De zeef van Sierpinski

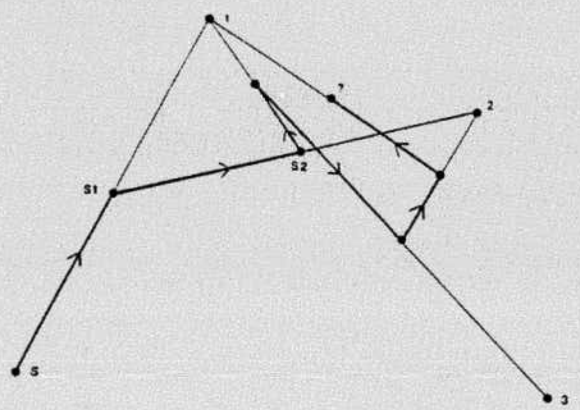
S.L. Kemme
Mathematatisch Instituut, RU Groningen

Samenvatting

Is wiskundig onderzoek slechts weggelegd voor hoog- en andere geleerden? Of kunnen gewone stervelingen, zoals leraren en leerlingen, daar ook nog wel eens mee in aanraking komen?
Is 'wiskunde in een context' altijd realistische toegepaste wiskunde? Heb je voor het bedrijven van wiskunde alléén maar je verstand, pen en papier nodig?
Lees onderstaand artikel voor een antwoord op deze vragen.

In het februari nummer van Pythagoras van 1988 stond de volgende opgave:

Waar gaat dat heen?



Zet op een velletje papier willekeurig drie punten (niet op één lijn) en geef ze bijvoorbeeld aan met 1, 2 en 3. Plaats (ook willekeurig) nog ergens een vierde punt S.
Maak drie lootjes, één met een 1, één met een 2 en één met een 3. Vouw de lootjes dicht, schud ze door elkaar en trek er één uit. Stel dat daar de 1 op staat. Trek dan een hulplijn van het startpunt S naar het punt 1 (zie figuur). Zet daarop precies in het midden een punt S_1 (en gum eventueel de hulplijn weer uit).
Doe de lootjes weer bij elkaar, schud ze en trek er opnieuw één uit. Als daar de 2 op staat, trek dan de hulplijn $S_1 2$ en zet daarop precies in het midden een punt S_2.
Ga zo door. Vanuit S_2 komt na loten een hulplijn naar één van de punten 1, 2 of 3, midden op die lijn een punt S_3, enzovoort. In de figuren zijn na de 1 en de 2 achtereenvolgens nog de cijfers 1, 3, 2 en 1 getrokken.
Hoe komt de verzameling van al die punten S eruit te zien, wanneer dit voorschrift maar vaak genoeg wordt herhaald?

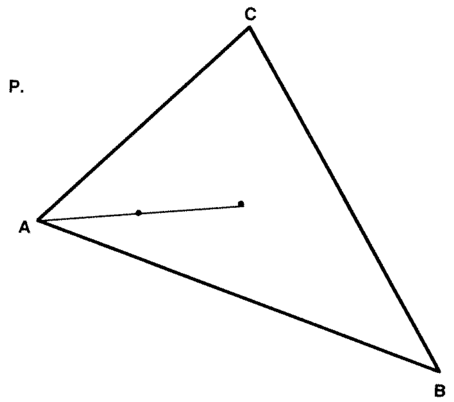
Het zal duidelijk zijn dat daar op de bovenbeschreven manier geen goed beeld van is te krijgen. Vooraf is nog wel eenvoudig in te zien dat wanneer een punt S_i eenmaal binnen de driehoek 123 is aangeland, alle daarop volgende punten binnen de driehoek blijven. Maar daar is op het eerste gezicht alles wel mee gezegd. Om snel een goed beeld te krijgen van die verzameling moet het voorschrift door een computer worden uitgevoerd. Het (mogelijk verrassende) resultaat verschijnt dan stapje voor stapje op het beeldscherm, of kan netjes worden geplot. Hoe vaker het voorschrift wordt herhaald, hoe beter het resultaat. Maar uiteindelijk zal het resultaat afhangen van het aantal pixels van het beeldscherm.

Fraaie (plotter)resultaten ziet de redactie met spanning tegemoet. Ze zullen samen met het gebruikte programma zeker worden afgedrukt. Vergeet bij inzending niet te vermelden met welk type computer er is gewerkt.
Zoals al is opgemerkt, zal het resultaat mogelijk verrassend zijn. Het kan ook met een ander tekenvoorschrift worden verkregen. Daarover in een volgend nummer meer.

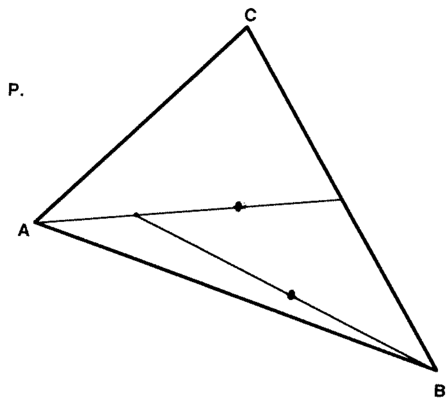
Tenslotte nog een lang niet eenvoudige vraag. Wie kan *zonder* bruut computergeweld aangeven hoe het resultaat eruit moet komen te zien?

Meestal gaan dergelijke opgaven bij mij het ene oog in en het andere weer uit. Maar nu zat ik eraan vast voordat ik er erg in had.

Aan de situatie kun je zien dat het wel iets met het zwaartepunt te maken heeft. Misschien is het wel het zwaartepunt? Maar dan moet een punt dat toevallig in dat zwaartepunt terecht is gekomen, op zijn plaats blijven liggen. En dat is niet het geval. Want dan ligt het volgende punt op een zwaartelijn halverwege een hoekpunt.

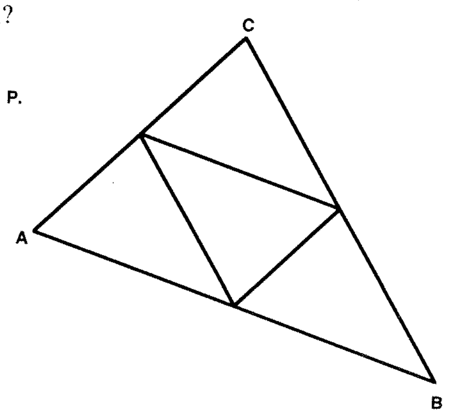


Misschien zijn de zwaartelijnen samen de limietfiguur. Even kijken. Wat gebeurt er als een punt op één van de zwaartelijnen ligt?



Dan gaat hij van de zwaartelijn af als B of C wordt gekozen. Dus de zwaartelijn is geen invariante lijn en daarmee afgevoerd als kandidaat voor de limietfiguur.

De driehoek gevormd door de middenparallellen misschien?



Nee ook niet, want het midden van AC gaat op een kwart van A afzitten als A wordt gekozen en verlaat daarmee deze 'middendriehoek'.

Ontdekking 1 (door raden):

Tot nu toe ben ik alleen nog maar dom aan het raden geweest. Mijn enige ontdekking is dat ik figuren in de driehoek moet zien op te sporen, die hun punten behouden bij de toepassing van een volgende stap.

Het wordt tijd voor wat meer systematiek!
In één of ander coördinatenstelsel van het vlak geef ik P de coördinaten (p_1, p_2) en A, B en C overeenkomstig: (a_1, a_2), (b_1, b_2) en (c_1, c_2). De coördinaten van het eerste punt P_1 zijn dan:

$$\tfrac{1}{2}((p_1, p_2) + (x, y))$$

waarbij (x,y) één van de punten A, B of C is, bijvoorbeeld:

$$\tfrac{1}{2}((p_1, p_2) + (a_1, a_2))$$

Als daarna C wordt gekozen dan ziet het volgende punt P_2 er zó uit:

$$\tfrac{1}{2}(\tfrac{1}{2}((p_1, p_2) + (a_1, a_2)) + (c_1, c_2))$$

Of, uitgewerkt:

$$\tfrac{1}{4}(p_1, p_2) + \tfrac{1}{4}(a_1, a_2) + \tfrac{1}{2}(c_1, c_2)$$

In een meer symbolische vorm wordt het allemaal wat overzichtelijker:

$$P_1 = \tfrac{1}{2}P + \tfrac{1}{2}A$$
$$P_2 = \tfrac{1}{4}P + \tfrac{1}{4}A + \tfrac{1}{2}C$$

Wordt daarna bijvoorbeeld weer A gekozen dan geldt voor P_3:

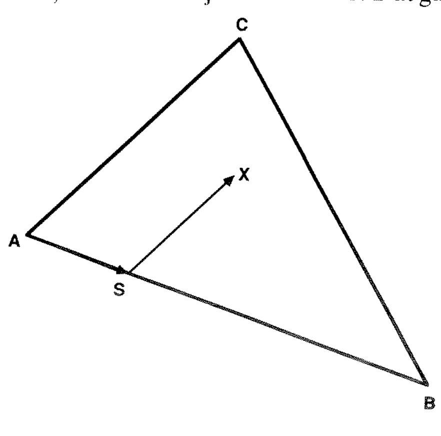
$$P_3 = \tfrac{1}{8}P + \tfrac{3}{4}A + \tfrac{1}{4}C$$

Ontdekking 2 (na enig gereken):

Voor P ontstaan steeds toenemende machten van $\tfrac{1}{2}$. Daarmee wordt de invloed van P uiteindelijk tot 0 gereduceerd. Het doet er in de limietfiguur dus niet toe waar je bent begonnen.
Een algemene gedaante ziet er nu als volgt uit:

$$P_n = (\tfrac{1}{2})^n P + aA + bB + cC$$

waarbij de getallen a, b en c sommen zijn van machten van $\tfrac{1}{2}$ waar helaas niet veel regelmaat in valt te bespeuren. In de limietfiguur blijven deze laatste termen over.
De vorm: $aA + bB + cC$ doet me denken aan de mogelijkheid om een coördinatensysteem binnen de driehoek in te voeren. Een punt binnen de driehoek kun je dan weergeven met behulp van de positieve getallen a, b en c waarbij $a + b + c = 1$. Dat gaat zo:



De plaats van X kan worden aangegeven met de vector AX ten opzichte van het punt A:

AX = AS + SX.

Nu is AS een 'gedeelte' van AB, bijvoorbeeld: AS = b·AB, en SX is een gedeelte van AC, bv: SX = c·AC.
Hierin zijn b en c getallen tussen 0 en 1. De punten A, B, C en X zijn op te vatten als de eindpunten van vectoren ten opzichte van één of andere oorsprong O. Dit levert:

OX = OA + b·AB + c·AC

Hierin kunnen we AB vervangen door: OB − OA en AC door: OC − OA. Schrijven we de vector weer gewoon als het eindpunt, dan komt er:

$$\begin{aligned} X &= A + b(B-A) + c(C-A) = \\ &= (1-b-c)A + bB + cC \end{aligned}$$

waarbij b en c het gedeelte van de vectoren B − A en C − A aangeven. Daarbij geldt: $0 \leqslant b \leqslant 1$, $0 \leqslant c \leqslant 1$ en $b+c \leqslant 1$. $1-b-c$ vervangen we door de letter a. Zo krijgen we:

$$X = aA + bB + cC$$

met: $0 \leqslant a,b,c \leqslant 1$ en $a+b+c=1$. Op deze manier kunnen we ieder punt binnen de driehoek ondubbelzinnig aangeven met de drie getallen a, b en c.

Vermoeden 1 (door oude kennis van barycentrische coördinaten):

Ik begin te vermoeden dat de limietfiguur er wel eens veel ingewikkelder uit kan gaan zien dan ik eerst dacht. Het gaat erom welke getallen a, b en c zijn te produceren met die machten van een $\frac{1}{2}$. Misschien wordt de hele figuur wel opgevuld doordat iedere keuze van a, b en c uiteindelijk wordt geproduceerd. Op dit punt aangekomen schiet mijn getaltheoretische kennis tekort en besluit ik de computer te hulp te roepen.

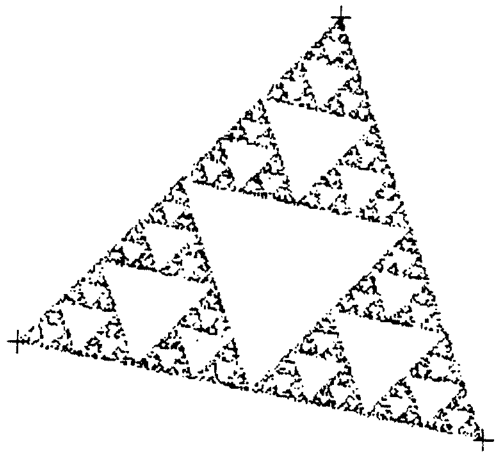
In een verloren uur schrijf ik het volgende programma in GW BASIC. Het grootste stuk van het programma bestaat uit het kiezen van de beginpunt en de driehoek.

```
REM Pythagoras Opgave

DIM X(3),Y(3)
CLS: SCREEN 2
WINDOW (−10,−10) − (10,10)
PRINT "Kies het beginpunt."
PRINT "(Kies de coördinaten tussen 10 en −10.)"
INPUT P,Q: PSET (P,Q)
PRINT "Kies de driehoek."
PRINT "(Weer tussen −10 en 10.)"
FOR 1 = 1 TO 3
  INPUT X(1),Y(1)
  IF 1 = 1 PSET(X(1),Y(1)) ELSE LINE −
  (X(1),Y(1))
NEXT 1
RANDOMIZE TIMER
FOR 1 = 1 TO 5000
  K = INT(3×RND) + 1
  P = 5×(P + X(K)): Q = 5×(Q + Y(K)))
  PSET (P,Q)
NEXT 1
END
```

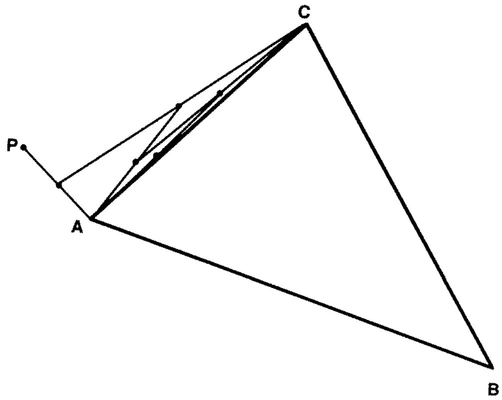
Ontdekking 3 (met de computer):

Het resultaat is verbluffend (dit plaatje is met de Macintosh gemaakt):



Er komt een fractal tevoorschijn. In het boekje van Lauwerier staat hij ook: het is de zeef van Sierpinski. De gaten in de driehoek tonen aan dat de getallen a, b en c niet iedere waarde tussen 0 en 1 kunnen krijgen.
Nu de verklaring nog:

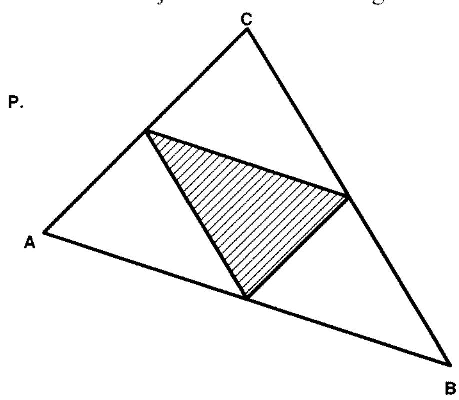
We zien aan de figuur dat alle punten door de driehoek worden 'vastgehouden', dat wil zeggen: als een punt eenmaal in de driekhoek zit, dan komt die er niet meer uit. Meetkundig is dat snel in te zien. Ligt P_n in de driehoek, dan ligt de verbindingslijn van P_n met A, B of C helemaal in de driehoek en daarmee ook het nieuwe punt P_{n+1}. Maar waarom komen bijna alle punten uiteindelijk in de driehoek terecht? Als je wilt kun je de punten er allemaal buiten houden. Ligt het beginpunt P bijvoorbeeld links van de lijn door A en C en kies je beurtelings A of C, dan blijven alle punten links van de lijn door A en C liggen. Ze wandelen naar het midden van AC toe, maar komen er alleen in het oneindige terecht.



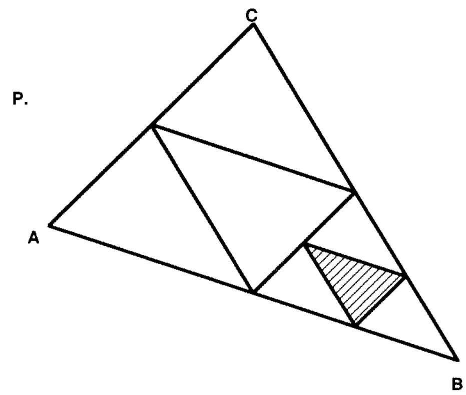
De kans dat dit echter in een randomkeuze gebeurt is na n stappen $(\frac{2}{3})^n$ en gaat dus snel naar 0. Ook door alleen maar A te kiezen kunnen we de punten buiten de driehoeken houden. De kans dat dit gebeurt gaat nog sneller naar 0. Aangezien dit de enige manieren zijn om de punten buiten de driehoek te houden, is de kans dat dit gebeurt dus te verwaarlozen.

Ontdekking 4 (na redeneren vanuit de kansrekening en de meetkunde):

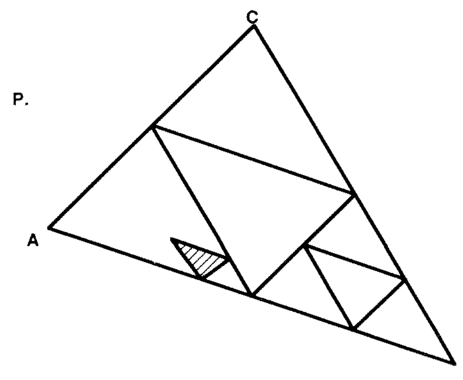
De punten worden na een voldoend aantal stappen door de driehoek ingevangen.
We zien dat de middendriehoek, op één punt na, leeg is gebleven. Een punt dat hierin terecht is gekomen kan er dus kennelijk niet meer in terugkeren.



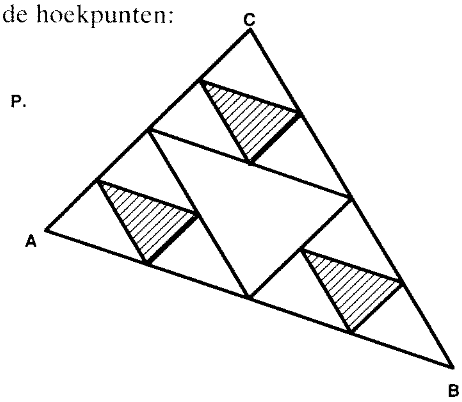
Stel dat een punt in de midden-driehoek ligt. Wordt B gekozen dan komt dit neer op een puntvermenigvuldiging vanuit B met een factor $\frac{1}{2}$. De toekomstige verzameling voor het nieuwe punt ziet er dan zó uit:



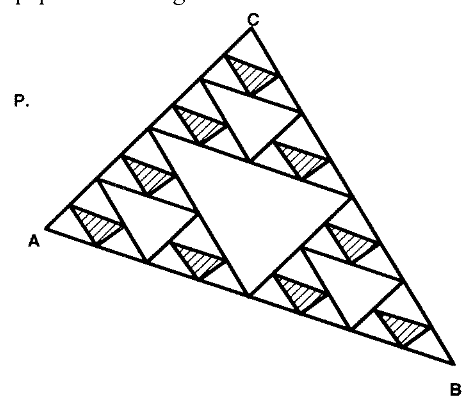
Wordt vervolgens A gekozen, dan wordt de mogelijkheid voor het weer volgende punt ook bepaald als een puntvermenigvuldiging vanuit A:



In de getekende situatie is eerst B en dan A gekozen. We moeten echter alle mogelijkheden bekijken. Maar iedere keuze van een hoekpunt betekent in feite een puntvermenigvuldiging van het hele vlak vanuit dat hoekpunt met een factor $\frac{1}{2}$.
Na één stap zijn de mogelijkheden voor punten die in de middendriehoek lagen de drie kleinere driehoeken bij de hoekpunten:



Bij de volgende stap zullen deze drie driehoeken zich weer opsplitsen in negen kleinere:

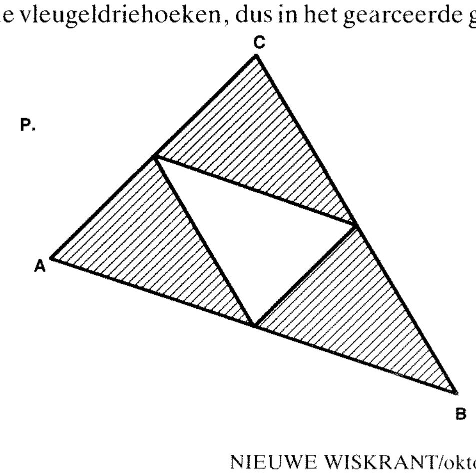


Enzovoorts. Belangrijk hierbij is de waarneming dat de punten niet meer in de middendriehoek terug zullen komen. Ze worden 'verspreid' over de drie hoeken bij de hoekpunten.

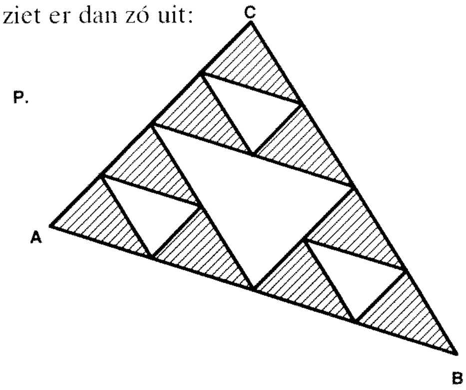
Ontdekking 5 (na enig tekenwerk en wat meetkunde):

Als een punt in een middendriehoek is terechtgekomen, dan zal die daar niet meer in terugkeren.

Maar misschien kan een punt vanuit één van drie 'vleugel'-driehoeken wel weer in de middendriehoek terechtkomen. Wie goed naar de vorige plaatjes heeft gekeken weet dat dat niet zo is. Voor de volledigheid gaan we toch nog even na wat er met die vleugeldriehoeken kan gebeuren. Stel een punt zit in één van de drie vleugeldriehoeken, dus in het gearceerde gebied:



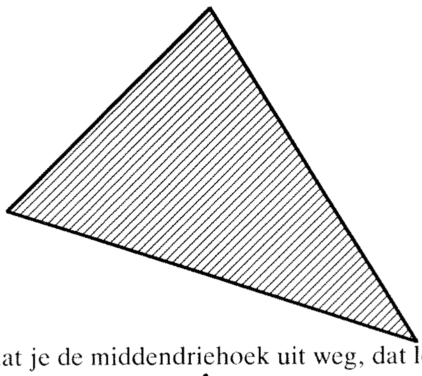
Door puntvermenigvuldiging vanuit A bijvoorbeeld, wordt deze hele figuur twee keer zo klein gekopieerd in de driehoek bij A. Hetzelfde gebeurt als vanuit B of C wordt gewerkt. Het resultaat van alle mogelijke punten die vanuit het gearceerde gebied zijn vertrokken, ziet er dan zó uit:



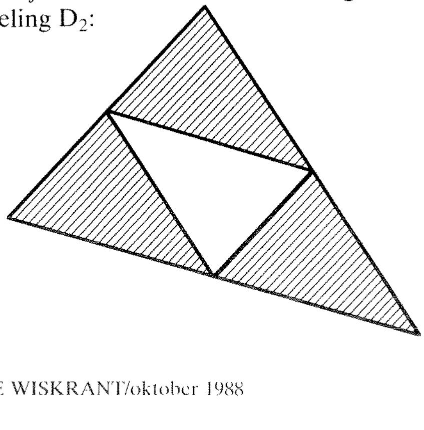
Het gearceerde gebied is kleiner geworden door het wegvallen van de kleinere middendriehoeken. In ieder geval zijn de punten in hun oorspronkelijke gearceerde gebied gebleven. En dit zelfde geldt vervolgens ook weer voor de nu gearceerde kleinere vleugeldriehoeken bij een volgende toepassing van de transformatie.

Ontdekking 6 (na enig geteken en wat meetkunde):

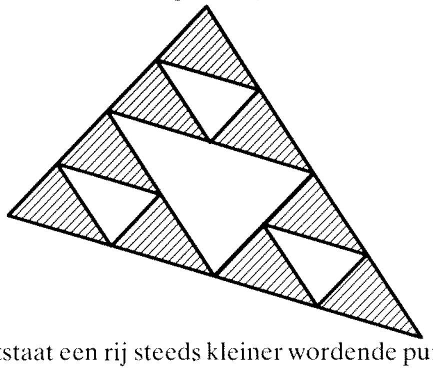
Een punt dat een keer in een 'vleugel'-driehoek is beland blijft in een vleugel-driehoek (niet noodzakelijkerwijs de oorspronkelijke).
Hiermee is voldoende verklaring gevonden voor het computerplaatje: de middendriehoeken kunnen hooguit één punt bevatten en blijven dus wit. De punten gaan zich ophopen in de vleugeldriehoeken.
Het computerplaatje is natuurlijk maar een eindige verzameling punten (50000) waarin ik één of ander patroon heb ontdekt. Wat is nu precies die zeef van Sierpinski?
Je kunt er op twee manieren tegen aan kijken. De eerste is dat je begint met de 'volle' driehoek D_1:



Daar laat je de middendriehoek uit weg, dat levert de verzameling D_2:



Uit de gearceerde driehoeken laat je ook weer de middendriehoeken weg om D_3 te maken:



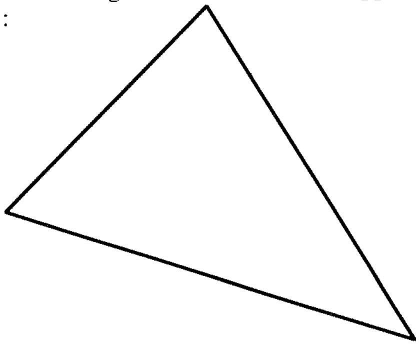
Zo ontstaat een rij steeds kleiner wordende puntverzamelingen:

$$D_1 \supset D_2 \supset D_3 \supset \ldots..$$

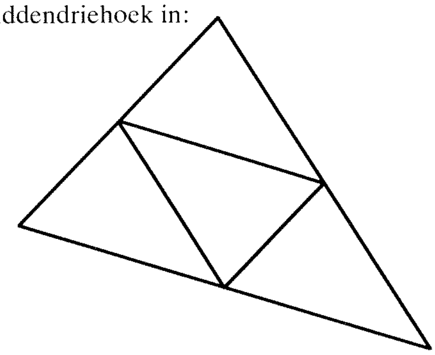
De limietfiguur hiervan wordt aangegeven met:

$$\bigcap_{i=1}^{\infty} D_i = Z$$

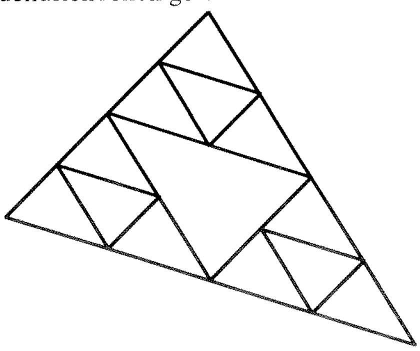
Dit is de wiskundige definitie van de '*zeef van Sierpinski*'. De naam zeef is heel beeldend: de figuur bevat steeds meer, steeds kleiner wordende gaten.
Een andere manier om naar de figuur te kijken is meer opbouwend van aard.
Begin met de lege driehoek, dat wil zeggen: alleen de zijden:



Dit noem ik S_1, van skelet. Vervolgens teken ik daar de middendriehoek in:



Dat is S_2. In de vleugeldriehoeken worden ook weer de middendriehoeken getekend.



Dat is natuurlijk S_3. Zo ontstaat een oneindige rij van steeds in grootte toenemende skeletten. De limietfiguur geven we aan met:

$$\bigcup_{i=1}^{\infty} S_i = S$$

Wie mocht denken dat S=Z heeft het mis. Z is aanzienlijk groter dan S, hoewel dat niet zo direct is in te zien.

De zeef van Sierpinski heeft een aantal opvallende eigenschappen. Ik noem er twee:
- Z heeft oppervlakte 0 terwijl S, als deelverzameling van Z, lengte oneindig heeft.
- Z is totaal samenhangend, dat wil zeggen: er bestaat een continue weg van ieder punt van Z naar ieder willekeurig ander punt van Z die volledig in Z ligt.

De moraal

Het bovenstaande vind ik een mooi voorbeeld van een wiskundig onderzoek (op mijn eigen niveau natuurlijk), waarin diverse gebieden van de wiskunde op een heel vanzelfsprekende manier aan de orde komen. Het mooiste dat ik tot nu toe ken.
Terugkijkend herken ik drie belangrijke fasen:
- De verkenning, waarin het probleem wordt opgepakt, vermoedens ontstaan, de onderzoeksmethode zich ontwikkelt en een gevoel van onmacht ontstaat omdat de oplossing niet in zicht komt.
- De computer als redder in de nood, die het antwoord op het probleem weet te leveren met bruut geweld, maar waarbij het onverwachte van dat antwoord tegelijk allerlei vragen naar de verklaring oproept.
- De bewijsvoering waarin, gesteund door het computerresultaat, een verklaring van het antwoord wordt ontwikkeld.

In de overgangen van deze fasen zit een harmonische lijn: van motiverende instap, via onverwachte oplossing naar afrondende verklaring.

Inmiddels heb ik dit probleem aan een tiental mensen uit mijn directe omgeving voorgelegd. Van wiskundige leken tot beroeps-wiskundigen. Ze reageren natuurlijk allemaal op hun eigen manier. Maar opvallende gemeenschappelijke reactie is het onverwachte van de oplossing en de fascinatie voor de limietfiguur. De gesprekken verspringen onopgemerkt van beschouwingen over de concrete meetkundige figuur, naar theoretische argumenten hoe die figuur er in het oneindige uit zal zien.

'Wiskunde in context' is in de mode. Iedereen denkt daarbij aan wiskunde in realistische situaties die meer of minder aansluiten bij de praktijk van alledag. Het bovenstaande sluit goddank niet aan bij de praktijk van alledag. Toch durf ik dit een briljant voorbeeld van wiskunde in context te noemen.

Als dank aan het tijdschrift Pythagoras heb ik een nieuwe opgave bedacht:
- Bewijs dat iedere S_i volledig te doorlopen is zonder het potlood van papier te halen en zodanig dat ieder lijnstuk maar één keer wordt afgelegd, en:
- Schrijf een computer-programma dat S_i ook zo kan tekenen.

Met dank aan: Maarten Bos, Henk Broer en Floris Takens.