze intuïtie kan ons echter in de problemen brengen: als één punt dimensie 0 heeft, dan heeft een verzameling bestaande uit twee punten natuurlijk ook dimensie 0. Hetzelfde geldt voor iedere verzameling van eindig veel punten. Maar zijn lijnen en vlakken niet verzamelingen van oneindig veel punten? Hoe en wanneer maak je dan de stap van 0 naar 1 of naar 2? Moeten de overgangen per se discontinu zijn? Kunnen verzamelingen misschien ook een dimensie hebben die tussen 0 en 1 of 1 en 2 inzit? Wat zijn dat dan voor verzamelingen?

Zulke in eerste instantie abstracte vragen spelen ook een belangrijke rol buiten de zuiver wiskundige context. Met de opkomst van de computer is het duidelijk geworden dat zeer veel natuurlijke processen gerelateerd kunnen worden aan verzame-



Figuur 1 Een zogenaamde chaotische attractor die optreedt in een model voor de eb en vloed-cyclus. Dit plaatje is verkregen met behulp van een computer-simulatie van het bijbehorende dynamische systeem.

lingen waarvan het niet a priori duidelijk is wat de dimensie zou moeten zijn. Deze verzamelingen ontstaan uit computer-experimenten met de dynamische systemen die deze processen wiskundig beschrijven. In dit verhaal nemen we als voorbeeld een representatie van de benadering van een zogenaamde chaotische attractor (zie figuur 1). Deze attractor speelt een rol bij het bestuderen van getijde-systemen en geeft aan hoe de op het eerste gezicht zeer regelmatige eb en vloed-cyclus chaotisch gemodelleerd kan worden, zoals we aan het eind van dit artikel zullen zien. De computer geeft ons vanzelfsprekend alleen een eindige hoeveelheid punten, maar de attractor bestaat 'natuurlijk' uit oneindig veel punten. De vraag is nu: wat is de dimensie van deze attractor? De puntenwolk heeft op meerdere plaatsen al een duidelijke lijnenstructuur, dus het lijkt erop dat de dimensie (minstens) 1 is. Op sommige plaatsen liggen deze lijnen echter zo dicht bij elkaar dat ze een tweedimensionaal gebiedje lijken te vullen en je dus eerder zou denken dat de dimensie van de attractor 2 moet zijn. Of is het zo dat deze 'tweedimensionale gebiedjes' toch weer blijken te bestaan uit vele aparte, maar zeer dichtbij elkaar liggende lijnen, zoals gebruikelijk is bij 'fractale objecten' (zie [4])? Maar als dat nu eens oneindig veel lijnen blijken te zijn?

Wat is eigenlijk de dimensie van deze puntenwolk? Hoe zou je de dimensie van dit wellicht 'fractale' object moeten 'meten'?

We zullen hier een methode presenteren die ons in staat stelt een soort (!) dimensie, ofwel de *limietcapaciteit*, van een verzameling, of een puntenwolk te bepalen, of te meten (vaak met behulp van de computer). Deze methode is in essentie op niets anders dan een nauwkeurige wiskundige (her)formulering van het intuïtieve dimensiebegrip gebaseerd. We zullen laten zien dat de methode voor mooie 'gladde' verzamelingen ook het antwoord geeft waarop we rekenen: een 'gewone' lijn heeft gewoon dimensie 1, etc. Echter, als we de methode gebruiken om de dimensie

van een fractale verzameling, zoals een Cantorverzameling, een Kochkromme of een chaotische attractor te bepalen, dan vinden we uitkomsten die niet per se geheeltallig zijn. Met andere woorden, zulke verzamelingen kunnen een zogenaamde 'fractale dimensie' hebben.

Hoewel het wiskundige fundament van de methode op een limietprocedure berust, kan de methode ook op een eenvoudige manier worden toegepast op 'experimentele' en dus eindige dataverzameling, zoals de output van een computer-experiment of waarnemingen van natuurlijke structuren. Omdat de verzameling in zo'n geval maar uit eindig veel elementen bestaat, of omdat de meetnauwkeurigheid begrensd is, kan nu de wiskundige limiet niet worden genomen. Echter, ook zonder de limiet te nemen kan de methode een benadering of een indicatie van de dimensie van een dataverzameling geven. Wij zullen bijvoorbeeld de dimensie van de attractor in figuur 1 schatten en vinden dat deze (ongeveer) gelijk aan 1.6 is.

Verzamelingen en limietprocedures

In deze, en de komende paragraaf, geven we voorbeelden van verzamelingen die aan de hand van een limietprocedure geconstrueerd worden. Mede als gevolg daarvan is het intuïtief niet duidelijk wat de dimensie van deze verzamelingen zou kunnen of moeten zijn. Met andere woorden: de verzamelingen die we in deze en de komende paragraaf construeren tonen aan hoe gebrek ons intuïtieve dimensie begrip eigenlijk is.

Allereerst construeren we de (standaard) Cantorverzameling C ; C is een deelverzameling van het gesloten interval $[0, 1]$. De Cantorverzameling C wordt stapsgewijs opgebouwd. Voor de eerste stap definiëren we:

$$V_0 = [0, 1].$$

Merk op dat de lengte van V_0 , L_0 , gelijk is aan 1. Vervolgens definiëren we V_1 door V_0 in 3 gelijke stukken te verdelen en het middelste stuk, het interval $(\frac{1}{3}, \frac{2}{3})$, 'weg te snijden' (zie figuur 2):

$$V_1 = \left[0, \frac{1}{3}\right] \cup \left[\frac{2}{3}, 1\right];$$

V_1 heeft (totale) lengte $L_1 = \frac{2}{3}$. We definiëren V_2 door hetzelfde te doen met elk van de twee componenten van V_1 : deel het in 3 gelijke stukken in en snij het middelste stuk weg. Dus,

$$V_2 = \left[0, \frac{1}{9}\right] \cup \left[\frac{2}{9}, \frac{1}{3}\right] \cup \left[\frac{2}{3}, \frac{7}{9}\right] \cup \left[\frac{8}{9}, 1\right],$$

met totale lengte $L_2 = \frac{2}{3} \times \frac{2}{3} = \frac{4}{9}$. In het algemeen wordt V_{n+1} uit V_n geconstrueerd door ieder component van V_n in 3 gelijke stukken te delen en de middelste daarvan weg te snijden (figuur 2). Het gevolg hiervan is dat $L_{n+1} = \frac{2}{3}L_n$. De Cantorverzameling C is nu gedefinieerd door

$$C = \lim_{n \rightarrow \infty} V_n.$$

Hoewel C is geconstrueerd uit elementen die stuk voor stuk 'duidelijk' dimensie 1 hebben, is het nu niet meer zo duidelijk of C zelf ook nog wel dimensie 1 heeft: de lengte van C is bijvoorbeeld niet gedefinieerd, want $L_n \rightarrow 0$ als $n \rightarrow \infty$, zodat C dus wel 'lengte 0' moet hebben. Toch is C nog een behoorlijk 'volle' verzameling, die veel meer punten bevat dan alleen de randpunten van

de vorm $\frac{m}{3^n}$ van de intervallen V_n (met $0 \leq m \leq 3^n$). De Cantorverzameling C bevat zelfs *overaftelbaar* veel punten. Dit betekent dat C ‘net zoveel’ punten bevat als het interval $[0, 1]$, hetgeen natuurlijk merkwaardig is omdat we C juist geconstrueerd hebben door oneindig veel stukken uit $[0, 1]$ weg te snijden (zie opmerking 1).

Een andere belangrijke eigenschap van C is dat de verzameling *zelfgelijkvormig* is, dat wil zeggen dat je door op de goede manier in te zoomen op C , steeds weer in essentie dezelfde verzameling C terugvindt. Bekijk hiervoor bijvoorbeeld de linkerhelft van C , dat wil zeggen de verzameling $C_1 = C \cap [0, \frac{1}{3}] \subset C$. Na vermenigvuldiging van ieder element met een factor 3 is deze verzameling identiek aan C zelf! Hetzelfde geldt voor iedere doorsnede van C met een van de componenten van een verzameling V_n : na verschuiving en vermenigvuldiging met de factor 3^n is deze identiek aan de oorspronkelijke verzameling C .

Natuurlijk kan je op een eenvoudige manier andere, maar soortgelijke, Cantorverzamelingen C^α construeren. Hiervoor kijken we nog eens naar de constructie van C . Bij iedere stap in de limietprocedure wordt ieder deelinterval van V_n , dat dus een lengte heeft van 3^{-n} , in 3 stukken gedeeld. Ieder deel heeft telkens dezelfde lengte, te weten $3^{-(n+1)}$. We kunnen dit natuurlijk ook anders doen: we beginnen weer bij V_0 en geven het middelste deelinterval een lengte α , waarbij α een willekeurig getal in $[0, 1]$ is. Vervolgens maken we de twee deelintervallen aan de uiteinden precies even lang, hetgeen betekent dat ze een lengte van $\frac{1}{2}(1 - \alpha)$ hebben. Net als bij de constructie van C snijden we nu het middelste interval weg, zodat

$$V_1^\alpha = \left[0, \frac{1}{2}(1 - \alpha)\right] \cup \left[\frac{1}{2}(1 + \alpha), 1\right]$$

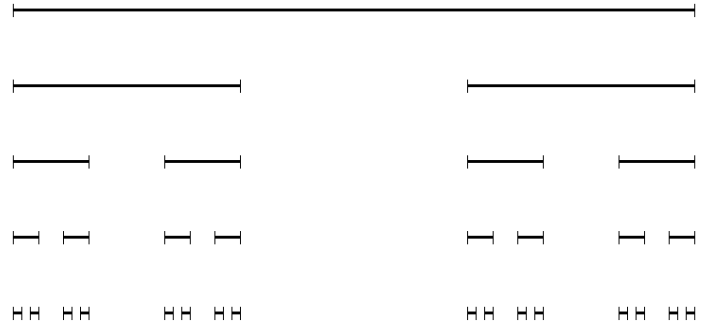
overblijft. Vervolgens snijden we uit beide deelintervallen van V_1^α het ‘middelste α -gedeelte’, etc. De *middelste- α Cantorverzameling* C^α wordt nu weer verkregen door de limiet te nemen. Net als in het oorspronkelijke geval geldt dat de totale lengte van V_n^α, L_n^α , naar 0 gaat als $n \rightarrow \infty$, tenminste, als $\alpha > 0$:

$$L_n^\alpha = (1 - \alpha)L_{n-1}^\alpha = (1 - \alpha)^n \rightarrow 0 \text{ als } n \rightarrow \infty,$$

dus, de ‘totale lengte’ van C^α is, net als in het oorspronkelijke geval, 0 (als $\alpha > 0$).

Merk op dat $C^0 = [0, 1]$, $C^1 = \{0, 1\}$ (ofwel: 2 losse punten) en dat $C = C^{\frac{1}{3}}$. In figuur 3 is (een benadering van) het zogenaamde *Cantorgordijn* geschetst: α loopt van 0 tot 1 van boven naar beneden, bovenaan staat C^0 , onderaan C^1 . Voor kleine α liggen de elementen van C^α zo’n beetje homogeen verdeeld in $[0, 1]$ en lijkt C^α wel eens een verzameling van dimensie 1 te kunnen zijn (hetgeen het natuurlijk duidelijk is voor $\alpha = 0$). Voor α dicht bij 1 bestaat C^α alleen uit ijle puntenverzamelingen dicht bij de randpunten 0 en 1: C^α lijkt nu meer op een verzameling met dimensie 0 (de dimensie van C^1 is natuurlijk 0).

De verzamelingen C^α lijken dus een mooie continue ‘verbinding’ te geven tussen een overduidelijk 0-dimensionale verzameling ($\alpha = 1$) en een ééndimensionale verzameling ($\alpha = 0$). Dit maakt het wel erg natuurlijk om het bestaan van verzamelingen met een niet-gehele dimensie te accepteren: indien de dimensie van C^α altijd 0 of 1 moet zijn, bij welke $\alpha \in [0, 1]$ zou dan de overgang van 0 naar 1 moeten zijn?



Figuur 2 De eerste 5 stappen in de constructie van de Cantor-verzameling C : van boven naar onderen zien we V_0, V_1, V_2, V_3 en V_4 .

Opmerking 1. Om in te zien dat $C = C^{\frac{1}{3}}$ overaftelbaar is, dat wil zeggen, ‘evenveel’ punten bevat als het interval $[0, 1]$, merken we op dat er een grote overeenkomst bestaat tussen de bovenstaande meetkundige constructie van C en de drietallige representatie van getallen in het interval $(0, 1)$. De vraag is dan eerst: hoe bepaal je de drietallige representatie van een getal $p \in [0, 1]$? Dat gaat als volgt:

- Kies $p \in [0, 1]$ en bereken $q_1 = 3p$.
- Definieer $p_1 \in \{0, 1, 2\}$ door $p_1 \leq q_1 = 3p < p_1 + 1$ (p_1 heet de ‘entier’ van q_1).
- Bereken $q_2 = 3(q_1 - p_1)$.
- Definieer $p_2 \in \{0, 1, 2\}$ door $p_2 \leq q_2 < p_2 + 1$.
- Bereken $q_3 = 3(q_2 - p_2)$.
- Definieer $p_3 \in \{0, 1, 2\}$ door $p_3 \leq q_3 < p_3 + 1$.
- Etcetera.

Merk op dat nu per constructie geldt dat

$$p = \frac{p_1}{3} + \frac{p_2}{3^2} + \frac{p_3}{3^3} + \dots = \sum_{j=1}^{\infty} \frac{p_j}{3^j}.$$

Het getal $p \in (0, 1)$ ligt dus eenduidig vast als we het rijtje $\{p_1, p_2, \dots\}$ — met $p_j \in \{0, 1, 2\}$ — kennen en kan nu worden gerepresenteerd door $p =_{(3)} 0.p_1p_2p_3 \dots$. Bijvoorbeeld

$$\frac{1}{2} =_{(3)} 0.11111\dots, \quad \frac{1}{3} =_{(3)} 0.1, \quad \frac{1}{4} =_{(3)} 0.020202\dots$$

Echter, verschillende rijtjes $\{p_1, p_2, \dots\}$ kunnen wel hetzelfde getal $p \in [0, 1]$ representeren:

$$\begin{aligned} \frac{1}{3} &=_{(3)} 0.1 =_{(3)} 0.02222\dots, \\ \frac{2}{3} &=_{(3)} 0.2 =_{(3)} 0.12222\dots, \\ 1 &=_{(3)} 1.0 =_{(3)} 0.22222\dots, \text{ etc.} \end{aligned}$$

Dit fenomeen kennen we natuurlijk al van de gewone 10-tallige representatie van getallen: $1 =_{(10)} 0.9999\dots$ etc. We kunnen nu terugkeren naar de constructie van de Cantorverzameling C en constateren dat

$$p \in V_1 \Leftrightarrow p_1 \neq 1$$

met andere woorden: $p \in V_1$ impliceert dat $p_1 = 0$ of 2 , en $p_1 = 0$



Figuur 3 Een grafische voorstelling van het Cantorgordijn; α loopt van boven naar beneden van 0 tot 1. Merk op dat in deze figuur de verzameling $V_3^\alpha \supset C^\alpha$ is getekend voor alle waarden $\alpha \in [0, 1]$.

of 2 betekent dat $p \in V_1$. Analogoos zien we dat

$$p \in V_n \Leftrightarrow p_1, p_2, \dots, p_n \neq 1 \text{ ofwel } p_1, p_2, \dots, p_n \in \{0, 2\}.$$

Hieruit volgt dat C als volgt kan worden gekarakteriseerd:

$$C = \{p = {}_{(3)}0.p_1p_2\dots, \text{ met } p_j \in \{0, 2\}\}.$$

Merk op dat we op deze manier veel 'concrete elementen' van C — die geen randpunt van een van de basis-intervallen V_m zijn — kunnen vinden, bijvoorbeeld: $\frac{1}{4} = {}_{(3)}0.020202\dots \in C$. Aan de hand van de zojuist geconstateerde niet-uniciteit van de representatie zien we tevens dat deze randpunten $\frac{m}{3^n}$, die natuurlijk ook allemaal elementen van C zijn, ook op deze manier gekarakteriseerd kunnen worden. Als we ons nu realiseren dat ten opzichte van het tweetalig stelsel het interval $[0, 1]$ als volgt kan worden gerepresenteerd:

$$[0, 1] = \{p = {}_{(2)}0.p_1p_2\dots, \text{ met } p_j \in \{0, 1\}\},$$

dan zien we hoe we aan ieder punt van C een uniek punt van $[0, 1]$ kunnen koppelen (en omgekeerd): verander elke '2' in de drietalige representatie van een element p van C in een '1' (of andersom). Op deze manier wordt bijvoorbeeld $\frac{1}{3} \in [0, 1]$ gekoppeld aan $\frac{1}{4} \in C$. Op deze manier heeft C dus 'precies evenveel' elementen als $[0, 1]$!

Opmerking 2. Op eenzelfde manier als in opmerking 1 kan worden aangetoond dat alle in deze paragraaf geconstrueerde Cantorverzamelingen C^α — natuurlijk met $0 \leq \alpha < 1$ — overaftelbaar zijn, dat wil zeggen in essentie net zoveel punten hebben als het interval $[0, 1]$.

Een kromme met onbegrensde lengte

De Cantorverzamelingen uit de vorige paragraaf kunnen nooit meer dan ééndimensionaal zijn, ze zijn tenslotte geconstrueerd als deelverzamelingen van $[0, 1]$. In deze paragraaf construeren we, wederom met behulp van een limietprocedure, een 'kromme' waarvan de lengte 'oneindig' moet zijn. Dit zou kunnen betekenen dat deze kromme een dimensie heeft die groter is dan 1.

We starten weer met het (gesloten) eenheids-interval, $W_0 = [0, 1]$, delen het weer in 3 gelijke stukken op en snijden het middelste interval, $(\frac{1}{3}, \frac{2}{3})$, weg. Echter, we vervangen dit interval nu door een tent- of driehoek-constructie, bestaande uit 2 lijnstukken met dezelfde lengte als het oorspronkelijke stuk, $\frac{1}{3}$ (zie figuur 4). Merk op dat de zo geconstrueerde 'tent' met de hoekpunten $(\frac{1}{3}, 0)$, $T_1^1 = (\frac{1}{2}, \frac{1}{6}\sqrt{3})$, $(\frac{2}{3}, 0)$ een gelijkzijdige driehoek is.

De verzameling W_1 bestaat nu dus uit 4 lijnstukken met lengte $\frac{1}{3}$: de verbindingen tussen de punten $(0, 0)$ en $(\frac{1}{3}, 0)$, $(\frac{1}{3}, 0)$ en T_1^1 , T_1^1 en $(\frac{2}{3}, 0)$, en tenslotte $(\frac{2}{3}, 0)$ en $(1, 0)$. Dit betekent dat de lengte M_1 van W_1 gelijk is aan $\frac{4}{3}$.

Vervolgens passen we dezelfde 'tent-constructie procedure' toe op ieder van de 4 deel-lijnstukken van W_1 . Hieruit ontstaat de verzameling W_2 , die uit 16 lijnstukjes met lengte $\frac{1}{9}$ bestaat; deze lijnstukjes verbinden 17 punten met elkaar, waaronder de 4 nieuwe 'tent-punten'

$$T_2^1 = \left(\frac{1}{6}, \frac{1}{18}\sqrt{3}\right), T_2^2 = \left(\frac{1}{3}, \frac{1}{9}\sqrt{3}\right), \\ T_2^3 = \left(\frac{2}{3}, \frac{1}{9}\sqrt{3}\right), T_2^4 = \left(\frac{5}{6}, \frac{1}{18}\sqrt{3}\right).$$

De totale lengte M_2 van W_2 is dus $\frac{16}{9} = \frac{4}{3} \times \frac{4}{3}$. In de volgende stap wordt natuurlijk op ieder van de 16 lijnstukjes weer een gelijkzijdig tentje 'opgezet', zodat W_3 uit 64 lijnstukjes van lengte $\frac{1}{27}$ bestaat: $M_3 = 64 \times \frac{1}{27} = (\frac{4}{3})^3$. Etcetera (zie figuur 4).

De *Kochkromme* K ontstaat nu door het nemen van de limiet $n \rightarrow \infty$ in W_n . Vanzelfsprekend geldt dat

$$\text{lengte}(K) = \lim_{n \rightarrow \infty} \text{lengte}(W_n) = \lim_{n \rightarrow \infty} \left(\frac{4}{3}\right)^n = \infty.$$

Echter, ondanks het feit dat K 'oneindig lang' is, is K wel keurig begrensd: K past bijvoorbeeld in z'n geheel in de rechthoek met hoekpunten $(0, 0)$, $(0, \frac{1}{6}\sqrt{3})$, $(1, \frac{1}{6}\sqrt{3})$, $(1, 0)$. Maar als K begrensd is en toch oneindig lang is, is K dan wel een kromme, dat wil zeggen, is K dan wel ééndimensionaal? Intuïtief was het hebben van een (eindige) lengte toch de bepalende eigenschap van ééndimensionale objecten. De 'kromme' K lijkt zeker ook niet tweedimensionaal te zijn, al was het maar omdat de Kochkromme 'duidelijk' geen oppervlak lijkt te hebben.

Net als de Cantorverzamelingen uit de vorige paragraaf is ook K weer zelfgelijkvormig. Ieder stukje van K dat opgebouwd wordt uit een van de deel-lijnstukjes van een W_n kan door middel van schuiven, draaien en 'opblazen' geïdentificeerd worden met de oorspronkelijke 'hele' K .

Opmerking 3. Natuurlijk kan je ook hier weer algemenere structuren dan K definiëren, net als in de vorige paragraaf waarin C^α de Cantorverzameling C generaliseerde. In het geval van een generaliseerde Kochkromme ontstaat er wel het 'probleem' — dat helemaal geen probleem hoeft te zijn — dat deze 'kromme' zichzelf gaat doorsnijden.

De limietcapaciteit van een verzameling

In deze paragraaf, en in de rest van dit verhaal, beperken we ons tot deelverzamelingen van het (platte) vlak. De verzamelingen die we beschouwen kunnen dus maximaal dimensie 2 hebben. Alles wat we hier bespreken kan worden ggeneraliseerd naar N dimensies.

Het basis-idee achter het begrip *limietcapaciteit* is dat de dimensie van een begrensde verzameling $U \subset \mathbf{R}^2$ bepaald hoeveel ‘kleine vierkantjes’ er nodig zijn om U te overdekken. Hierbij bedoelen we met ‘overdekken’ dat ieder element van U in minstens een van de overdekkende vierkantjes zit. De keuze om met vierkantjes te overdekken en dus bijvoorbeeld niet met cirkelschijfjes is in eerste instantie willekeurig.

Het lijnstuk $[0, 1]$ kan natuurlijk (precies) worden overdekt met 2 vierkantjes $S(\frac{1}{2})$, dat zijn vierkantjes waarvan beide zijden lengte $\frac{1}{2}$ hebben. In het algemeen definiëren we $S(\varepsilon)$ als een vierkantje met zijden van lengte ε . Er zijn 4 vierkantjes $S(\frac{1}{4})$ en, in het algemeen, $n = n^1$ vierkantjes $S(\frac{1}{n})$ nodig om $[0, 1]$ te bedekken. Natuurlijk zijn er n^2 vierkantjes $S(\frac{1}{n})$ nodig om het ‘grote’ vierkant $S(1)$ te overdekken. Aan de andere kant is er altijd maar $1 = n^0$ vierkantje $S(\frac{1}{n})$ nodig om het punt $(0, 0)$ te overdekken.

Als we nu ε schrijven in plaats van $\frac{1}{n}$ en $N(\varepsilon) = N(\varepsilon; U)$ introduceren als het getal dat aangeeft hoeveel vierkantjes $S(\varepsilon)$ er nodig zijn om een gegeven verzameling $U \subset \mathbf{R}^2$ te overdekken, dan kunnen we concluderen dat $N(\varepsilon) = \varepsilon^{-2}$ voor het duidelijk tweedimensionale vierkant $S(1)$, dat $N(\varepsilon) = \varepsilon^{-1}$ voor het 1-dimensionale interval $[0, 1]$ en dat $N(\varepsilon) = 1 = \varepsilon^0$ voor een 0-dimensionaal punt. We zullen verderop zien dat er nog een aantal haken en ogen aan deze conclusies zitten, maar, het lijkt duidelijk dat de dimensie van een verzameling U iets te maken heeft met de (absolute waarde van de) exponent van ε in de uitdrukking voor $N(\varepsilon; U)$. Nu hebben de zojuist gegeven drie voorbeelden natuurlijk wel een zeer simpele, en bovenal regelmatige, structuur. Als gevolg daarvan zijn deze voor relatief grote waarden van ε al zeer ‘netjes’ te overdekken. Dit geldt natuurlijk niet voor alle waarden van ε : $S(1)$ wordt bijvoorbeeld ‘sluitend’ overdekt met 4 $S(\frac{1}{2})$'s, maar voor waarden van ε die ook maar iets kleiner zijn dan $\frac{1}{2}$ geldt meteen dat $N(\varepsilon; S(1)) \geq 7$ (maar niet automatisch 9, zie hieronder), met als gevolg dat er een aantal vierkantjes $S(\varepsilon)$ zullen zijn die flink ‘uitsteken’ — we komen hier verderop nog uitgebreid op terug. Met name voor verzamelingen met een wat meer grillig karakter zal gelden dat de overdekking passender/nauwkeuriger wordt naarmate de waarde van ε afneemt. Gebaseerd op deze constatering definiëren we de *limietcapaciteit* $\text{limcap}(U)$ van een verzameling $U \subset \mathbf{R}^2$ als volgt:

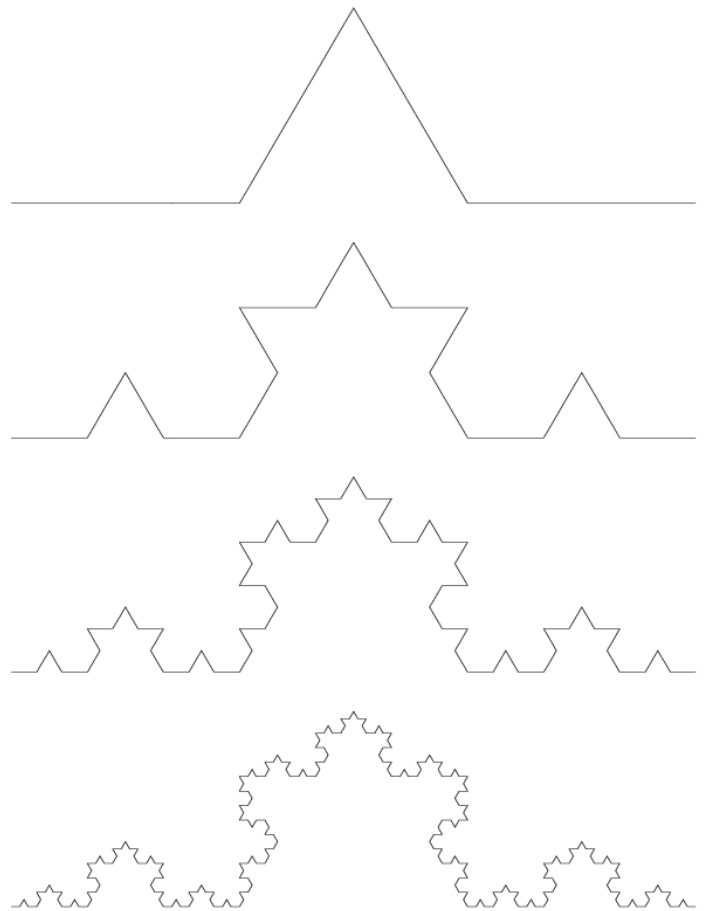
$$\text{limcap}(U) = -\lim_{\varepsilon \rightarrow 0} \frac{\log N(\varepsilon; U)}{\log \varepsilon}. \tag{1}$$

Merk op dat in ieder geval geldt dat $0 \leq \text{limcap}(U) \leq 2$, voor iedere (begrensde) $U \subset \mathbf{R}^2$ (omdat zeker geldt dat $0 \leq N(\varepsilon; U) \leq C\varepsilon^{-2}$, voor zekere positieve C). Opvallend is dat het niet uitmaakt welk grondtal we gebruiken voor de logaritmen in (1), de extra omrekeningsfactor (bijvoorbeeld van e naar 10) treedt in teller en noemer op en valt dus weg. Omdat $N(\varepsilon; S(1)) = \varepsilon^{-2}$ geldt nu natuurlijk dat

$$\text{limcap}(S(1)) = -\lim_{\varepsilon \rightarrow 0} \frac{\log \varepsilon^{-2}}{\log \varepsilon} = 2 = \text{dim}(S(1))$$

en, natuurlijk,

$$\begin{aligned} \text{limcap}([0, 1]) &= 1 = \text{dim}([0, 1]), \\ \text{limcap}(P = (p_1, p_2)) &= 0 = \text{dim}(P). \end{aligned}$$



Figuur 4 De eerste 5 stappen (de verzamelingen W_0, W_1, W_2, W_3, W_4) in de constructie van de Kochkromme K .

Dus, voor dit soort eenvoudige verzamelingen lijkt de formule (1) een correcte, berekenbare uitdrukking voor de dimensie van een verzameling te geven.

Echter, tot nu toe hebben we de situatie iets eenvoudiger voorgesteld dan hij in werkelijkheid is. We moeten eigenlijk het getal $N(\varepsilon; U)$ precieser definiëren: het is het *minimale* aantal vierkantjes $S(\varepsilon)$ waarmee de verzameling $U \subset \mathbf{R}^2$ in z'n geheel overdekt kan worden. Deze voorwaarde dat $N(\varepsilon)$ minimaal is kan a priori aanleiding geven tot een flink gepuzzel.

Laten we daarom nog eens terugkeren naar het vierkant $S(1)$. Zoals we al opmerkten geldt namelijk in het algemeen niet dat $N(\varepsilon; S(1)) = \varepsilon^{-2}$, dat is alleen het geval voor $\varepsilon = \frac{1}{n}$ voor zekere $n \in \mathbf{N}$. Als we ε gelijk aan $\frac{1}{2} - \sigma$ kiezen, en aannemen dat σ ‘klein genoeg’ is dan zien we meteen dat $4 < N(\varepsilon) \leq 9$. Echter, $N(\varepsilon)$ is kleiner dan 9. Om dit in te zien maken we eerst van 4 vierkantjes $S(\varepsilon)$ te samen het vierkant $S(2\varepsilon)$, en geven dat de hoekpunten $(0, 0), (0, 2\varepsilon), (2\varepsilon, 2\varepsilon), (2\varepsilon, 0)$. Als we deze $S(2\varepsilon)$ over $S(1)$ met de hoekpunten $(0, 0), (0, 1), (1, 1), (1, 0)$ leggen dan blijft er nog een strookje ter breedte van σ over langs de zijdes van $S(1)$ die de punten $(1, 0), (1, 1)$ en $(0, 1)$ met elkaar verbinden. De ‘zuijnigste’ methode om dit (smalle) strookje te overdekken is om 3 vierkantjes $S(\varepsilon)$ te nemen: eentje leg je met het middelpunt op $(1, 1)$ zo dat de diagonalen samenvallen met de twee onbedekte zijdes van $S(1)$, de andere 2 kunnen nu de resterende overgeble-

ven stukjes van de nog open strookjes overdekken door ook deze een kwartslag te draaien (en dus de diagonalen weer parallel aan de zijdes van $S(1)$ te kiezen). Dus, we hebben 7 $S(\varepsilon)$'s nodig gehad om $S(1)$ te overdekken. Merk op dat dit alleen kan voor ε behoorlijk dicht bij $\frac{1}{2}$: zodra ε te klein wordt (of σ te groot) zal het niet meer lukken om het overgebleven strookje op deze manier te overdekken (het volgt vrij eenvoudig dat er zeker ruimte zal overblijven als $\varepsilon \leq \frac{1}{3}\sqrt{2} \approx 0.471\dots$). Verder hebben we nog niet bewezen dat er nu ook werkelijk geldt dat $N(\varepsilon) = 7$ voor ε dicht genoeg bij $\frac{1}{2}$: wie weet lukt het toch met 6 vierkantjes $S(\varepsilon)$ als we de overdekking geheel anders aanpakken.

Nu willen we een 'werkbaar' — en voor een computer berekenbaar — dimensie begrip hanteren, en dan willen we bovengstaand combinatorisch gepuzzel natuurlijk vermijden. En dat kan ook met het in (1) geïntroduceerde begrip limietcapaciteit. We zijn namelijk niet in de exacte waarde van $N(\varepsilon)$ geïnteresseerd, maar alleen in het (exponentiële) groeigedrag van $N(\varepsilon)$ als functie van ε als $\varepsilon \rightarrow 0$. We kijken daarom nog eens naar $S(1)$. Als ε voor zekere n tussen $\frac{1}{n+1}$ en $\frac{1}{n}$ in zit, dan zal in ieder geval gelden dat

$$n^2 \leq N(\varepsilon) < (n+1)^2.$$

Ofwel, voor ieder $\varepsilon (\leq 1)$ geldt dat

$$\left(\frac{1}{\varepsilon} - 1\right)^2 \leq N(\varepsilon; S(1)) \leq \left(\frac{1}{\varepsilon} + 1\right)^2.$$

Dus

$$-\lim_{\varepsilon \rightarrow 0} \frac{\log\left(\frac{1}{\varepsilon} - 1\right)^2}{\log \varepsilon} \leq \text{limcap}(S(1)) \leq -\lim_{\varepsilon \rightarrow 0} \frac{\log\left(\frac{1}{\varepsilon} + 1\right)^2}{\log \varepsilon}.$$

Beide limieten zijn gelijk aan 2: het maakt dus helemaal niet uit of we de preciese waarde van $N(\varepsilon; S(1))$ kennen, het vinden van een geschikte boven- en ondergrens is al voldoende. Met andere woorden, het nemen van de limiet maakt, in combinatie met de logaritmen, het exact bepalen van $N(\varepsilon)$ overbodig! De combinatorische subtiliteiten worden er als het ware uitgefilterd.

De vraag is natuurlijk of dit geldt voor alle mogelijke verzamelingen $U \subset \mathbf{R}^2$. Allereerst beschouwen we verzamelingen U waarvan het 'intuïtief duidelijk' is dat $\dim(U) = 2$. Wat betekent dat? In ieder geval dat U zo is dat er een $\delta > 0$ te vinden is zodat er een (wellicht zeer klein vierkantje) $S(\delta)$ bestaat dat in z' n geheel in U past. Lukt het niet om zo'n $S(\delta)$ te vinden, dan is het namelijk meteen de vraag of U wel tweedimensionaal is. Echter, alleen al om $S(\delta)$ te overdekken zijn er minstens een zekere constante $\times \varepsilon^{-2}$ vierkantjes $S(\varepsilon)$ nodig (als ε klein genoeg is, en zeker kleiner dan δ), zoals we zojuist gezien hebben (voor $\delta = 1$, maar dat maakt natuurlijk niet echt uit). Er geldt dus dat $N(\varepsilon; U) \geq C\varepsilon^{-2}$ (voor zekere C), zodat $\text{limcap}(U) = 2$ (omdat $\text{limcap}(U) \leq 2$ voor alle begrensde $U \subset \mathbf{R}^2$). We concluderen dat $\dim(U) = \text{limcap}(U) = 2$ voor elk van deze verzamelingen U .

Hetzelfde geldt voor verzamelingen die 'duidelijk' 1-dimensionaal zijn. Laat $\gamma \subset \mathbf{R}^2$ zo'n duidelijk 1-dimensionale (begrensde) verzameling zijn: laten we aannemen dat γ een zogenaamde gladde kromme is. In dat geval heeft γ een welgedefinieerde lengte L_γ (zie opmerking 4). Kies nu ε klein genoeg en plaats een vierkantje $S(\varepsilon)$ over γ . De doorsnede $\gamma \cap S(\varepsilon)$ zal er 'bijna' uitzien als een

recht lijnstukje (door steeds dieper in te zoomen — ofwel: door ε steeds kleiner te kiezen — wordt het net alsof γ lokaal niet langer gekromd is — dit komt natuurlijk omdat we hebben aangenomen dat γ 'glad' is). We kunnen $S(\varepsilon)$ dus zo neerleggen dat $\gamma \cap S(\varepsilon)$ ingaat en uitgaat via twee diagonaal tegenover elkaar liggende hoekpunten van $S(\varepsilon)$; $\gamma \cap S(\varepsilon)$ valt dan dus bijna samen met een diagonaal van $S(\varepsilon)$, hetgeen betekent dat de lengte van $\gamma \cap S(\varepsilon)$ 'bijna gelijk is' aan $\varepsilon\sqrt{2}$. Door alle $S(\varepsilon)$ op deze manier neer te leggen, en ze netjes op elkaar te laten aansluiten zien we dat

$$N(\varepsilon; \gamma) \approx \frac{L_\gamma}{\sqrt{2}\varepsilon} \text{ als } \varepsilon \rightarrow 0,$$

waarbij het nu wel duidelijk is dat deze $N(\varepsilon; \gamma)$ daadwerkelijk het minimale aantal vierkantjes $S(\varepsilon)$ is (hetgeen ook betekent dat de hierboven gevonden $N(\varepsilon) = \varepsilon^{-1}$ voor $\gamma = [0, 1]$ veel te grof was!). We zien nu dus ook dat voor iedere gladde kromme γ geldt dat $\dim(\gamma) = \text{limcap}(\gamma) = 1$.

Ook voor verzamelingen P_N bestaande uit eindig veel, ofwel N , losse punten geldt natuurlijk dat $\dim(P) = \text{limcap}(P) = 0$ (eenvoudigweg omdat $N(\varepsilon; P_N) = N$, onafhankelijk van ε , als ε klein genoeg is). Dus, we mogen concluderen dat het begrip limietcapaciteit inderdaad overeenkomt met ons intuïtieve dimensie begrip, tenminste, voor verzamelingen U waarvoor er geen twijfel kan bestaan over de dimensie van U (zie opmerking 5). In de volgende paragrafen zullen we zien dat de interpretatie van het begrip dimensie als limietcapaciteit (1) van het dimensie begrip ook een zeer berekenbare (en benaderbare) grootheid maakt.

Opmerking 4. Een gladde kromme γ wordt vastgelegd door een afbeelding $F = (f_1, f_2) : [0, 1] \rightarrow \mathbf{R}^2$, zodat $\gamma = \{(f_1(t), f_2(t)) : t \in [0, 1]\}$. Voor beide componenten f_i moet nu gelden dat de afgeleide bestaat en continu is als functie van t , voor iedere $t \in [0, 1]$. Dit betekent dat L_γ bestaat en eindig is (zie bijvoorbeeld [8]). Merk ook op dat dit impliceert dat de Kochkromme K onmogelijk door zo'n gladde afbeelding F gegeven kan worden.

Opmerking 5. De limietcapaciteit is zeker niet de 'beste' generalisatie van het dimensie begrip. Het is bijvoorbeeld zeer eenvoudig om te laten zien dat de limietcapaciteit van de verzameling van rationale getallen tussen 0 en 1, $\mathbf{Q} \cap (0, 1)$, gelijk is aan die van het hele interval $[0, 1]$: $\text{limcap}(\mathbf{Q} \cap (0, 1)) = 1$. Echter, $\mathbf{Q} \cap (0, 1)$ bevat slechts aftelbaar oneindig veel punten (zie bijvoorbeeld [8]) en is dus zeer 'ijl'. In het bijzonder geldt dat $\mathbf{Q} \cap (0, 1)$ dus veel en veel minder punten dan de Cantorverzameling C^α bevat, voor iedere $\alpha < 1$ — zie opmerkingen 1 en 2 en figuur 3 (!). In de volgende paragraaf zullen we laten zien dat $0 < \text{limcap}(C^\alpha) < 1$ voor iedere $\alpha \in (0, 1)$ en het feit dat $\text{limcap}(\mathbf{Q} \cap (0, 1)) = 1$ is dan dus niet erg 'logisch'. Dit probleem kan worden weggevangen door het begrip Hausdorff-dimensie in te voeren, dat gezien kan worden als een generalisatie van de limietcapaciteit en waarvan de waarde samenvalt met die van de limietcapaciteit voor alle andere in dit artikel genoemde verzamelingen $U \subset \mathbf{R}^2$, zie [1, 7, 9].

Opmerking 6. We hebben de wiskundige vraag of de in de definitie van de limietcapaciteit genomen limiet voor iedere mogelijke $U \subset \mathbf{R}^2$ bestaat niet gesteld. Het is inderdaad niet uitgesloten dat deze limiet niet bestaat, en dat het quotiënt in (1) tussen twee uitersten blijft op en neer schommelen als $\varepsilon \rightarrow 0$. In dat geval is

de limietcapaciteit gedefinieerd als de maximale waarde die het quotiënt aanneemt in de limiet $\varepsilon \rightarrow 0$.

Verzamelingen met fractionele dimensie/limietcapaciteit

De limietcapaciteit van de Cantorverzameling C kan op een eenvoudige manier expliciet berekend worden.

Allereerst constateren we dat $N(\varepsilon_n; C)$ direct te bepalen is als $\varepsilon_n = \frac{1}{2}\sqrt{2} \times (\frac{1}{3})^n$ voor zekere $n \geq 1$. Voor iedere waarde van n kunnen we de verzameling $V_n \supset C$ volledig overdekken met 2^n vierkantjes $S(\varepsilon_n)$ door ieder van de 2^n deel-lijnstukjes van V_n precies passend te bedekken met de diagonaal van een vierkantje $S(\varepsilon_n)$ — deze heeft voor deze waarde van ε namelijk exact de lengte $\frac{1}{3^n}$ (zie figuur 2). Het is duidelijk dat deze waarde van $N(\varepsilon, C)$ voor deze keuzes van ε inderdaad minimaal moet zijn. Als we ons nu even niets aantrekken van alle andere mogelijke waarden van ε dan vinden we:

$$\text{limcap}(C) = - \lim_{\varepsilon_n \rightarrow 0} \frac{\log N(\varepsilon_n; C)}{\log \varepsilon_n} =$$

$$\lim_{n \rightarrow \infty} \frac{\log 2^n}{\log 3^n \sqrt{2}} = \lim_{n \rightarrow \infty} \frac{n \log 2}{n \log 3 + \log \sqrt{2}} = \frac{\log 2}{\log 3} \approx 0.6309 \dots$$

Het preciese bewijs dat inderdaad $\text{limcap}(C) = \log 2 / \log 3 \in (0, 1)$ (!) heeft nog wat meer voeten in de aarde, omdat we voorsnog alleen maar naar a priori heel speciale waarden van ε gekeken hebben. Maar, net als in het geval van het vierkant $S(1)$, volgt er dat deze andere waarden van ε geen leidende orde invloed hebben op de exponent van ε in de uitdrukking voor $N(\varepsilon; C)$ in de limiet $\varepsilon \rightarrow 0$. We zullen hier niet verder op ingaan.

Dus, de Cantorverzameling C heeft een dimensie/limietcapaciteit die niet gelijk is aan een geheel getal! Mede om deze reden wordt C wel een *fractale verzameling* genoemd. De verzameling C is dus duidelijk meer dan een verzameling losse punten (die zou dimensie 0 hebben), maar, zoals verwacht, ook minder dan een ‘kromme’, want dan zou de dimensie 1 hebben moeten zijn.

Op eenzelfde manier kunnen we de dimensie/limietcapaciteit van de Kochkromme K bepalen. Ook hier trekken we ons weinig meer aan van de vraag of we wel voor iedere waarde van ε de optimale — en dus ‘echte’ — waarde van $N(\varepsilon; K)$ kennen. We zullen ons verder ook niet meer bezig houden met deze kwestie, uiteindelijk kan met wat meer moeite worden laten zien dat dit — alweer — geen invloed heeft op de dimensie van K . Net als bij het bepalen van $\text{limcap}(C)$ maken we ook hier weer gebruik van het feit dat K aan de hand van een limietprocedure geconstrueerd is.

We kiezen wederom $\varepsilon_1 = \frac{1}{2}\sqrt{2} \times \frac{1}{3}$ en zien dat we de verzameling W_1 netjes kunnen overdekken met 4 vierkantjes $S(\varepsilon_1)$ (zie figuur 4): ieder van de 4 lijnstukjes waaruit W_1 is opgebouwd laten we precies samenvallen met een diagonaal van een overdekkende vierkantje $S(\varepsilon_1)$, die per constructie exact de goede lengte ($= \frac{1}{3}$) heeft (dit is natuurlijk volledig analoog aan de overdekking van V_n hierboven). Anders dan bij C moeten we nu nog wel expliciet nagaan of deze overdekking van W_1 ook een overdekking van K geeft. Dat was in het geval van C automatisch duidelijk, omdat $C \subset V_n$, voor iedere n . Dit is natuurlijk niet het geval voor K en de W_n 's (figuur 4) en heeft daarom nog enige voeten in de aarde.

De overdekking van W_1 is zeker ook een overdekking van W_2 , want de 4 nieuwe ‘tentjes’ die in de volgende iteratie-stap op ie-

der van de deel-lijnstukjes van W_1 worden ‘opgezet’ liggen keurig binnen het vierkantje $S(\varepsilon_1)$ dat het deel-lijnstukje overdekt. Dit is eenvoudig na te rekenen, bijvoorbeeld, de ‘top’ $T_2^1 = (\frac{1}{6}, \frac{1}{18}\sqrt{3})$ van het ‘tentje’ op het eerste deel-lijnstukje $[0, \frac{1}{3}]$ van W_1 ligt duidelijk onder het hoekpunt van de overdekkende $S(\varepsilon_1)$ (dat natuurlijk op $(\frac{1}{6}, \frac{1}{6})$ ligt). Dus ligt ook het hele ‘tentje’ binnen de overdekkende $S(\varepsilon_1)$. Dit geldt dan automatisch ook voor de eerste iteraties van de 3 andere deel-lijnstukjes van W_1 (deze zijn namelijk op verschuiving en draaiing na identiek aan de iteratie van $[0, \frac{1}{3}]$). De vraag is nu: overdekken de $S(\varepsilon_1)$'s ook de volgende iteraties W_3, W_4 , etc. en uiteindelijk K ? Dit is a priori niet duidelijk: er worden meer en meer ‘tentjes’ dichter en dichter bij de randen van de $S(\varepsilon_1)$'s ‘opgezet’ (figuur 4).

Laten we eerst eens naar de volgende, fijnere, overdekking kijken. Door de zelfgelijkvormigheid van K kunnen we de constructie van zojuist een stapje dieper in het iteratie-proces herhalen: we kiezen $\varepsilon_2 = \frac{1}{2}\sqrt{2} \times (\frac{1}{3})^2$ en overdekken ieder van de 16 deel-lijnstukjes van W_2 met een $S(\varepsilon_2)$ (weer zodat de diagonaal van $S(\varepsilon_2)$ samenvalt met het deel-lijnstukje). Door de gelijkvormigheid in de constructie van K geldt nu automatisch dat deze overdekking ook de volgende iteratie W_3 van W_2 overdekt.

We kunnen nu terugkeren naar de vraag of de overdekking van W_1 door de 4 $S(\varepsilon_1)$'s ook W_3, W_4 , etc. overdekt. We beschouwen hiervoor wederom de overdekking van het eerste deel-lijnstukje $[0, \frac{1}{3}]$ door $S(\varepsilon_1)$. Het is direct expliciet na te gaan dat alle 4 de vierkantjes $S(\varepsilon_2)$ die de W_2 -iteratie van $[0, \frac{1}{3}]$ overdekken binnen $S(\varepsilon_1)$ liggen. Dit betekent dat de overdekking van W_1 door de 4 $S(\varepsilon_1)$'s ook een overdekking geeft van W_3 (omdat de $S(\varepsilon_2)$ overdekking van W_2 hierbinnen ligt en we weten dat deze W_3 overdekt). Analoog zien we een iteratie-stap verder dat de $S(\varepsilon_2)$ overdekking van W_2 ook W_4 overdekt (dit is weer het gelijkvormigheids-argument). Dit betekent dan weer dat de $S(\varepsilon_1)$ overdekking van W_1 de W_4 overdekt. Etcetera.

Deze constructie van ‘geneste overdekkingen’ kunnen we dus in z'n geheel itereren: voor alle $\varepsilon_n = \frac{1}{2}\sqrt{2} \times (\frac{1}{3})^n$ kunnen we de bijbehorende W_n , en alle daar op volgende W_{n+m} 's, overdekken met $4^n S(\varepsilon_n)$'s. Dit betekent dat al deze overdekkingen tevens K overdekken, zodat we kunnen concluderen dat

$$\text{limcap}(K) = - \lim_{\varepsilon_n \rightarrow 0} \frac{\log N(\varepsilon_n; K)}{\log \varepsilon_n} =$$

$$= \lim_{n \rightarrow \infty} \frac{\log 4^n}{\log 3^n \sqrt{2}} = \frac{2 \log 2}{\log 3} \approx 1.2618 \dots$$

De Kochkromme K heeft dus inderdaad een dimensie (of beter: limietcapaciteit) die niet geheeltallig is en in ligt tussen 1 en 2!

De dimensie van de algemene Cantorverzameling C^α zou ook weer op eenzelfde manier als die van C en K kunnen worden bepaald. Echter, er is een makkelijker manier, tenminste, als we de formule (1) nog iets ‘ruimer’ interpreteren. Deze relatie lijkt namelijk te impliceren dat voor ε ‘klein genoeg’

$$N(\varepsilon; U) \approx C\varepsilon^{-\text{limcap}(U)}, \text{ voor zekere } C > 0. \tag{2}$$

Dat dit niet per se zo hoeft te zijn was al in opmerking 6 geconstateerd, echter, voor de meeste verzamelingen U is dit een zeer bruikbare benadering, die we ook in de rest van dit verhaal zullen

gebruiken. Met deze formule, en het feit dat C^α zelfgelijkvormig is, is $\text{limcap}(C^\alpha)$ zonder at teveel expliciet rekenwerk te bepalen.

De linkerkant van de verzameling C^α , de verzameling $C_1^\alpha = C^\alpha \cap V_1^\alpha$, vormt natuurlijk de helft van de gehele C^α en kan dus ook met half zoveel vierkantjes $S(\varepsilon)$ worden overdekt:

$$N(\varepsilon, C_1^\alpha) = \frac{1}{2} N(\varepsilon, C^\alpha).$$

Aan de andere kant kunnen we C_1^α ook 'opblazen', ofwel vermenigvuldigen met een factor, zodat C_1^α precies samenvalt met de oorspronkelijke C^α (C^α is tenslotte zelfgelijkvormig). Omdat de lengte van V_1^α gelijk is aan $\frac{1}{2}(1-\alpha)$, zal de opblaasfactor $\frac{2}{1-\alpha}$ zijn. Door dit opblazen verandert ook de grootte van de overdekkende vierkantjes $S(\varepsilon)$: $S(\varepsilon)$ blaast op tot $S(\frac{2}{1-\alpha}\varepsilon)$. Dus, een overdekking van C_1^α met $S(\varepsilon)$'s geeft door 'opblazing' een overdekking van C^α met $S(\frac{2}{1-\alpha}\varepsilon)$'s, hetgeen betekent dat

$$N(\varepsilon, C_1^\alpha) = N\left(\frac{2}{1-\alpha}\varepsilon, C^\alpha\right)$$

Deze laatste twee formules te samen geven de relatie:

$$\frac{1}{2}N(\varepsilon, C^\alpha) = N\left(\frac{2}{1-\alpha}\varepsilon, C^\alpha\right)$$

Als we nu aannemen dat (2) exact geldt, dan betekent dit dat

$$\frac{1}{2}C\varepsilon^{-d} = C\left(\frac{2}{1-\alpha}\varepsilon\right)^{-d} = C\left(\frac{2}{1-\alpha}\right)^{-d}\varepsilon^{-d}$$

met $d = \text{limcap}(C^\alpha)$.

Hieruit kunnen C en ε^{-d} worden weggedeeld, zodat volgt:

$$d = \text{limcap}(C^\alpha) = \frac{\log 2}{\log 2 - \log(1-\alpha)}.$$

Deze enigszins formele afleiding van $\text{limcap}(C^\alpha)$ kan ook wiskundig hard worden gemaakt, maar daar gaan we hier verder niet op in (zie ook [1, 7, 9]).

We kennen al drie expliciete waarden van $\text{limcap}(C^\alpha)$: $\text{limcap}(C^1) = 0$ (want $C^1 = \{0, 1\}$), $\text{limcap}(C^0) = 1$ (want $C^0 = [0, 1]$) en natuurlijk $\text{limcap}(C^{\frac{1}{3}}) = \log 2 / \log 3$. Deze 3 waardes kloppen keurig met de algemene uitdrukking (al moet voor $\alpha = 1$ wel een limiet genomen worden). Verder valt op dat $\text{limcap}(C^\alpha)$ monotoon afneemt als α van 1 naar 0 loopt (met precies tussenin $\text{limcap}(C^{\frac{1}{2}}) = \frac{1}{2}$!): C^α geeft ons dus de mogelijkheid om voor ieder getal d tussen 0 en 1 een verzameling te construeren waarvan de dimensie/limietcapaciteit precies gelijk is aan d !

Het benaderen van de dimensie van een attractor/puntenwolk

We keren nu terug naar de oorspronkelijke vraag uit de Inleiding: hoe moet je de dimensie (of nu: limietcapaciteit) van een verzameling, of attractor, of puntenwolk, A 'meten'? Het grootste probleem hierbij is dat deze puntenwolken 'in de praktijk' door een computer gegenereerd worden, en dat je dus altijd maar eindig veel van deze punten kent (zie figuur 1). Natuurlijk gaan we er

hierbij vanuit dat deze verzameling, of attractor in het geval van figuur 1, weldegelijk uit oneindig veel punten zal bestaan, het is alleen praktisch onmogelijk om al deze punten te bepalen.

De 'officiële definitie' van de limietcapaciteit, (1), kan hiervoor niet gebruikt worden: er wordt een limiet $\varepsilon \rightarrow 0$ genomen, en voor een verzameling van eindig veel punten zal dat altijd betekenen dat de limietcapaciteit $\equiv 0$ is. Echter, we kunnen wel weer gebruik maken van de benadering als in (2): voor ε klein genoeg, maar niet te klein, zal gelden dat

$$\text{limcap}(A) \approx -\frac{\log \tilde{N}(\varepsilon; A)}{\log \varepsilon}. \quad (3)$$

En hierbij hebben we het ons zelfs nog makkelijker gemaakt. We hebben $N(\varepsilon; A)$ namelijk vervangen door $\tilde{N}(\varepsilon; A)$. Dit betekent dat we niet langer 'gedwongen' willen zijn om te gaan 'spelen' met vierkantjes $S(\varepsilon)$ om zo te proberen de verzameling A zo 'zuinig mogelijk' te overdekken (hetgeen nodig is om $N(\varepsilon; A)$ te bepalen). In plaats daarvan fixeren we gewoon de vierkantjes $S(\varepsilon)$ in een rooster (als op een ruitjespapier, zie figuur 5) en tellen de 'ruitjes' die een stukje van A overdekken. De uitkomst noemen $\tilde{N}(\varepsilon; A)$. Dit getal geeft natuurlijk alleen een bovengrens op $N(\varepsilon; A)$, maar we hebben in de vorige paragraaf toch steeds al gezien dat het hebben van de preciese waarde van $N(\varepsilon; A)$ er niet echt toe doet bij het bepalen van $\text{limcap}(A)$ (zie opmerking 7).

Het 'berekenen' van $\tilde{N}(\varepsilon; A)$ is natuurlijk zeer eenvoudig te programmeren, hetgeen duidelijk maakt waarom deze, a priori grotendeels formele benadering zo populair is. De op deze manier bepaalde dimensie van een object heeft zelfs een eigen naam gekregen: de *box-counting dimension*, een naam die precies zegt wat het is (opmerking 7). De box-counting aanpak geeft ons (en vele met ons) een ideaal handvat om een schatting te maken van attractors als in figuur 1 en maakt van het begrip dimensie een experimenteel meetbare grootheid.

Voor we hier aan gaan beginnen moeten we nog even stilstaan bij een gevaarlijk aspect van deze methode. We beschouwen hiertoe (alweer) het eenheidsvierkant $S(1)$, op een rooster dat is opgebouwd uit vierkantjes $S(0.1)$. Als we de zijdes van $S(1)$ precies samen laten vallen met de assen van het rooster dan is $\tilde{N}(0.1; S(1)) = 100$ en geeft de box-counting aanpak ons met behulp van (3) de volgende 'benadering'

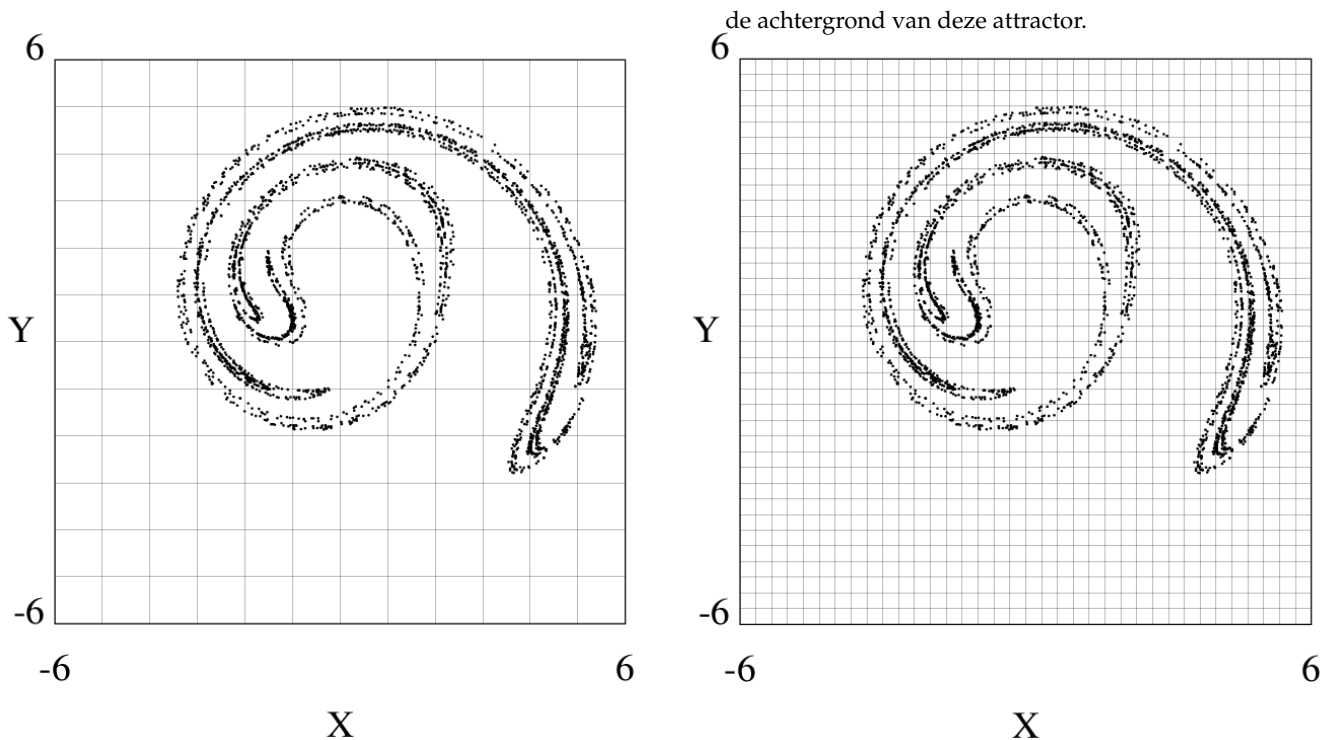
$$\text{limcap}(S(1)) \approx -\frac{\log 100}{\log 0.1} = 2.0$$

Echter, de dimensie van $S(1)$ mag natuurlijk niet veranderen als we $S(1)$ met de een of ander factor 'opblazen'. Maar, als we $S(1)$ twee keer zo groot maken, dan vinden we, op hetzelfde rooster (dat we dus niet mee opblazen):

$$\text{limcap}(S(2)) \approx -\frac{\log 400}{\log 0.1} = 2.602\dots,$$

En, een opblaasfactor $\frac{1}{2}$ geeft (nog steeds op het $S(0.1)$ rooster): de waarde

$$\text{limcap}(S(0.5)) \approx -\frac{\log 25}{\log 0.1} = 1.397\dots,$$



Figuur 5 Twee overdekkingen van de attractor A uit figuur 1. De eerste overdekking geeft de schatting $\text{limcap}(A) \approx 1.62$, de tweede (meer nauwkeurige) $\text{limcap}(A) \approx 1.58$.

Dat zijn nogal verschillen. Waar ligt dat aan? Laten we nog eens naar benadering (2) kijken en deze in de vorm van (3) schrijven:

$$\text{limcap}(A) \approx -\frac{\log \tilde{N}(\varepsilon; A) - \log C}{\log \varepsilon}.$$

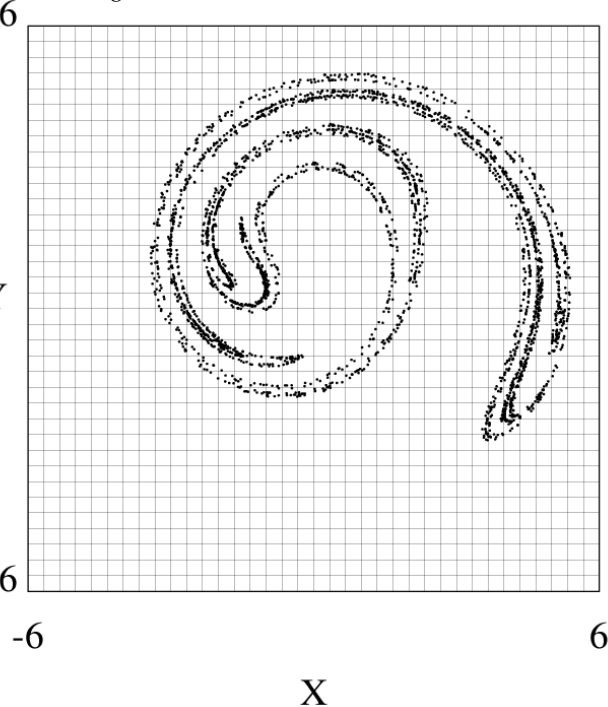
Dus, benadering (3) is veel 'grover' dan (2). Natuurlijk komen beide uitdrukkingen overeen als ε echt heel klein wordt, maar voor niet al te kleine ε heeft de extra term $-\log C$ een significante invloed op de uitkomst. In de bovenstaande voorbeelden is het natuurlijk duidelijk wat de waarden voor C moeten zijn: $C_{S(2)} = 4$, $C_{S(1)} = 1$ en $C_{S(0.5)} = 0.25$. Geen wonder dat $S(1)$ de beste benadering geeft: $\log C_{S(1)} = 0$! Met deze extra correctie-term volgt nu

$$\begin{aligned} \text{limcap}(S(2)) &\approx -\frac{\log 400 - \log 4}{\log 0.1} = 2.0 = \\ &= -\frac{\log 25 - \log 0.25}{\log 0.1} \approx \text{limcap}(S(0.5)). \end{aligned}$$

Dit betekent in de praktijk dat we er voor moeten proberen te zorgen dat we onze attractor A zo schalen dat de waarde C_A — die we niet kennen en eigenlijk ook niet willen kennen — zo dicht mogelijk bij 1 ligt. Het bovenstaande voorbeeld geeft aan dat dit betekent dat we de absolute afmeting van A niet 'te groot' (zoals bij $S(2)$) en niet 'te klein' (zoals bij $S(0.5)$) mogen maken, tenminste, als we de ε niet zo klein kunnen maken als we maar willen (hetgeen eigenlijk altijd het geval is). In de praktijk betekent dit dat we A zo schalen dat A 'ongeveer passend' overdekt wordt door één vierkant $S(1)$.

We kunnen nu de tot dusver ontwikkelde theoretische kennis toepassen op de attractor A in figuur 1. Eerst gaan we kort in op

de achtergrond van deze attractor.



Deze is verkregen met behulp van een computer-simulatie van een model, in de vorm van een dynamisch systeem, dat de modulatie van de getijde-golf in een zeker geïdealiseerd bekken (denk hierbij aan een baai, een fjord of een haven) beschrijft (zie [5]). Met modulatie bedoelen we hier dat de punten in figuur 1 niet de exacte waterstanden representeren, maar de veranderingen ten opzichte van de door 'iedereen' verwachte perfect periodieke — en dus zeker niet chaotische — eb en vloed cyclus. Grofweg gesproken kan je zeggen dat ieder puntje in figuur 1 (of figuur 5) een 'meting' van de waterstand tijdens springvloed representeert. Dit betekent dat de computer het dynamische systeem, in dit geval een (gewone) differentiaalvergelijking, simuleert en, in de fysische tijd, eens in de 28-dagen een waarde van X en Y geeft; zo'n waarde bepaalt een puntje in figuur 1. Het volgende puntje volgt 28 dagen later. Het getal 28 staat natuurlijk voor de periode van de springvloed-cyclus. In figuur 1 staan meer dan 5000 van deze puntjes, dus, deze figuur beschrijft de eb en vloed dynamica in het bekken over een periode van zo'n 400 jaar. De waarden X en Y representeren de amplitude en de fase van (een getransformeerde en geschaalde versie van) het getij; de amplitude, of wel de geschaalde hoogte van de waterstand, volgt uit $\sqrt{X^2 + Y^2}$ (zie [6] voor meer details en voor de oorspronkelijke versie van figuur 1). De voor de hand liggende vraag 'Waarom zou het getij chaotisch zijn?' kan hier niet zo een, twee, drie worden beantwoord (en het antwoord valt ook volledig buiten de context van dit verhaal). In de artikelen [6] en [2] wordt uitgelegd hoe de nietlineariteit in het model die veroorzaakt wordt door de geometrie van het bekken (ofwel de vorm en structuur van de bodem) aanleiding kan geven tot chaotisch, en dus onregelmatig, gedrag.

De assen in figuur 1 en 5 geven aan dat $X, Y \in (-6, 6)$. Dit betekent dat de attractor A 'behoorlijk groot' is: grof gezegd wordt A zo'n beetje passend overdekt door het vierkant $S(12)$. Om de zojuist geconstateerde vertroebelende invloed van de constante C_A

zoveel mogelijk te reduceren schalen we A zodat we $S(12)$ kunnen vervangen door $S(1)$. Dit komt in figuur 5 neer op het negeren van de ± 6 -en langs de assen: we definiëren de punten links onderaan de twee (grote) vierkanten in figuur 5 als $(0, 0)$ en rechts bovenaan als $(1, 1)$.

In het eerste plaatje in figuur 5 is het vierkant overdekt met 144 kleinere vierkantjes. Dit betekent dat $\varepsilon = \frac{1}{12}$: $S(1)$ wordt overdekt door een rooster van 144 $S(\frac{1}{12})$'s. Enig telwerk levert dat 56 van deze $S(\frac{1}{12})$'s een of meer elementen van A bevatten: $\tilde{N}(\frac{1}{12}; A) = 56$. Dit geeft met behulp van (3) de eerste benadering van $\limcap(A)$: $\limcap(A) \approx (\log 56)/(\log 12) = 1.6199\dots$. Fijnere roosters geven: $\tilde{N}(\frac{1}{24}; A) = 158$ (het bijbehorende plaatje staat niet in deze tekst) en $\tilde{N}(\frac{1}{36}; A) = 289$ (het tweede plaatje in figuur 5). Dit geeft de volgende, meer nauwkeurige, benaderingen: $\limcap(A) \approx 1.5929\dots$ en $\limcap(A) \approx 1.5812\dots$. Deze drie opeenvolgende benaderingen komen dus uitstekend overeen. Zonder na te denken over de nauwkeurigheid van de benaderingen lijkt in ieder geval de conclusie dat $\limcap(A) \approx 1.6$ gerechtvaardigd.

Heeft het nu nog zin om een nog fijnere overdekking te beschouwen, en zo een nog betere benadering van $\limcap(A)$ proberen te vinden? Het antwoord hierop is 'nee'. Er zijn namelijk al een aantal vierkantjes $S(\frac{1}{36})$ aan te wijzen waarin geen van de ruim 5000 punten van A liggen, en die dus niet meetellen voor $\tilde{N}(\frac{1}{36}, A)$, maar waarvan het duidelijk is dat er wel punten van A in zouden liggen als we meer punten van A hadden gehad (zie bijvoorbeeld in figuur 5 de $S(\frac{1}{36})$'s in de 6-de rij, 29-ste kolom en de 8-ste rij, 31-ste kolom). Dus, terwijl je zou denken dat er zoveel punten zijn dat een fijner rooster zeker nog zou kunnen, blijkt voor deze waarde van ε al dat we 'last hebben' van het feit dat we zo 'weinig' punten hebben. Dit probleem wordt natuurlijk alleen maar groter als we ε nog kleiner kiezen. Dit betekent dat $\tilde{N}(\varepsilon; A)$ voor kleinere ε (veel) te klein gaat worden en we dus een steeds grotere fout in de benadering van $\limcap(A)$ gaan maken.

Opmerking 7. Voor een wiskundig preciesere definitie van de box-counting dimension dienen we $N(\varepsilon; A)$ in (1) te vervangen door $\tilde{N}(\varepsilon; A)$. Natuurlijk is het a priori weer niet duidelijk of de limiet 'altijd' dezelfde uitkomst zal geven. We gaan hier wederom niet verder op in en verwijzen naar [1, 7, 9] voor meer details.

Discussie

We hebben in dit verhaal laten zien hoe het intuïtieve begrip dimensie gemathematiseerd en gekwantiseerd, ofwel berekenbaar, kan worden. Met behulp van de formules (1), (2) en (3) kunnen we nu voor allerlei verzamelingen de dimensie bepalen. Soms

kan dat analytisch, soms is hiervoor een benaderingsprocedure en een computer nodig. Het is ook duidelijk dat er eigenlijk geen goede reden meer is om aan te nemen dat de dimensie van een 'object' /verzameling geheeltallig moet zijn.

Voor de laatste formule (3) geeft ons ook de mogelijkheid om de dimensie van een door de computer gegenereerde of zelfs van een experimenteel waargenomen 'data-set' te schatten. Met behulp van (3) en gedetailleerde kaarten kan nu zelfs een schatting worden gemaakt van een 'natuurlijk fractaal object' als de kustlijn van Engeland (zie [7]). Bij alle toepassingen van (3) dient men wel ten alle tijde bewust te zijn van de beperkte geldigheid van deze uitdrukking. We hebben gezien dat je zelfs bij een behoorlijk grote en nauwkeurige data-set als die van figuur 1 al vrij snel 'in de problemen' komt. Bij experimentele geografische of biologische 'toepassingen' van de box-counting dimension aanpak zijn deze problemen nog eens veel groter: er is meestal een relatief grote meetfout en het aantal betrouwbare 'waarnemingen' is vaak klein. In dit soort situaties is het dan ook maar de vraag wat de betrouwbaarheid of waarde van uitspraken over, bijvoorbeeld, 'de dimensie van de kustlijn van Engeland' is. Echter, dat neemt niet weg dat het met de hier gepresenteerde wiskundige ideeën mogelijk is om gestructureerd over deze kwesties na te denken.

Je zou verwachten dat de begrippen limietcapaciteit en box-counting dimension vooral binnen de theorie en de praktijk van het interdisciplinaire vakgebied dynamische systemen van groot belang zijn. Echter, om attractors in dynamische systemen goed te beschrijven en te begrijpen zijn subtielere en meer geavanceerde begrippen noodzakelijk. In die zin is het begrip limietcapaciteit slechts de eerste stap in de richting van de ontwikkeling van een 'dynamisch' dimensie begrip. Een van de belangrijkste redenen hiervoor is het feit dat de limietcapaciteit geen rekening kan houden met inhomogeniteiten in de verdeling van punten over een attractor A als in figuur 1. Het is zeer duidelijk dat sommige gebieden van A veel vaker 'bezoekt' worden dan anderen (zie figuur 5). De limietcapaciteit kan dit niet 'meten' en geeft daarom vaak een onbetrouwbaar beeld 'de dimensie' van een attractor. We verwijzen naar de inleidende teksten [1, 3, 9–10] waarin meer geschikte begrippen als correlatie-dimensie en Lyapunov-exponent worden geïntroduceerd. \Leftarrow

Verantwoording

Een belangrijk gedeelte van het 'materiaal' in dit verhaal is gebaseerd op Barteld Braaksma's hoofdstuk 'Gebroken dimensies en fractale verzamelingen' in [1], dat dieper op de materie ingaat. De figuren zijn gemaakt door Harmen van der Ploeg.

Referenties

- 1 H.W. Broer, F. Verhulst, *Dynamische Systemen en Chaos*, Epsilon Uitgaven **14** (1990).
- 2 A. Doelman, A.F. Koenderink, L.R.M. Maas, 'Quasi-periodically forced nonlinear Helmholtz oscillators', *Physica D* **164** (1-2), 1–27 (2002).
- 3 J. Guckenheimer, P. Holmes, *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*, Third edition, Applied Mathematical Sciences **42**, Springer (1990).
- 4 H. Lauwerier, *Fractals. Meetkundige Figuren in Eindeloze Herhaling*, Aramith (1987).
- 5 L.R.M. Maas, On the nonlinear Helmholtz response of almost-enclosed tidal basins with sloping bottoms, *J. Fluid Mech.* **349**, 361–380 (1997).
- 6 L.R.M. Maas, A. Doelman, 'Chaotic tides', *J. Phys. Oceanogr.* **32**(3), 870–890 (2002).
- 7 B. Mandelbrot, *The Fractal Geometry of Nature*, Freeman (1982).
- 8 W. Rudin, *Principles of Mathematical Analysis*, Third edition, McGraw-Hill (1976).
- 9 H.G. Schuster, *Deterministic Chaos. An Introduction*, Second edition, VCH Verlag (1988).
- 10 S. Wiggins, *Introduction to Applied Nonlinear Dynamical Systems and Chaos*, Texts in Applied Mathematics **2**, Springer (1990).