John Simons
Rijksuniversiteit Groningen
j.l.simons@rug.nl

Onderzoek

Het $3x+1$-probleem, recente ontwikkelingen en oplossingsrichtingen

Voor een natuurlijk getal x definiëren we de opvolgerfunctie $T : x \to \frac{x}{2}$ als x is even en $x \to \frac{3x+1}{2}$ als x is oneven. Het (onbewezen) $3x+1$-vermoeden zegt dat vroeger of later de cykel $1 \rightleftharpoons 2$ wordt bereikt. In dit artikel bespreekt John Simons enkele recente ontwikkelingen en oplossingsrichtingen.

Het $3x+1$-probleem werd in 1937 voor het eerst geformuleerd door Lothar Collatz. Er bestaat veel onderzoek over [4,6,9], desondanks is het probleem nog steeds onopgelost. Voor een beginwaarde x wordt de rij $x \to T(x) \to T^2(x) \cdots$ de baan van x genoemd. De cykel $1 \rightleftharpoons 2$ wordt de triviale cykel genoemd. Er kunnen nu twee $3x+1$-vermoedens worden geformuleerd:

Het cykelvermoeden. *Voor iedere x bestaat er een k zodanig dat $T^k(x) = 1$, dat wil zeggen er bestaan geen niet-triviale cykels.*

Het convergentievermoeden. *Voor iedere x bestaat er een bovengrens $U(x)$ zodanig dat voor iedere k geldt $T^k(x) \leq U(x)$, dat wil zeggen er bestaan geen divergente banen.*

Naar aanleiding van een vakantiecursus schreef Benne de Weger in 2014 in het NAW een interessant artikel over herformuleringen van het $3x+1$-probleem in de grafentheorie, in de theorie van eigenwaarden van oneindige matrices en in de theorie van functionaalvergelijkingen [16]. Het huidige artikel behandelt recente oplossingsrichtingen en deelresultaten voor de twee onderscheidenlijke vermoedens:

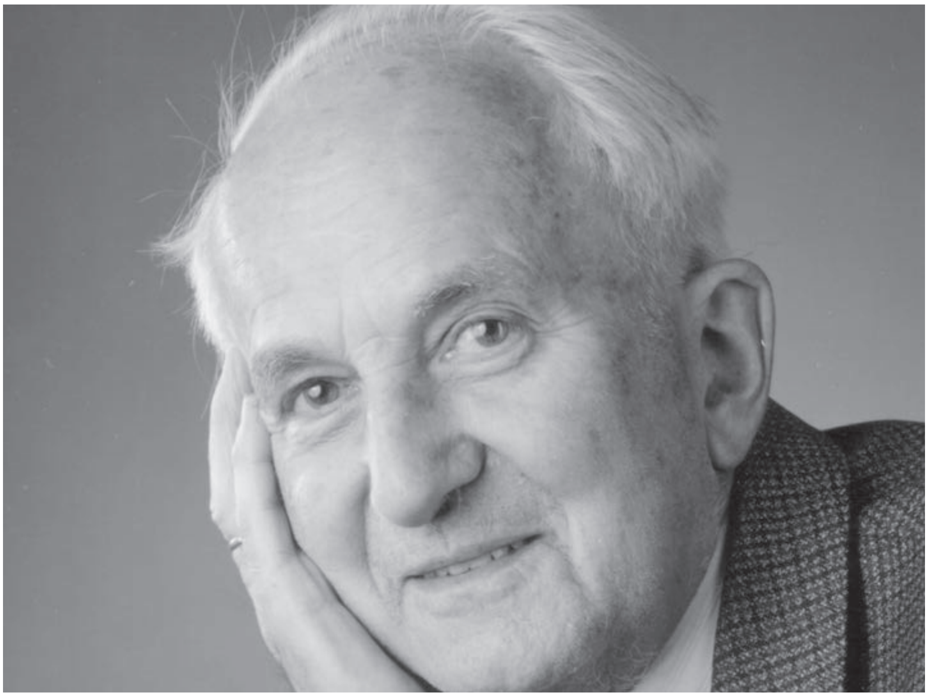
- Voor het *cykelvermoeden* worden voorwaarden gedefinieerd voor het bestaan van een niet-triviale cykel. Als zo'n cykel K oneven en L even getallen bevat, dan moet $\frac{3^K}{2^{K+L}}$ ongeveer 1 zijn. Dit betekent dat $(K+L)\log 2 - K\log 3$ klein is. De transcendente getaltheorie levert stellingen over de grootte van zulke lineaire log-expressies in termen van de coëfficiënten. De opvolgerfunctie T wordt hier als een deterministische functie beschouwd.
- Voor het *convergentievermoeden* wordt voor een beginwaarde $x > 2$ het concept stoptijd geanalyseerd, dit is de kleinste k waarvoor $T^k(x) < x$. De theorie van ergodische processen levert uitspraken over de dichtheid van de verzameling beginwaarden met een eindige stoptijd. De opvolgerfunctie T wordt hier als een stochastische (niet-deterministische) functie beschouwd.

In de volgende paragraaf bespreken we de geschiedenis van het $3x+1$-probleem. Daarna worden enkele generalisaties van het $3x+1$-probleem genoemd. Deze gebruiken we om recente ontwikkelingen en oplossingsrichtingen voor het $3x+1$-probleem te beoordelen op hun algemenere toepasbaarheid.

In de paragraaf daarna gaan we dieper in op een oplosmethode [10] om het (niet) bestaan van cykels te bewijzen. Als een cykel m lokale minima (en dus ook m lokale maxima) heeft, noemen we dat een m-*cykel*. Met behulp van transcendente getaltheorie en diofantische approximatietechnieken is het (wonderlijk schijnende) resultaat bewezen dat voor $2 \leq m \leq 75$ geen m-cykel bestaat. Die aanpak is ook toepasbaar op generalisaties van het $3x+1$-probleem, zij het met de nodige aanpassingen.

Vervolgens bespreken we het recente bewijs van Tao [14] dat bijna alle banen bijna begrensd zijn. Verder gaan we in op de generaliseerbaarheid van de stoptijd. Voor een speciale generalisatie van het $3x+1$-probleem wordt het bestaan van divergente banen bewezen.

Hoewel de onderzoekslijnen voor de twee vermoedens geheel verschillend zijn, gaan we in de laatste paragraaf in op een merkwaardige overlap, namelijk het be-



Lothar Collatz (1910–1990)

staan van bepaalde cykels en het niet bestaan van bepaalde divergente banen voor het $3x+1$-probleem.

De literatuur over het $3x+1$-probleem is omvangrijk en groeit nog steeds. Dit overzicht is derhalve onvolledig.

Een hardnekkig probleem

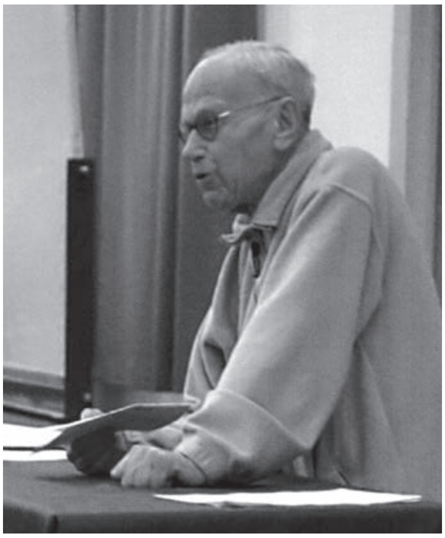
Algemeen wordt aangenomen dat Lothar Collatz (Duits wiskundige) in de dertiger jaren dit probleem heeft geïntroduceerd.

In 1972 publiceert Martin Gardner erover in de *Scientific American* en gaat het probleem 'viral'. Het probleem wordt meestal het Collatz-probleem genoemd, maar staat ook onder andere namen bekend zoals probleem van Hasse, Kakutani, Ulam, Syracuse. Herman Duparc introduceerde het probleem bij zijn wiskundestudenten (waaronder ik) tijdens een Christiaan Huygens-borrel in café Het Noorden in Delft in 1967. Sindsdien houdt het $3x+1$-probleem mij bezig.

Waarom is het $3x+1$-probleem zo taai? Dit is eerder een wetenschapsfilosofische dan een wiskundige vraag. Volgens Lagarias [6] zijn er twee oorzaken: (i) Het pseudo-random gedrag van een deterministisch proces. Een klassiek getaltheoretisch bewijs dat er geen divergente banen zijn loopt vast op het pseudo-random gedrag en een stochastische analyse loopt vast op de deterministische opvolgerfunctie die het proces niet pseudo-random maakt. (ii) De onbeslisbaarheid van het probleem. Conway [2] heeft bewezen dat er geen algoritme bestaat om voor iedere $x \in \mathbb{N}$ te bepalen of er een k bestaat met $T^k(x) = 1$.

Hierbij kunnen kanttekeningen worden geplaatst. Het concept pseudo-random heeft een intuïtieve connotatie. De wiskundige definitie is niet van toepassing op sommige generalisaties van het $3x+1$-probleem. Zie de paragraaf 'Het convergentievermoeden' verderop. Een algoritme kan geen bewijs uit het ongerijmde produceren.

Het Koninklijk Wiskundig Genootschap voert de zinspreuk dat een onvermoeide arbeid alles te boven komt. Volgelingen van Brouwer zouden kunnen vragen: "Wanneer?" Erdős beoordeelde het $3x+1$-probleem als volstrekt hopeloos. Hetzelfde zal lange tijd gezegd zijn van de grote stelling van Fermat (geen positief gehele oplossingen van $x^n+y^n=z^n$ voor $n>2$). Dat pessimisme werd in 1997 gelogenstraft door Wiles. Zijn onvermoeide arbeid resulteerde in een bewijs van 190 pagina's!



Herman Duparc (1918–2002)

In navolging van Hilbert worden de bewijzen en deelresultaten in dit artikel gespiegeld aan generalisaties van het $3x+1$-probleem:

> "If we do not succeed in solving a mathematical problem, the reason frequently consists in our failure to recognize the more general standpoint from which the problem before us appears only as a single link in a chain of related problems." – David Hilbert (1862–1943)

Uitzoomen: het $px+q$-probleem

We beschouwen een algemener probleem. Voor oneven p, q kiezen we de opvolgerfunctie

$$T(x) = \begin{cases} \dfrac{x}{2} & \text{als } x \text{ is even,} \\ \dfrac{px+q}{2} & \text{als } x \text{ is oneven.} \end{cases} \qquad (1)$$

Dit probleem is uitgebreid geanalyseerd [7, 11]. Het cykelvermoeden (ook voor het $px+q$-probleem onbewezen) is dat het aantal cykels eindig is. Noem het aantal cykels $C(p,q)$. $C(3,1)=1$ is het cykelvermoeden voor het $3x+1$-probleem. Er is bewezen dat voor iedere p er oneindig veel $px+q$-problemen zijn met $C(p,q) \geq 1$ [12]. Rob Tijdeman bestudeerde priemgetallen q waarvoor k, l bestaan zodanig dat $2^{k+l}-3^k=q$. Noem q een gegeneraliseerde Tijdeman-priem indien er k, l bestaan zodanig dat $2^{k+l}-p^k=q$. Dan is $C(p,q) \geq 1$. Een exotisch voorbeeld is het $97x+32641759$-probleem met tenminste 92 cykels.

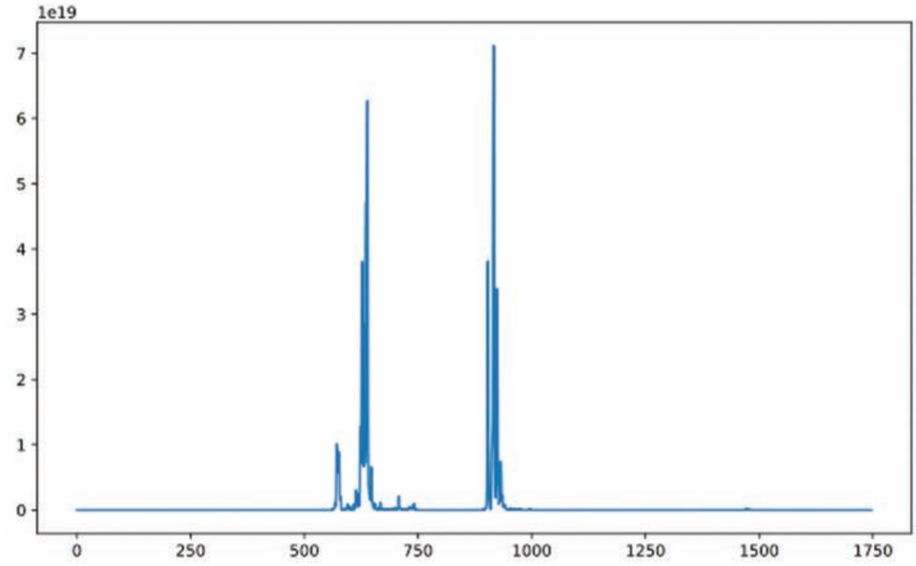
Het convergentievermoeden (onbewezen) is dat voor $p=3$, $U(x)$ altijd bestaat en voor $p \geq 5$, $U(x)$ bijna nooit bestaat.

Het gegeneraliseerde Collatz-probleem

Een algemenere versie van het $px+q$-probleem werd geïntroduceerd door Matthews [8]. Beschouw daartoe de opvolgerfunctie

$$x_{n+1} = \frac{p_i \cdot x_n + (q-p_i)r_i}{q} \quad \text{als } x_n \equiv i \pmod{q} \qquad (2)$$

met p_i, r_i $(i = 0, \ldots, q-1)$ en q zodanig dat $(p_i, p_j) = 1$, $(p_i, q) = 1$, $r_i \equiv i \pmod{q}$.



Figuur 1 Getallen in de cykel van opvolgerfunctie (3) beginnend met 114.

Zij $s = |p_0 p_1 \cdots p_{q-1}| - q^q$. Vanwege de eisen aan p_i, q, r_i geldt altijd $s \neq 0$. Het cykelvermoeden en het convergentievermoeden (beide onbewezen) zijn afhankelijk van het teken van s. Als $s < 0$ dan bestaat $U(x)$ voor iedere x. Er zijn dan geen divergente banen en er is een eindig aantal cykels (minimaal 1). Als $s > 0$ dan bestaat $U(x)$ bijna nooit, dat wil zeggen bijna alle banen zijn divergent en het aantal cykels is eindig.

Er zijn veel voorbeelden om deze vermoedens empirisch te ondersteunen. Als $\prod p_i = q^q - 1$ zijn er grote cykels. Beschouw bijvoorbeeld de opvolgerfunctie

$$T(x) = \begin{cases} \frac{x}{4} & \text{als } x \equiv 0 \pmod 4, \\ \frac{3x-3}{4} & \text{als } x \equiv 1 \pmod 4, \\ \frac{5x-2}{4} & \text{als } x \equiv 2 \pmod 4, \\ \frac{17x-3}{4} & \text{als } x \equiv 3 \pmod 4. \end{cases} \tag{3}$$

Er zijn tenminste 17 cykels, sommige met lengte 1747 zoals $\{114, 142, 177, 132, 33, \ldots, 27\}$ en een maximum waarde van $71\,133\,256\,567\,324\,334\,017$, zie Figuur 1.

Matthews formuleert nog een extra vermoeden gebaseerd op een waarschijnlijkheidsargument. De getallen in een divergente baan zijn uniform verdeeld modulo q^α voor iedere $\alpha \geq 1$.

Het cykelvermoeden

We keren nu weer terug naar het $3x+1$-probleem. Hoe kun je het bestaan of niet bestaan van cykels bewijzen? Dat gaat via het vinden van noodzakelijke voorwaarden voor het bestaan van een cykel. Veronderstel dat er een cykel bestaat met k oneven en l even getallen. Als we ons beperken tot cykels die bestaan uit een rij oneven getallen gevolgd door een rij even getallen (een 1-cykel) dan kan zo'n bestaansvoorwaarde eenvoudig worden afgeleid. Het minimum van de cykel is een oneven getal x, dat kan worden geschreven als $x = a2^k - 1$. Na k keer toepassen van $T(x) = \frac{3x+1}{2}$ ontstaat het even getal $a3^k - 1$, het maximum van de cykel. Na l keer toepassen van $T(x) = \frac{x}{2}$ zijn we weer terug bij het getal $a2^k - 1$. Als zo'n cykel bestaat dan voldoen a, k, l aan de koppelvergelijking

$$\frac{a3^k - 1}{2^l} = a2^k - 1.$$

Hieruit volgt

$$0 < 2^{k+l} - 3^k \leq 2^l - 1,$$

$$0 < 1 - \frac{3k}{2^{k+l}} < \frac{1}{2^k}$$

en

$$1 < \frac{2^{k+l}}{3^k} < 1 + \frac{1}{2^{k-1}}.$$

Een noodzakelijke voorwaarde voor zo'n 1-cykel is dat k, l voldoen aan

$$0 < (k+l)\log 2 - k \log 3 < \frac{1}{2^k - 1}. \tag{4}$$

Voor dit type van diofantische vergelijkingen levert de transcendente getaltheorie, aanvankelijk ontwikkeld door Baker [1], een expliciete ondergrens die polynomiaal is in k. Als $|(k+l)\log 2 - k\log 3| \neq 0$ dan is $|(k+l)\log 2 - k \log 3| > ck^{-d}$ met $c, d > 0$. Confrontatie van deze polynomiale ondergrens met de exponentiële bovengrens van vergelijking (4) levert een $k_{\max}$ op. Met de theorie van de kettingbreuken (*continued fractions*) kun je vervolgens slim zoeken naar oplossingen $k \leq k_{\max}$. In 1978 bewees Steiner [13] op deze wijze dat $k = 1$, $l = 1$ en dus $a = 1$ de enige oplossing is van vergelijking (4), overeenkomend met de cykel $1 \rightleftharpoons 2$.

De aanpak van Simons en De Weger

We definiëren een m-cykel als een cykel met m lokale maxima (en dus ook m lokale minima). De baan tussen twee lokale minima begint met een stijgende rij oneven getallen, gevolgd door een dalende rij even getallen. De lokale minima kun je schrijven als $x_i = a_i 2^{k_i} - 1$ voor $i = 0, \ldots, m-1$. Na k_i keer toepassen van de opvolgerfunctie ontstaat het lokale maximum $y_i = a_i 3^{k_i} - 1$ en na l_i keer toepassen van de opvolgerfunctie ontstaat het minimum $x_{i+1} = a_{i+1} 2^{k_{i+1}} - 1$. Steiner bewees dat $1 \rightleftharpoons 2$ de enig mogelijke 1-cykel is. In 2005–2010 bewezen Simons en De Weger [10] dat er geen m-cykels bestaan voor $2 \leq m \leq 75$. Het bewijs steunt op de volgende constatering: een noodzakelijke en voldoende voorwaarde voor het bestaan van een m-cykel met $K = \sum k_i$ oneven getallen en $L = \sum l_i$ even getallen is het bestaan van een oplossing (a_i, k_i, l_i) voor $i = 0, \ldots, m-1$ van de matrixvergelijking

$$M \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_{m-1} \end{pmatrix} = \begin{pmatrix} 2^{l_0} - 1 \\ 2^{l_1} - 1 \\ \vdots \\ 2^{l_{m-1}} - 1 \end{pmatrix},$$

met

$$M = \begin{pmatrix} -3^{k_0} & 2^{k_1 + l_0} & 0 & \cdots & 0 \\ 0 & -3^{k_1} & 2^{k_2 + l_1} & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & \ddots & -3^{k_{m-2}} & 2^{k_{m-1} + l_{m-2}} \\ 2^{k_0 + l_{m-1}} & 0 & \cdots & 0 & -3^{k_{m-1}} \end{pmatrix}.$$

Dit is een generalisatie van de koppelvergelijking in Steiners bewijs. Met een niet-triviale approximatie van $\frac{x_{i+1}}{x_i}$ waarbij de minima groter zijn dan een numeriek verkregen ondergrens X_0, kan een noodzakelijke voorwaarde voor het bestaan van een m-cykel worden afgeleid.

Lemma 1. *Veronderstel dat de matrixvergelijking een integer oplossing (a_i, k_i, l_i) heeft. Dan geldt voor K, L, m,*

$$0 < (K+L)\log 2 - K \log 3 < \sum_{i=0}^{m-1} \frac{1}{x_i} < mc_m 2^{-\frac{\delta - 1}{\delta^m - 1}K},$$

waarbij $\delta = \log 3 / \log 2$ en c_m een constante is die marginaal van X_0 afhangt.

Met transcendente getaltheorie (sinds 1978 zijn scherpere grenzen afgeleid, maar de essentie blijft [3]) kan wederom een $K_{\max}$ worden berekend. Bovendien kan vanuit de ondergrens $x_i > X_0$ een $K_{\min}$ worden berekend. Vervolgens worden geavanceerde diofantische approximatie-technieken gebruikt om het interval $(K_{\min}, K_{\max})$ te verkleinen en voor $m \leq 75$ wordt aldus het niet bestaan van m-cykels bewezen.

Recente ontwikkelingen

De bovengrens in Lemma 1 kan op twee manieren worden verlaagd.

(i) Het bewijs voor m-cykels maakt gebruik van de schatting $x_i > X_0$ voor elke i. Dat kan scherper. De worst case in het bewijs is dat $x_0 < x_1 < \cdots < x_{m-1}$ en het verschil tussen x_{i+1} en x_i kan worden afgeschat. We introduceren het concept m-nykel (nykel van nagenoeg cykel).

Definitie. Een m-*nykel* is een baan vanaf x_0 met $m+1$ lokale minima $x_0, x_1, \ldots, x_m$ die voldoen aan $x_0 < x_i$. In tegenstelling tot een m-cykel geldt niet $x_m = x_0$.

Voor een 1-nykel (een baan vanaf het globale minimum x_0 naar het lokale minimum x_1) kan een resultaat worden bewezen dat vergelijkbaar is met Steiners resultaat.

Lemma 2. *Voor een* 1-*nykel met* $x_1 = x_0 + 2v$ *met* $v \geq 1$ *bestaat een eenvoudig berekenbare bovengrens* $x_0 \leq C_v$, *met andere woorden bij gegeven* v *is er een eindig aantal oplossingen voor* x_0.

Bewijs. a, k, l zijn gedefinieerd zoals voor een 1-cykel, zie vergelijking (4). De koppelvergelijking is nu

$$\frac{a3^k - 1}{2^l} = a2^k - 1 + 2v$$

en die leidt tot de noodzakelijke voorwaarde

$$0 < k \log 3 - (k+l) \log 2 < \frac{v}{2^{k-1}}. \quad (5)$$

Transcendente getaltheorie leidt tot een $k_{\max}$ en dus tot een C_v. □

Analoog aan Steiners bewijs voor 1-cykels kan worden aangetoond dat de enige 1-nykel met $v = 1$ is $11, 17, 26, 13$. Voor $v = 4052$ echter zijn er twee 1-nykels met als kleinste element $64\,827$ en $593\,919$. Bij gegeven v geldt $x_0 \leq C_v$, dus als $x_0 > C_v$ (waarbij v, zoals $v = 4052$, met de theorie der kettingbreuken slim moet worden gekozen!) dan is $x_1 > x_0 + 2v$. Voor de worst case van oplopende minima geldt $x_i > x_0 + i \cdot 2v$. Het effect is dat de factor m in Lemma 1 wordt verlaagd tot $\sim \log m$.

(ii) In het bewijs van Simons en De Weger wordt voor de coëfficiënten a_i in de matrixvergelijking de afschatting $a_i \geq 1$ gebruikt. Dat kan eveneens scherper. Inductief kan worden bewezen dat $a_i \neq a_j$. We nemen aan dat met behulp van Lemma 1 het niet bestaan van een m-cykel is bewezen voor gegeven m. Veronderstel nu dat er een $m+1$-cykel bestaat met $a_0 = a_q$ en $x_0 < x_q$. Voor de variabelen $a_0, \ldots, a_{q-1}$ volgt uit de matrixvergelijking de voorwaarde

$$0 < (k_1 + \cdots + k_q + l_0 + \cdots + l_{q-1}) \log 2 - (k_0 + \cdots + k_{q-1}) \log 3 < \sum_{i=0}^{q-1} \frac{1}{x_i}.$$

Met de substitutie

$$K = k_0 + \cdots + k_{q-1},$$
$$L = l_0 + \cdots + l_{q-1} + k_q - k_0,$$

ontstaat een tegenspraak met vergelijking (1) voor een kleinere m. Deze voorwaarde leidt tot

$$\prod_{i=0}^{m-1} a_i \sim \left(\frac{m}{e}\right)^m.$$

Het effect op c_m is voor kleine m verwaarloosbaar en voor $m > 50$ een factor ~ 3.

Deze scherpere afschattingen verlagen de bovengrens in Lemma 1 en leiden tot uitsluiting van m-cykels voor grotere m. Dit onderzoek loopt nog.

Het convergentievermoeden

Het convergentievermoeden zegt dat voor iedere x een bovengrens $U(x)$ bestaat zodanig dat voor iedere k, $T^k(x) \leq U(x)$. Empirisch convergeren alle getallen, maar $U(x)$ kan relatief groot zijn. Zo is

$$U(27) = 4616$$

en

$$U(274\,133\,054\,632\,352\,106\,267\,113\,298) = 124\,744\,388\,651\,798\,242\,538\,014\,293\,435\,290\,632.$$

Er is bewezen dat het aantal getallen $1 \leq n \leq n_{\max}$ waarvan de baan naar $1 \rightleftharpoons 2$ convergeert tenminste $n_{\max}^{84}$ is voor voldoende grote $n_{\max}$ [5]. Hoe kun je bewijzen dat voor het $3x+1$-probleem een baan $x, T(x), T^2(x), \ldots$ convergeert? Het concept pseudo-random levert een intuïtief 'bewijs':

'Bewijs'. Kies random een oneven getal x en bereken $T(x)$. In 50% van de gevallen zal $T(x)$ even zijn, en in 50% van de gevallen zal $T(x)$ oneven zijn. De verwachte waarde van de vermenigvuldigingsfactor van x is $\frac{3}{2} \cdot \frac{1}{2} = \frac{3}{4} < 1$. De verwachting is daarom dat de baan $x, T(x), T^2(x), \ldots$ convergeert.

Intuïtief moge dit overtuigend zijn, het is geen wiskundig bewijs. Dat komt onder andere omdat random trekken van een oneven positief getal een kansfunctie verlangt op de verzameling positieve oneven getallen. Die kan niet uniform zijn omdat $\sum_{n=0}^{\infty} \mathbb{P}(x = 2n+1) = 1$. Wat dan wel? En wat is het argument voor een gekozen kansfunctie?

In het convergentie-onderzoek wordt gebruikgemaakt van de begrippen stoptijd en dichtheid. Voor $x > 2$ is de kleinste k met $T^k(x) < x$ de stoptijd en de kleinste k met $T^k(x) = 1$ de totale stoptijd. De dichtheid van een verzameling (in de verzameling der natuurlijke getallen $\mathbb{N}$) wordt als volgt gedefinieerd. Beschouw een deelverzameling $A \subseteq \mathbb{N}$ en definieer $a(n)$ als het aantal elementen in A kleiner of gelijk aan n. Als $\lim_{n\to\infty} \frac{a_n}{n}$ bestaat dan heet die limiet de *natuurlijke* (ook wel *asymptotische*) *dichtheid* van A. De dichtheid van de even getallen is $0{,}5$, van de priemgetallen 0 en van de samengestelde getallen 1. Het zijn alle drie oneindige verzamelingen. Een bewering wordt bijna altijd waar genoemd voor $\mathbb{N}_+$ indien de bewering waar is voor een deelverzameling met dichtheid 1.

Terras [15] bewees in 1976 dat de natuurlijke dichtheid $F(k)$ van de getallenverzameling met een stoptijd $\leq k$ eindig is. Als $k \to \infty$ dan $F(k) \to 1$. Dit betekent dat bijna alle getallen een eindige stoptijd hebben. De verzameling getallen die geen eindige stoptijd hebben is niet per definitie leeg. Als $\pi^*(x)$ het aantal getallen $n < x$ is met een eindige stoptijd, dan bestaat voor het aantal getallen zonder eindige stoptijd een bovengrens $cx^{1-\eta}$ met $c, \eta > 0$.

Voor het $3x+1$-probleem is een eindige stoptijd een noodzakelijke voorwaarde voor een eindige totale stoptijd. Bij generalisaties is de relatie tussen de aldus gedefinieerde stoptijd en totale stoptijd niet bepalend voor het eindigen in een cykel. Neem het $3x+7$-probleem met de cykel

$5,11,20,10$. Het getal 11 heeft stoptijd 2 en heeft geen eindige totale stoptijd. Om het al of niet eindigen in een cykel te analyseren is een generalisatie van het concept stoptijd nodig:

Definitie. *Stoptijd voor $px+q$-problemen.* Beschouw de baan $x \to T(x) \to T^2(x) \cdots$. Voeg als voorganger van x toe $2x$, dat wil zeggen $T(2x)=x$. Beschouw twee lokale minima $T^j(x)$ en $T^k(x)$ met $0 \le j < k$. Definieer de *stoptijd van x* als de minimale k waarvoor een j bestaat zodanig dat $T^k(x) < T^j(x)$. Definieer de *totale stoptijd* als de minimale k waarvoor een j bestaat zodanig dat $T^k(x) = T^j(x)$.

De nieuwe stoptijd is nu weer een noodzakelijke voorwaarde voor de nieuwe totale stoptijd. Met deze definities wordt het bestaan van niet-triviale cykels ondervangen.

Recente ontwikkelingen

Er zijn alternatieve definities voor de dichtheid, bijvoorbeeld de logaritmische dichtheid. Dan wordt niet het aantal elementen met n vergeleken, maar $\sum_{n\in A, n<x} \frac{1}{n}$ met $\log x$. Tao bewees in 2020 (zijn woorden) dat: *Almost all orbits of the Collatz map attain almost bounded values*. Wat bewijst Tao [14] nu precies?

Tao gebruikt een eigen definitie voor de logaritmische dichtheid. Voor $R, A \subset \mathbb{N}_+$ wordt een random variabele $\mathrm{Log}(R)$ gedefinieerd door

$$\mathbb{P}(\mathrm{Log}(R) \in A) = \frac{\sum_{x\in A\cap R} \frac{1}{x}}{\sum_{x\in R} \frac{1}{x}}.$$

Zijn logaritmische dichtheid van de verzameling A wordt dan

$$\lim_{x\to\infty} \mathbb{P}(\mathrm{Log}(\mathbb{N}_+ \cap [1,x] \in A)).$$

Tao hanteert de gebruikelijke definitie van *almost all*. Een bewering $P(x)$ is waar voor bijna alle x indien de (logaritmische) dichtheid de limiet 1 heeft als $x \to \infty$. Zij $\mathrm{Lo}(x) = \min_k T^k(x)$. Tao bewijst

Lemma 3. *Voor iedere $f : \mathbb{N}_+ \to \mathbb{R}$ met $\lim_{x\to\infty} f(x) = \infty$ geldt voor bijna alle x, $\mathrm{Lo}(x) \le f(x)$.*

Neem als voorbeeld $f(x) = \log x$. Veronderstel dat we van alle getallen $\le X_0$ weten dat de baan convergeert naar de cykel $1 \rightleftharpoons 2$. Dan zegt dit lemma dat voor $x = X_0 + 1$ almost always de baan van

Terence Tao

x een minimum bevat dat kleiner is dan $\log x$ en dus ook kleiner dan x. Inductief is dan bewezen dat de baan van x eveneens convergeert naar de cykel $1 \rightleftharpoons 2$. Dat is de betekenis van *almost bounded values* in de titel van zijn artikel. Tao claimt niet het cykelvermoeden of het convergentievermoeden te hebben bewezen. Hij bewijst wel een sterke uitspraak over een combinatie van de twee vermoedens. Het bewijs omvat 49 pagina's, gebruikt transcendente getaltheorie en ergodische theorie en is geen sinecure (ik kwam er niet doorheen).

Los van Tao's resultaat kan in bijzondere gevallen voor een generalisatie van het $3x+1$-probleem divergentie worden bewezen. Lothar Collatz bestudeerde aanvankelijk getallen, voortgebracht door

$$T(x) = \begin{cases} \dfrac{2x}{3} & \text{als } x \equiv 0 \pmod 3, \\ \dfrac{4x-1}{3} & \text{als } x \equiv 1 \pmod 3, \\ \dfrac{4x+1}{3} & \text{als } x \equiv 2 \pmod 3. \end{cases} \tag{6}$$

In de terminologie van Matthews is $s = 32-27 = 5$ met de daarbij horende vermoedens over cykels en convergentie. De verwachte vermenigvuldigingsfactor is $32/27$. T is een permutatie van de verzameling $\mathbb{N}$ en de inverse T^{-1} is

$$T^{-1}(x) = \begin{cases} \dfrac{3x}{2} & \text{als } x \equiv 0 \pmod 2, \\ \dfrac{3x+1}{4} & \text{als } x \equiv 1 \pmod 4, \\ \dfrac{3x-1}{4} & \text{als } x \equiv 3 \pmod 4. \end{cases} \tag{7}$$

$s = 81-64 = 17$ en de hierbij verwachte vermenigvuldigingsfactor is niet $27/32$, maar $81/64 > 1$. Volgens Matthews gelden voor T en T^{-1} dus dezelfde vermoedens. Divergente banen zijn dubbel divergent: zij komen van oneindig, dalen af naar een minimum en stijgen weer naar oneindig. Dezelfde vermenigvuldigingsfactor (bijvoorbeeld $32/27$) zorgt kennelijk voor de oneindig lange daling en de oneindig lange stijging. Empirisch lijken er willekeurig veel dubbel divergente banen te zijn.

De permutaties van vergelijkingen (6) en (7) kunnen alternatief worden beschreven als de permutatie

$$\begin{aligned} 2n &\leftrightarrow 3n, \\ 4n+1 &\leftrightarrow 3n+1, \\ 4n+3 &\leftrightarrow 3n+2. \end{aligned} \tag{8}$$

Elk natuurlijk getal x valt zowel links als rechts in precies één categorie (regel). In bijzondere gevallen kan het bestaan van divergente banen worden bewezen. Beschouw het voorbeeld

$$\left.\begin{array}{c} 6n \\ 12n+1 \\ 12n+2 \\ 12n+3 \\ 12n+4 \\ 12n+5 \\ 12n+7 \\ 12n+8 \\ 12n+9 \\ 12n+10 \\ 12n+11 \end{array}\right\} \leftrightarrow \left\{\begin{array}{c} 3n \\ 15n+1 \\ 15n+2 \\ 15n+4 \\ 15n+5 \\ 15n+7 \\ 15n+8 \\ 15n+10 \\ 15n+11 \\ 15n+13 \\ 15n+14 \end{array}\right. . \tag{9}$$

Vanaf de tweede regel geldt (van links naar rechts): Als $x > 2$ dan $T(x) > x$ en $T(x) \not\equiv 0 \pmod 3$. Hieruit volgt dat er drie triviale cykels $(0), (1), (2)$ zijn en een oneindig aantal dubbeldivergente banen met minima $3, 9, 15, \ldots$. In een divergente baan is geen uniforme distributie van getallen zoals vermoed door Matthews.

In vergelijking (6) zijn de verzamelingen $\{(2,0),(4,1),(4,2)\}$ en $\{(3,0),(3,1),(3,2)\}$ van gelijke grootte. Hetzelfde geldt voor de verzamelingen links en rechts in vergelijking (9). Het zijn zogenaamde *CC-sets*:

Definitie. Zij C een verzameling van $k \ge 2$ geordende paren $\{(a_i,b_i), i = 1,\ldots,k\}$ met $a_i \in \mathbb{N}_+$ and $b_i \in \mathbb{N}_0$. We noemen C een *CC-set* (complete coverage set) indien voor iedere $x \in \mathbb{N}$ er precies een i bestaat waarvoor $x \equiv b_i \pmod{a_i}$. Voorbeelden zijn:

$$\{(2,0),(4,1),(8,3),(16,7),(16,15)\},$$
$$\{(3,0),(3,1),(3\cdot 2^{k+1}, 3\cdot 2^k - 1) \mid k \ge 0\}.$$

Gelijke grootte CC-sets zijn een voldoende voorwaarde voor de aanpak van Simons en De Weger. Op deze wijze kan het (niet) bestaan van cykels worden bewezen, evenwel niet het bestaan van divergente banen.

Permutaties zijn in twee opzichten bijzondere generalisaties van de $3x+1$-functie T.

1. Als er een eindig aantal cykels is, dan moeten er divergente banen zijn. Er kunnen namelijk geen aanloopsprieten zijn zoals in het $3x+1$-probleem.
2. Na verwisseling van twee expressies aan de rechter (of linker) zijde (bijvoorbeeld $15n+7$ en $15n+8$ in vergelijking (9) blijft de functie een permutatie. We hebben dus te maken met een klasse van $(N+1)!$ permutaties waarvoor dezelfde CC-set-voorwaarde en lineaire logexpressie geldt, maar waarvan de cykels per gekozen permutatie in die klasse verschillen.

Hebben de twee vermoedens overlap?

De twee onderzoekslijnen om het cykelvermoeden en het convergentievermoeden voor het $3x+1$-probleem te bewijzen gebruiken verschillende takken van de wiskunde, respectievelijk transcendente getaltheorie en ergodische theorie. Zij hebben weinig overlap. Hierna wordt uitgelegd dat het nykel-concept in beide onderzoekslijnen een rol speelt.

Het verschil tussen x_0 en x_1 in een 1-nykel is $2v$. Voor iedere v bestaat er een 1-nykel met $k=2$, $l=1$. Die begint met $x_0=16v-5$ en eindigt met $18v-5$. De hele 1-nykel is $(16v-5, 24v-7, 36v-10, 18v-5)$ en dan valt de overeenkomst met de 1-cykel $(-5,-7,-10,-5)$ voor $x<0$ op. Een analoge opmerking betreft de 2-cykel $(-17,-25,-37,-55,-82,-41,-61,-91,-136,-68,-34,-17)$ die correspondeert met de 2-nykel die begint met $x_0=4096w-17=2^{11}2w-17$ en eindigt met $4374w-17=3^7 2w-17=x_0+2\times 139$. Algemeen geldt: indien er voor $x<0$ een m-cykel bestaat met $k_i, l_i, i=0,\dots,m-1$, dan bestaan er voor $x>0$ oneindig veel m-nykels met diezelfde $k_i, l_i, i=0,\dots,m-1$.

De eindwaarde $18v-5$ van de 1-nykel kan in de vorm $16w-5$ worden geschreven als $v=8t$ en $w=9t$. w bevat dan een factor 8 minder dan v. Dit proces is eindig. Hieruit volgt dat er geen divergente baan bestaat die een opeenvolging is van 1-nykels met $k_i=2$, $l_i=1$. Hetzelfde geldt voor een opeenvolging van 2-nykels met $k_0=4$, $l_0=2$, $k_1=3$, $l_1=3$. De triviale constatering dat er voor het $3x+1$-probleem geen divergente baan bestaat met uitsluitend oneven getallen, hangt samen met het bestaan van de cykel (-1).

Een divergente baan kan niet bestaan uit m-nykels met dezelfde k_i, l_i vanwege zo'n koppelvergelijking. Een divergente baan bestaat (afgezien van een aanloopspriet) per definitie uit een opeenvolging van m-nykels waarvan de beginwaarden $\bar{x}_j$ voldoen aan $\bar{x}_j<\bar{x}_{j+1}$. De rij k_i, l_i (per m-nykel zijn er m minima) kan geen cyclisch patroon bezitten. Lemma 2 kan worden gegeneraliseerd voor m-nykels en dat heeft gevolgen voor de structuur van een divergente baan

Lemma 4. *Een divergente baan bestaat uit een qua grootte van v en/of m oneindige verzameling m-nykels.*

Bewijs. Veronderstel dat de verzamelingen van v waarden en van m waarden allebei eindig zijn. Voor iedere m en v bestaat er een maximale x_0. Over alle m-nykels en alle v waarden bestaat er dan een maximale x_0 en dat is in tegenspraak met de divergentie. □

Conclusies

1. De kracht van een methode om een vermoeden van het $3x+1$-probleem te bewijzen, kan worden getoetst aan de toepasbaarheid op gegeneraliseerde problemen. Bij het cykelvermoeden gaat $\Lambda=(K+L)\log 2-K\log 3$ van het $3x+1$-probleem over in $\Lambda=(K+L)\log 2-K\log p$ voor het $px+q$-probleem en over in een lineaire vorm van meer logaritmen voor andere generalisaties. Transcendente getaltheorie blijft toepasbaar voor het bewijs van (non)-existentie van cykels. Bij het convergentievermoeden speelt s in de formulering van Matthews een cruciale rol. Het concept pseudo-random is niet altijd van toepassing op divergente banen. Terras' definitie van de stoptijd is specifiek voor het $3x+1$-probleem.
2. Voor het cykelvermoeden kan de bovengrens in Lemma 1 worden verscherpt. Dat leidt tot een grotere m voor de non-existentie van m-cykels. Een nieuw idee, bijvoorbeeld een m-onafhankelijk bewijs voor een eindig aantal cykels, is nodig om significante voortgang te bereiken.
3. Voor het convergentievermoeden kan (vermoedelijk) de dichtheid van de getallenverzameling waarvoor het convergentievermoeden geldt, worden verhoogd. Een doorbraak zou zijn als bewezen kan worden dat divergente banen een bepaalde structuur moeten bezitten, en vervolgens het bestaan van zulke banen tot een tegenspraak leidt.
4. En overigens is het $3x+1$-probleem nog niet opgelost.

Referenties

1 A. Baker, Linear forms in the logarithms of algebraic numbers IV, *Mathematika* 15 (1968), 204–216.

2 J.H. Conway, Unpredictable iterations, in *Proc. Number Theory Conference, University of Colorado, Boulder, Colorado*, 1972, pp. 49–52.

3 J.H. Evertse, Linear forms in logarithms I : Complex and p-adic, 2007, https://www.math.leidenuniv.nl/~evertse/linearforms.pdf.

4 R.K. Guy, *Unsolved Problems in Number Theory*, Problem Books in Mathematics, Vol. 1, Springer, 194, second edition.

5 I. Krasikov and J.C. Lagarias, Bounds for the $3x+1$ problem using difference inequalities, *Acta Arithmetica* 109(3) (2003), 237–258.

6 J.C. Lagarias, ed., *The Ultimate Challenge: The $3x+1$ Problem*, AMS, 2010.

7 J.C. Lagarias, The $3x+1$ problem: An overview (2021), arXiv:2111.02635.

8 K.R. Matthews, The generalized $3x+1$ mappings, Markov chains and ergodic theory, 2010, http://www.numbertheory.org/PDFS/matthews-final-revised.pdf.

9 A. van der Poorten e.a., *Recurring Sequences*, AMS Mathematical Surveys and Monographs, Volume 104, AMS, 2000.

10 J.L. Simons en B.M.M. de Weger, Theoretical and computational bounds for m-cycles of the $3n+1$ problem, *Acta Arithmetica* 117(1) (2005), 51–70.

11 J.L. Simons, On the non-existence of m-cycles for generalized Syracuse sequences, *Acta Arithmetica* 131(3) (2007), 217-254.

12 J.L. Simons, Exotic Collatz cycles, *Acta Arithmetica* 134(3) (2008), 201–209.

13 R.P. Steiner, A theorem on the Syracuse Problem, in *Proc. 7th Manitoba Conference on Numerical Mathematics 1977, Winnipeg*, 1978, pp. 553–559.

14 T. Tao, Almost all orbit of the Collatz map attain almost bounded values, 2020, arXiv: 1909.03562 .

15 R. Terras, A stopping time problem on the positive integers, *Acta Arithmetica* 30 (1972), 241–252.

16 B. de Weger, Het $3n+1$-vermoeden, *Nieuw Archief voor Wiskunde* 5/15(3) (2014), 40–50.