Experimentele Getaltheorie

Frits Beukers

1 De vergelijking van Mordell

In de 17e eeuw merkte Fermat al op dat $5^2 = 3^3 - 2$ en vroeg zich af of er meer gehele getallen x, y bestaan zó dat $y^2 = x^3 - 2$. Al spoedig liet Fermat zien dat er geen andere oplossingen bestaan. Fermat's vraag laat zich gemakkelijk generaliseren. Gegeven een getal $k \neq 0$, positief of negatief, bestaan er gehele getallen x, y zó dat

$$y^2 = x^3 + k.$$

Dit is een voorbeeld van een *diophantische vergelijking*, dat wil zeggen een vergelijking waarvan we de oplossing in gehele getallen willen weten. De vergelijking $y^2 = x^3 + k$ staat bekend als de *vergelijking van Mordell*, naar de eerste wiskundige die aantoonde dat voor gegeven $k \neq 0$ er hooguit eindig veel oplossingen kunnen zijn. Het bewijs hiervan, dat niet makkelijk is, is te vinden in Mordell, *Diophantine Equations*, Academic Press 1969. Mordell's vergelijking wordt gezien als een van de meest "klassieke" diophantische vergelijkingen. Hoewel we weten dat er voor gegeven k hoogstens eindig veel oplossingen zijn, is het vaak een lastige zaak om die oplossingen ook inderdaad te bepalen. In het boek van London en Finkelnstein, *On Mordell's equation*, Bowling Green State University 1973, wordt voor de meeste waarden van k met $|k| \leq 100$ de volledige oplossing gegeven. Het feit dat hieraan een heel boek besteed wordt, geeft al aan dat het geen eenvoudige klus is! Hier is een voorbeeld, de vergelijking $y^2 = x^3 + 17$ heeft de oplossingen

$$(x, y) = (-2, 3), (-1, 4), (2, 5), (4, 9), (8, 23)$$

$$(43, 282), (52, 375), (5234, 378661)$$

waarbij we alleen de oplossingen met $y > 0$ hebben opgeschreven. De andere krijg je gewoon door het teken van y te veranderen. Aan de andere kant zijn er voor veel waarden van k helemaal geen oplossingen. Sinds een twintigtal jaren zijn er systematische methoden ontwikkeld om Mordell's vergelijking aan te pakken. Deze ontwikkeling is gestimuleerd door de beschikbaarheid van steeds grotere hoeveelheden rekencapaciteit in de vorm van steeds sneller wordende computers. In een artikel uit 1998 van Gebel, Pethö en Zimmer,

On Mordell's equation, Compositio Math. 110(1998), p335-367 worden alle vergelijkingen met $|k| < 10.000$ opgelost en het merendeel van de vergelijkingen met $|k| < 100.000$. Deze laatste lijst is afgemaakt door K.Wildanger. De lijst met $|k| < 10.000$ is te vinden op `http://emmy.math.uni-sb.de/ simath/MORDELL/`. Uit deze tabel reproduceren we in de Appendix A de oplossingen voor alle k met $0 < |k| \leq 100$.
Sleutelwoorden in de gebruikte technieken zijn: lineaire vormen in elliptische logaritmen en LLL-algoritme, ingredienten die pas de laatste vijftien jaar gevonden zijn.

2 Hall's vermoeden

Ondanks het feit dat er nu uitgebreide tabellen voor de Mordell vergelijking bestaan, blijft er nog een groot aantal vragen over. Bijvoorbeeld, om de Mordellvergelijking op te lossen zou het handig zijn om te weten hoe klein het verschil tussen een kwadraat en een derde macht kan zijn. Het bekendste vermoeden in die richting is het volgende.

Vermoeden 2.1 (Hall) *Er bestaat een getal $C > 0$ zó dat voor elk tweetal $x, y \in \mathbb{Z}_{>0}$ met $x^3 \neq y^2$ geldt $|x^3 - y^2| > C\sqrt{x}$.*

Dit vermoeden werd naar voren gebracht door Marshall Hall tijdens een symposium *Computers in Number Theory* [H] in 1969. Dit symposium was een van de eerste waarin computer-experimenten op getaltheoriegebied voorop staan.
Laten we eens aannemen dat Hall's vermoeden waar is en kijken wat het gevolg is voor de Mordell vergelijking. Aangezien we nog geen enkel voorbeeld van x, y kennen met $0 < |x^3 - y^2| < 0.01\sqrt{x}$, nemen we even $C = 0.01$. Voor een oplossing van de vergelijking $y^2 = x^3 + k$ zou dit betekenen dat $|k| = |x^3 - y^2| > 0.01\sqrt{x}$ en dus, $x < (100k)^2$. Om de vergelijking van Mordell op te lossen met bijvoorbeeld $k = 17$ zouden we voor $x = -2, -1, 0, 1, 2, \ldots 1700^2$ moeten proberen of $x^3 + 17$ een kwadraat is. Voor de meeste PC's is dit tegenwoordig een paar seconden werk.
Het vermoeden van Hall is, zoals de naam zegt, een vermoeden. De ondergrenzen voor $|x^3 - y^2|$ die werkelijk bewezen zijn in de getaltheorie liggen daar nog heel ver af. Een recente ondergrens is de volgende

Stelling 2.2 (Sprindzuk, 1982) *Als* $x^3 \neq y^2$ *en* $x > 10$, *dan geldt*

$$|x^3 - y^2| > \gamma \log x/(\log\log x)^6$$

waarin γ *een berekenbaar positief getal is.*

De techniek om deze stelling te bewijzen, is A.Baker's theorie van lineaire vormen in logaritmen uit 1970 en verbeteringen daarop in latere jaren. Dit resultaat is het beste resultaat dat we kunnen krijgen met de hedendaagse technieken uit de getaltheorie. Om echter in de buurt van Hall's ongelijkheid te komen zullen er waarschijnlijk fundamenteel nieuwe technieken in de getaltheorie ontwikkeld moeten worden.
Omdat we natuurlijk geen zin hebben om op die nieuwe technieken te wachten, gaan we aan het experimenteren. Allereerst zouden we eens kunnen proberen om getallen x, y te vinden zo dat $0 < |x^3 - y^2| < \sqrt{x}$. De lezer is bij deze uitgenodigd een poging hiertoe te wagen, zonder naar de tabellen verderop te kijken. Hij of zij zal waarschijnlijk verbaasd zijn om te moeten constateren dat dit vrijwel niet lukt. De hardnekkigheid waarmee $|x^3 - y^2| > \sqrt{x}$ voor elk doorsnee tweetal x, y is opvallend. Het lijkt erop dat we te maken hebben met een harde wetmatigheid die zich in de getallenwereld afspeelt. Een systematische manier om erachter te komen hoe hardnekkig dit verschijnsel is, is natuurlijk voor $x = 1, 2, 3, \ldots$ de waarde x^3 uitrekenen en het verschil met het dichtstbijzijnde kwadraat y^2 bepalen. Dit experiment is talloze malen op vele computers gedaan. Een aantal jaren geleden heb ik zelf ook een dergelijke proef genomen en alle getallen $x < 10^{11}$ getest. De waarde $|x^3 - y^2|/\sqrt{x}$ zullen we de *Hallwaarde* van het paar x, y (of van x) noemen. We spreken ook af dat we x, y met Hallwaarde nul, dus $x^3 = y^2$ voortaan buiten beschouwing laten. Hieronder volgen alle $x < 10^{11}$ met Hallwaarde< 1.

x	$x^3 - y^2$	Hallwaarde
5234	-17	0.234
8158	-24	0.265
93844	-297	0.969
367806	207	0.341
421351	-618	0.952
720114	-225	0.265
939787	307	0.316
28187351	-1090	0.205
110781386	-8569	0.814
154319269	-11492	0.925
384242766	-14668	0.748
390620082	-14857	0.751
3790689201	-28024	0.455
65589428378	-117073	0.456

Een fysicus zou, aangemoedigd door deze tabel, kunnen concluderen dat $C = 0.2$ in Hall's vermoeden. Er is echter geen enkele garantie daarvoor. Laten we dus verder gaan en grotere waarden van x testen. Nu treedt er echter een probleem op. Vroeger of later botsen we tegen de beperkingen van onze hardware op. Om een idee te geven op welk moment we met deze beperkingen te maken krijgen, stellen voor het gemak dat het testen van één x-waarde ongeveer 10^{-7} seconde kost. Dit zijn ongeveer 100 klok cykels met de tegenwoordige GigaHertz processoren, amper genoeg om een getal van zo'n tien cijfers tot de derde macht te verheffen, een bijbehorende y te vinden en $y^2 - x^3$ te bepalen. Voor het testen van 10^{12} x-waarden hebben we dus ongeveer 10^5 seconden nodig, ofwel 30 uur. Om alle $x < 10^{15}$ te testen hebben we 30.000 uur nodig. Het moge duidelijk zijn dat we hier aan de grens van onze mogelijkheden zitten.
De enige optie om verder te komen is een andere methode te bedenken die ons verder kan brengen. Dit is tevens de grootste uitdaging voor de computationele wiskunde. Goede ideeën die ons rekenbereik significant vergroten zijn zeldzaam en meestal zeer lastig te vinden. Daarom zijn ze goud waard. Een dergelijke doorbraak op het gebied van het experimentele Hall vermoeden werd in 1998 door Noam Elkies gedaan. Door herordening van het zoekbereik en een aantal ideeën uit de diophantische approximatie te gebruiken, slaagde hij erin alle $x < 10^{18}$ te testen. De grens 10^{11} is hiermee letterlijk verpletterd. Als we met de naïeve methode alle getallen $x < X$ willen testen

dan hebben we daarvoor ongeveer γX seconden nodig, waarin γ afhangt van de gebruikte hardware. Een functie die begrensd wordt door een constante te maal X noemen we een *functie van orde* X. Notatie: $O(X)$. De methode van Elkies daarentegen heeft een looptijd van de orde $O(X^{1/2}\log X)$. Een fors verschil met X, vooral als X groot is. Hier volgt een tabel van de nieuwe instanties met Hallwaarde< 1.

x	$x^3 - y^2$	Hallwaarde
12438517260105	2767769	0.784
35495694227489	5190544	0.871
53197086958290	-4401169	0.603
5853886516781223	1641843	0.021
12813608766102806	87002345	0.768
23415546067124892	105077952	0.686
38115991067861271	30032270	0.153
322001299796379844	548147655	0.965
471477085999389882	-497218657	0.724
810574762403977064	-193234265	0.214

In deze tabel komt zelfs een paar x, y met Hallwaarde 0.021 voor! Het even simpele als geniale idee van Elkies zullen we in dit stukje niet uiteenzetten. Daarvoor verwijzen we naar het artikel van Elkies in *Lecture Notes in Computer Science 1838*, Springer Verlag. Maar zelfs met Elkies' nieuwe ideeën stuiten we uiteindelijk ook weer op de beperkingen die de hardware ons oplegt. Uit de tabel krijgen we echter wel het gevoel dat er oneindig veel x met Hallwaarde < 1 zijn.
De Russische wiskundige Danilov baarde in 1980 groot opzien toen hij erin slaagde oneindig veel voorbeelden met Hallwaarde < 1 te construeren. Danilov's constructie is even verrassend als eenvoudig. Hij begint met de identiteit

$$((u-3)^2-5)^3 - (u^2+1)(u^2-9u+19)^2 = 27(2u-11).$$

Stel nu dat we een gehele waarde van u kunnen vinden zó dat $u^2+1 = 125v^2$ voor zekere gehele v. In het bijzonder betekent dit dat $u^2+1 \equiv 0 \pmod 5$. Hieruit volgt dat $u \equiv \pm 3 \pmod 5$. Door het teken van u geschikt te kiezen kunnen we ervoor zorgen dat $u \equiv 3 \pmod 5$. Met deze waarde van u volgt nu dat

$$\left(5\left(\frac{u-3}{5}\right)^2 - 1\right)^3 - v^2(u^2-9u+19)^2 = \frac{27}{125}(2u-11).$$

Stel nu $x = 5\left(\frac{u-3}{5}\right)^2 - 1$ en $y = v(u^2 - 9u + 19)$. Dan hebben we hiermee een paar x, y met Hallwaarde

$$\frac{27|2u - 11|\sqrt{5}}{125\sqrt{(u-3)^2 - 5}}$$

en dit gaat naar $\frac{54}{25\sqrt{5}} = 0.965\ldots < 1$
De vergelijking $u^2+1 = 125v^2$ kan herschreven worden als $u^2-125v^2 = -1$ en staat bekend als de *vergelijking van Pell*. Het is bekend dat deze vergelijking oneindig veel gehele oplossingen heeft. Ze kunnen verkregen worden door de oneven machten van $682 + 61\sqrt{125}$ uit te werken in de vorm $u + v\sqrt{125}$. Bijvoorbeeld, de eerste macht van $682 + 61\sqrt{125}$ geeft $u = 682, v = 61$. We nemen nu $u = -682$ om ervoor te zorgen dat $u \equiv 3 \pmod 5$. Met deze u vinden we $x = 5\left(\frac{u-3}{5}\right)^2 - 1 = 93844$ en $y = v(u^2 - 9u + 19) = 28748141$. Merk op dat $93844^3 - 28748141^2 = -297$. Deze komt inderdaad in de eerste tabel voor. Uitwerking van de derde macht geeft

$$(682 + 61\sqrt{125})^3 = 1268860318 + 113490317\sqrt{125}.$$

We kiezen $u = 1268860318$. De bijbehorende x-waarde is dan

$$x = 322001299796379844,$$

welke in de tweede tabel voorkomt.
Tenslotte nog een waarde die niet in de tabellen voorkomt.

$$(682 + 61\sqrt{125})^5 = 2360712083917682 + 211148507797805\sqrt{125}$$

We kiezen $u = -2360712083917682$. De resulterende x-waarde is

$$x = 1114592308630995805123571151844.$$

We kunnen ons afvragen of er misschien oneindig veel x met Hallwaarde < 0.5 of zelfs < 0.2 kunnen voorkomen. Een expliciete constructie van een dergelijke rij x-waarden zou bijzonder spectaculair zijn!
Bij voorlopig gebrek aan een dergelijke constructie zullen we een experiment uitvoeren waarvan de uitslag suggereert dat er inderdaad zulke rijen zijn. Sterker nog, we zullen argumenten aanvoeren waarom Hall's vermoeden niet juist kan zijn. Anders gezegd, we zullen betogen, maar niet bewijzen, dat er bij elke $C > 0$ een paar $x, y \in \mathbb{N}$ bestaat zó dat $|x^3 - y^2| < C\sqrt{x}$.

3 Een experiment

Kies een getal $C > 0$ en beschouw de ongelijkheid

$$|y^2 - x^3| < C\sqrt{x} \qquad \text{(A)}$$

waarin x, y positief gehele getallen zijn. Als x, y inderdaad aan een dergelijke ongelijkheid voldoen, dan zal, als $x > 4C$, het getal y het unieke gehele getal zijn dat het dichtst bij $x^{3/2}$ ligt. We kiezen nu een interval, bijvoorbeeld $10^6 < x < 10^7$. Vervolgens bepalen we voor elke x daarin het gehele getal y dat het dichtst bij $x^{3/2}$ ligt en kijken of x, y voldoen aan (A). We tellen het aantal x-waarden waarbij dat inderdaad gebeurt. In de volgende tabel vinden we de resultaten voor $x < 10^{11}$ en $C = 20, 10, 5, 2, 1$.

interval	$C = 20$	$C = 10$	$C = 5$	$C = 2$	$C = 1$
$10^3 \leq x < 10^4$	43	16	7	3	2
$10^4 \leq x < 10^5$	37	19	5	2	1
$10^5 \leq x < 10^6$	43	21	12	7	4
$10^6 \leq x < 10^7$	40	18	6	1	0
$10^7 \leq x < 10^8$	40	15	6	2	1
$10^8 \leq x < 10^9$	64	38	20	8	4
$10^9 \leq x < 10^{10}$	55	25	14	3	1

Uit deze tabel wordt het volgende patroon zichtbaar. Voor gegeven $C > 0$ bestaat er een getal n_C zó dat de gevonden aantallen oplossingen van (A) in elk interval $10^n \leq x < 10^{n+1}$ rond deze waarde n_C fluctueren. Voor $C = 20$ ligt n_{20} rond de veertig. In elk interval lijkt zich eenzelfde aantal "hits" af te tekenen. Er zijn natuurlijk fluctuaties, zoals in het interval $10^8 \leq x < 10^9$ waarin de aantallen ver boven het gemiddelde uitsteken. De vraag is natuurlijk of bij uitbreiding van onze tabel de gevonden gemiddelden aan blijven houden of niet. Met veel moeite zouden we de tabel met nog een paar regels kunnen uitbreiden, maar op een gegeven moment houdt het op. Er zijn grenzen, zelfs voor de snelste hedendaagse computers.
We willen dit gedrag verklaren met behulp van een redelijk klinkende aanname. Maar eerst vertalen we ongelijkheid (A) in een iets andere. Uit deze ongelijkheid volgt namelijk dat $|y - x^{3/2}|(y + x^{3/2}) < C\sqrt{x}$ en dus $|y - x^{3/2}| < C\sqrt{x}/(y + x^{3/2})$. We nemen aan dat x groot is ten aanzien van C. In dat geval is namelijk y vrijwel gelijk aan $x^{3/2}$ en we vervangen y in de rechterzijde van onze afschatting gewoon door $x^{3/2}$. We vinden,

$$|y - x^{3/2}| < C\sqrt{x}/(y + x^{3/2}) \approx C/2x$$

De fout die we bij de benadering gemaakt hebben, is miniem en van geen enkele consequentie voor ons verhaal. Uit onze ongelijkheid mogen we afleiden dat y het gehele getal is dat het dichtst bij $x^{3/2}$ ligt. Het verschil van een getal ξ met het dichtstbijgelegen gehele getal zullen we aangeven met $\langle \xi \rangle$. Bijvoorbeeld $\langle \pi \rangle = 0.14159\ldots$ en $\langle e \rangle = -0.2817\ldots$. We vinden nu

$$|\left\langle x^{3/2} \right\rangle| < C/2x \tag{B}$$

Voor practische doeleinden en voor grote x is (B) vrijwel equivalent met ongelijkheid (A).
We formuleren nu onze hypothese. Deze berust erop dat in de rij getallen $\left\langle x^{3/2} \right\rangle$ voor $x = 1, 2, 3, \ldots$ geen enkel herkenbaar patroon zit. Er is geen enkele reden waarom sommige deelintervallen van $[-1/2, 1/2]$ relatief meer waarden van $\left\langle x^{3/2} \right\rangle$ zouden bevatten dan andere deelintervallen. Bij gebrek aan enig zichtbaar patroon maken we de volgende aanname:

De rij getallen $\left\langle x^{3/2} \right\rangle$ *voor* $x = 1, 2, 3, \ldots$ *is verdeeld over het interval* $[-1/2, 1/2]$ *via een uniforme kansverdeling op dit interval.*

Anders gezegd:

Gegeven een interval $I \subset [-1/2, 1/2]$*, dan is de kans dat* $\left\langle x^{3/2} \right\rangle$ *in* I *zit gelijk aan de lengte van* I*.*

Met behulp van deze hypothese kunnen we een verklaring geven van de uitslag van ons experiment. Kies twee grote gehele getallen X, Y met $X < Y$. In onze toepassingen zal Y een bepaalde factor maal X zijn, bijvoorbeeld $Y = 2X$ of $Y = 10X$, zoals in ons experiment. We gaan kijken hoeveel getallen x met $X \leq x < Y$ volgens verwachting voldoen aan ongelijkheid (B). De kans dat x aan (B) voldoet is volgens onze hypothese gelijk aan de lengte van het bijbehorende interval, C/x. De kans dat x niet aan (B) voldoet is derhalve $1 - C/x$.
Met P_n geven we de kans aan dat precies n getallen x met $X \leq x < Y$ aan (B) voldoen. Om te beginnen de kans dat geen enkele x voldoet. Deze is gelijk aan

$$P_0 := \prod_{X \leq x < Y} \left(1 - \frac{C}{x}\right).$$

We zullen de waarde van P_0 niet uitrekenen. De kans P_1 is gelijk aan de som over y van alle kansen dat y aan (B) voldoet en alle andere waarden niet.

Dus

$$P_1 = \sum_{X \le y < Y} \frac{C}{y} \left(1 - \frac{C}{y}\right)^{-1} \prod_{X \le x < Y} \left(1 - \frac{C}{x}\right).$$

Omdat we X zeer groot nemen ten opzichte van C, kunnen we veronderstellen dat de factor $\left(1 - \frac{C}{y}\right)^{-1}$ vrijwel gelijk aan 1 is. Hiermee krijgen we de zeer accurate benadering

$$P_1 \approx \sum_{X \le y < Y} \frac{C}{y} P_0.$$

Uit de appendix van deze tekst blijkt dat we een fout van hooguit $1/X$ maken als we $\sum_{X \le y < Y} \frac{1}{y}$ vervangen door $\log(Y/X)$. Dus,

$$P_1 \approx C \log(X/Y) P_0.$$

De kans dat er twee oplossingen voor (B) zijn in het interval $X \le x < Y$ is gelijk aan

$$P_2 \approx \sum_{X \le y_1 < y_2 < Y} \frac{C}{y_1} \frac{C}{y_2} P_0.$$

Merk op dat

$$\sum_{X \le y_1 < y_2 < Y} \frac{1}{y_1 y_2} = \frac{1}{2} \sum_{\substack{X \le y_1, y_2 < Y \\ y_1 \ne y_2}} \frac{1}{y_1 y_2}.$$

Als we hier $\sum_{X \le y < Y} \frac{1}{y^2}$ bij optellen, dan kunnen we de eis $y_1 \ne y_2$ in de sommatie laten vervallen. Uit de appendix blijkt dat

$$\sum_{X \le y < Y} \frac{1}{y^2} < \frac{1}{X - 1}.$$

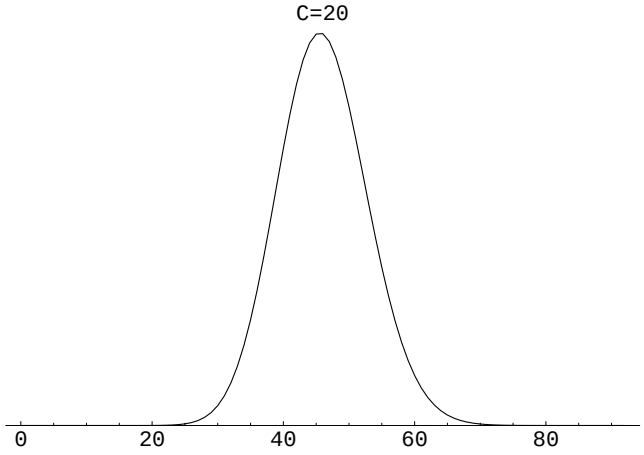
Een zeer klein getal dus. We vinden daarom de accurate benadering

$$\begin{aligned} P_2 &\approx \frac{1}{2} \sum_{X \le y_1, y_2 < Y} \frac{C}{y_1} \frac{C}{y_2} P_0 \\ &\approx \frac{1}{2} \left(\sum_{X \le y < Y} \frac{C}{y} \right)^2 P_0 \\ &\approx \frac{1}{2} (C \log(Y/X))^2 P_0 \end{aligned}$$

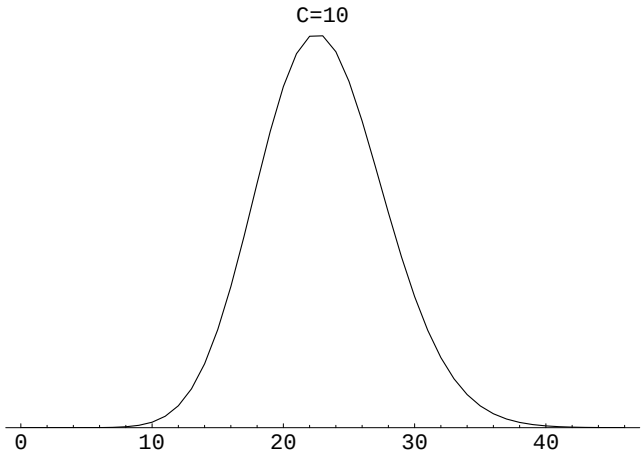
Op analoge wijze vinden we dat

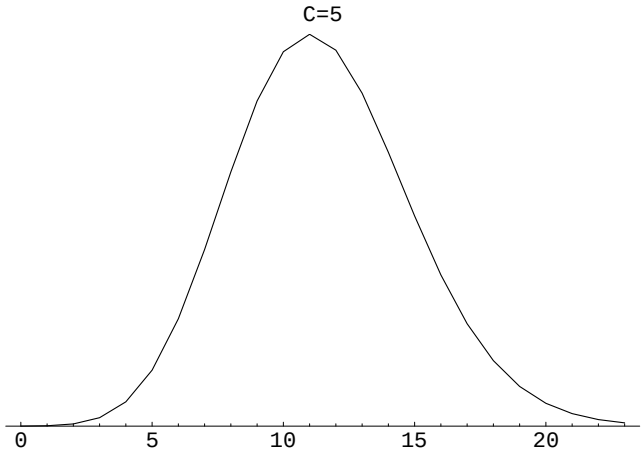
$$P_n \approx \frac{1}{n!}(C \log(Y/X))^n P_0.$$

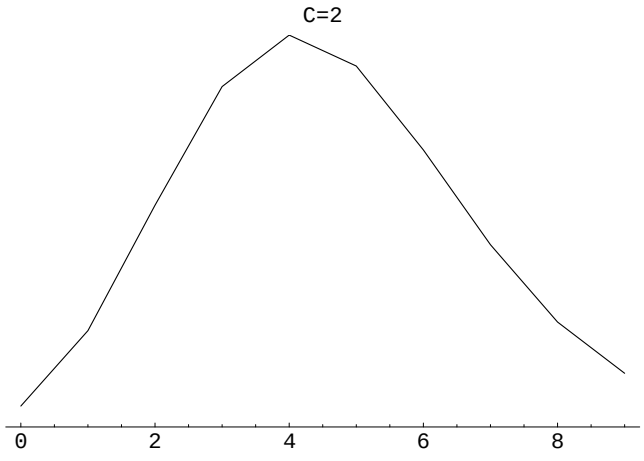
Met andere woorden, de aantallen oplossingen van (B) zijn verdeeld volgens een Poisson verdeling met verwachting $C \log(Y/X)$. We willen hier niet ingaan op de details van het begrip Poissonverdeling. Daarvoor refereren we naar de voordracht van professor Tijms, of naar zijn boek *Spelen met Kansen* in de Epsilonreeks. Wel geven we hier een aantal plots van de grafiek van P_n. In ons experiment hadden we $Y = 10X$ genomen. De grafiek van $\frac{1}{n!}(C \log 10)^n$ met $C = 20$ ziet er als volgt uit.

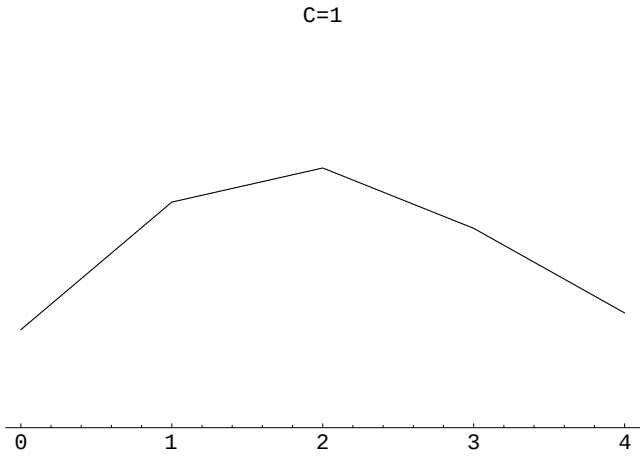


De waarden die we in ons experiment vonden, passen goed in dit plaatje. Enige uitzondering is misschien de waarde 64 die we in het interval $10^8 \leq x < 10^9$ vonden. Deze valt nog net in het bereik waar de kans significant positief is. Hieronder volgen de grafieken van de kansverdelingen bij de overige waarden van C.


C=10
0
10
20
30
40


C=5
0
5
10
15
20


C=2
0
2
4
6
8



Het blijkt dat onze experimentele resultaten redelijk binnen het bereik van bovenstaande grafieken vallen. De spreiding van de grafieken wordt groter naarmate de waarde van C kleiner wordt. Dit hangt samen met het feit dat de standaarddeviatie van een Poissonverdeling met gemiddelde Λ gelijk is aan $\sqrt{\Lambda}$. De verhouding standaarddeviatie/gemiddelde is dus $1/\sqrt{\Lambda}$, een dalende functie in Λ.

4 Conclusies

Het feit dat de voorspellingen in de vorige paragraaf in redelijke overeenstemming zijn met ons experiment, suggereert dat de voorspellingen ook opgaan voor andere waarden van C, X, Y. Neem bijvoorbeeld $C = 0.1$ en als interval $X, Y = 10^{30}X$. Dan geldt $C\log(Y/X) = 6.90\ldots$ en standaarddeviatie $2.6\ldots$. Dit suggereert dat we in elk interval van de vorm $X, 10^{30}X$ een zestal oplossingen van (B), en dus ook (A), mogen verwachten. Nu is de waarde $C = 0.1$ willekeurige gekozen en hadden we ook iedere andere positieve waarde kunnen nemen. Onze conclusie luidt dus dat we verwachten dat het vermoeden van Hall niet waar is. Iets anders gezegd, we verwachten dat er een oneindige rij paren $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n), \ldots$ bestaat waarvan de bijbehorende Hall-waarde naar nul gaat als $n \to \infty$.
Tegenwoordig is men voorzichtiger en formuleert het Hallvermoeden als volgt.

Vermoeden 4.1 (Gewijzigd Hall vermoeden) *Bij elke $\epsilon > 0$ is er een $C(\epsilon) > 0$ zó dat*

$$|y^2 - x^3| > C(\epsilon)x^{1/2-\epsilon}$$

voor elk tweetal $x, y \in \mathbb{N}$ *met* $y^2 \neq x^3$.

We willen dit verhaal besluiten met een kleine waarschuwing die illustreert dat de formulering van aannemelijke hypothesen binnen de getaltheorie een uiterst precaire bezigheid is. Vaak laat men op het Hall-vermoeden een heuristische redenatie toe die maakt gebruik van het feit dat in de buurt van een groot geheel getal X het verschil van twee opeenvolgende kwadraten ongeveer $2\sqrt{X}$ is. Stel namelijk $n^2 \leq X < (n+1)^2$. Dan is n ongeveer $\sqrt{X}$ en het verschil tussen n^2 en $(n+1)^2$ is $2n+1 \approx 2\sqrt{X}$. Kies x willekeurig. In de buurt van x^3 hebben kwadraten een onderling verschil van ongeveer $2x^{3/2}$. De kans dat het interval $[x^3 - C\sqrt{x}, x^3 + C\sqrt{x}]$ een kwadraat bevat is dus de lengte $2C\sqrt{x}$ van het interval gedeeld door $2x^{3/2}$. Met andere woorden, C/x. De verwachting van het aantal x tussen twee grenzen X en Y waarvan de Hallwaarde kleiner dan C is, is dus $\sum_{x=X}^{Y} \frac{C}{x} \approx C \log(Y/X)$. Dit komt goed overeen met onze beschouwingen in de voorgaande paragrafen, en is dus bemoedigend.
Het blijkt echter dat in veel gevallen soortgelijke redenaties volkomen foutieve antwoorden geven. Bijvoorbeeld, de vergelijking $x^3+y^3 = z^3$ met $\text{ggd}(x,y,z) = 1$ heeft alleen maar de triviale oplossingen met $xyz = 0$, terwijl de vergelijking $x^3+y^3+1 = z^3$ er oneindig veel heeft. Kies bijvoorbeeld $x = 9t^4+3t, y = 9t^3-1, z = 9t^4$ met t willekeurig. Vanuit statistisch oogpunt zouden we echter verwachten dat de functies $\sqrt[3]{x^3+y^3}$ en $\sqrt[3]{x^3+y^3+1}$ elkaar niet zoveel ontlopen. Dit geeft al een indicatie dat als er zoiets als zinvolle experimentele getaltheorie bestaat, dan zullen we zeer voorzichtig moeten zijn in het formuleren van goede uitgangspunten.

5 Diversen

Behalve het Hall-vermoeden bestaan er nog vele andere gebieden waarbij men met de computer probeert het antwoord te zoeken. Een bekend voorbeeld is dat van generalisaties van het vermoeden van Fermat. Gegeven een getal $n > 2$, dan weten we sinds 1994 dat de vergelijking van Fermat: $x^n + y^n = z^n$ geen oplossing in positieve gehele getallen x, y, z heeft. Generalisaties van de Fermatvergelijking zijn er ook altijd geweest. Bijvoorbeeld $x^4 + y^4 + z^4 = z^4$. Sinds de 18e eeuw heeft men zich afgevraagd of er behalve de voor de hand liggende oplossing $0^4 + 0^4 + 0^4 + 1^4 = 1^4$ nog andere oplossingen zijn. Ook in het computertijdperk is door veel mensen naar zo'n oplossing gezocht.

Totdat Noam Elkies in 1988 ontdekte dat er oneindig veel oplossingen zijn, met als "kleinste":

$$95800^4 + 217519^4 + 414560^4 = 422481^4.$$

Deze oplossingen werden op puur theoretische gronden ontdekt, er is dus geen computer aan te pas gekomen. De enig bekende methode om via de computer oplossingen van $x^4+y^4+z^4 = u^4$ te bepalen is alle mogelijke drietallen x, y, z met $x \leq y \leq z$ aflopen en testen of de som van hun vierde machten weer een vierde macht is. Nemen we als bovengrens $x, y, z < 500.000$ dan moeten we $\frac{1}{3!}(500.000)^3 \approx 21 \times 10^{15}$ mogelijkheden testen. Veel te veel voor zelfs de snelste PC's. Hoewel er voor 1988 vele pogingen zijn gedaan om oplossingen voor het vierde graads probleem te vinden, is het nu duidelijk waarom dit niet gelukt is.
Dit betekent niet dat computers nutteloos zijn bij het vinden van oplossingen. Er is namelijk een generalisatie van Fermat's vergelijking waarvoor in 1966 door Lander en Parkin wel een oplossing per computer werd gevonden, namelijk:

$$27^5 + 84^5 + 110^5 + 133^5 = 144^5.$$

Ook hier geldt weer dat naïeve zoekacties al gauw stranden op het feit dat looptijden voor het computerprogramma uit de hand lopen. Ondanks de vele pogingen daartoe zijn er nog steeds geen andere oplossingen dan bovenstaande voor het vijfde graads geval gevonden. Overigens zijn er wel oneindig veel gevallen van vijf vijfde machten waarvan de som een vijfde macht is. Deze worden bijvoorbeeld gegeven door de identiteit

$$\begin{aligned}&(75y^5 - x^5)^5 + (x^5 + 25y^5)^5 + (x^5 - 25y^5)^5 + (10x^3y^2)^5 + (50xy^4)^5 \\ &= (x^5 + 75y^5)^5.\end{aligned}$$

Op de website `http://euler.free.fr` wordt een verzameling aangelegd van gevallen waarin sommen van k-de machten gelijk zijn aan sommen van andere k-de machten. Iedereen die zin en klokcykels op de PC over heeft, kan zich aanmelden en meedoen met de zoektocht naar deze rariteiten. Een andere beroemd internetproject is het GIMPS-project, waarin gezocht wordt naar de grootste bekende priemgetallen. De laatste vier records zijn afkomstig uit dit project. Zie `http://www.mersenne.org`, waar ook vele links staan naar andere projecten waar volop aan gerekend wordt.

6 Appendix A

Oplossingen van de Mordell vergelijking $y^2 = x^3 + k$ voor alle k met $0 < |k| \leq 100$. Als een waarde k niet in de tabel voorkomt dan betekent dit dat er geen oplossingen zijn. We schrijven alleen de oplossingen met $y \geq 0$ op en beginnen met $k > 0$.

k	Oplossingen	k	Oplossingen
1	$(-1,0),(2,3),(0,1)$	44	$(-2,6),(5,13)$
2	$(-1,1)$	48	$(1,7)$
3	$(1,2)$	49	$(0,7)$
4	$(0,2)$	50	$(-1,7)$
5	$(-1,2)$	52	$(-3,5)$
8	$(-2,0),(1,3),(2,4),(46,312)$	54	$(3,9)$
9	$(0,3),(-2,1),(3,6),(6,15),(40,253)$	55	$(9,28)$
10	$(-1,3)$	56	$(2,8)$
12	$(-2,2),(13,47)$	57	$(-2,7),(7,20),(4,11)$
15	$(1,4),(109,1138)$	63	$(-3,6),(1,8)$
16	$(0,4)$	64	$(-4,0),(8,24),(0,8)$
17	$(-2,3),(-1,4),(2,5),(8,23),(4,9),$ $(43,282),(52,375),(5234,378661)$	65	$(-4,1),(-1,8),(14,53),$ $(584,14113)$
18	$(7,19)$	68	$(-4,2),(152,1874)$
19	$(5,12)$	71	$(5,14)$
22	$(3,7)$	72	$(-2,8)$
24	$(-2,4),(1,5),(10,32),$ $(8158,736844)$	73	$(-4,3),(2,9),(3,10),(6,17),$ $(72,611),(356,6717)$
25	$(0,5)$	76	$(-3,7)$
26	$(-1,5)$	79	$(45,302)$
27	$(-3,0)$	80	$(-4,4),(1,9),(4,12),(44,292)$
28	$(-3,1),(2,6)$	81	$(0,9)$
30	$(19,83)$	82	$(-1,9)$
31	$(-3,2)$	89	$(-4,5),(-2,9),(10,33),(55,408)$
33	$(-2,5)$	91	$(-3,8)$
35	$(1,6)$	92	$(2,10)$
36	$(-3,3),(0,6),(4,10),(12,42)$	94	$(3,11)$
37	$(-1,6),(3,8),(243,3788)$	97	$(18,77)$
38	$(11,37)$	98	$(7,21)$
40	$(6,16)$	99	$(1,10)$
41	$(2,7)$	100	$(-4,6),(0,10),(5,15),(20,90),$ $(24,118),(2660,137190)$
43	$(-3,4)$		

In de volgende tabel staan de oplossingen van $y^2 = x^3 + k, y \geq 0$ met $k < 0$.

k	Oplossingen	k	Oplossingen
-1	$(1,0)$	-47	$(6,13),(12,41),(63,500)$
-2	$(3,5)$	-48	$(4,4),(28,148)$
-4	$(2,2),(5,11)$	-49	$(65,524)$
-7	$(2,10),(32,181)$	-53	$(9,26),(29,156)$
-8	$(2,0)$	-54	$(7,17)$
-11	$(3,4),(15,58)$	-55	$(4,3),(56,419)$
-13	$(17,70)$	-56	$(18,76)$
-15	$(4,7)$	-60	$(4,2),(136,1586)$
-18	$(3,3)$	-61	$(5,8)$
-19	$(7,18)$	-63	$(4,1),(568,13537)$
-20	$(6,14)$	-64	$(4,0)$
-23	$(3,2)$	-67	$(23,110)$
-25	$(5,10)$	-71	$(8,21)$
-26	$(3,10),(35,207)$	-72	$(6,12)$
-27	$(3,0)$	-74	$(99,985)$
-28	$(4,6),(8,22),(37,225)$	-76	$(5,7),(101,1015)$
-35	$(11,36)$	-79	$(20,89)$
-39	$(4,5),(10,31),(22,103)$	-81	$(13,46)$
-40	$(14,52)$	-83	$(27,140)$
-44	$(5,9)$	-87	$(7,16)$
-45	$(21,96)$	-89	$(5,6)$
		-95	$(6,11)$
		-100	$(5,5),(10,30),(34,198)$

7 Appendix B

In deze appendix geven we afschattingen voor de sommaties

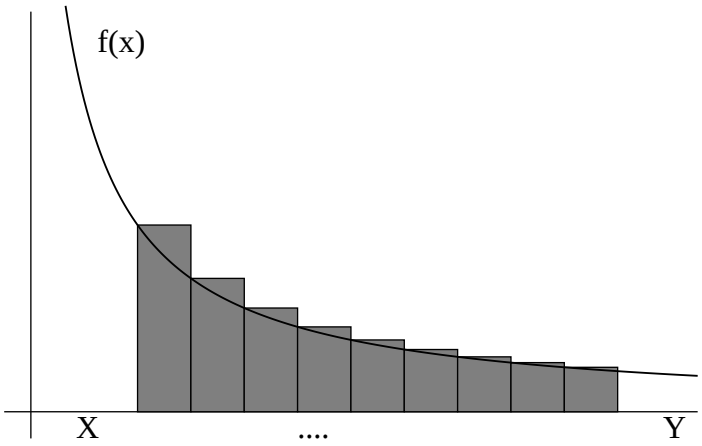
$$S = \sum_{X \leq n < Y} \frac{1}{n^k}$$

voor $k = 1, 2, 3, \ldots$. We gebruiken hiervoor het zogenaamde *integraalcriterium* voor de sommatie van reeksen.

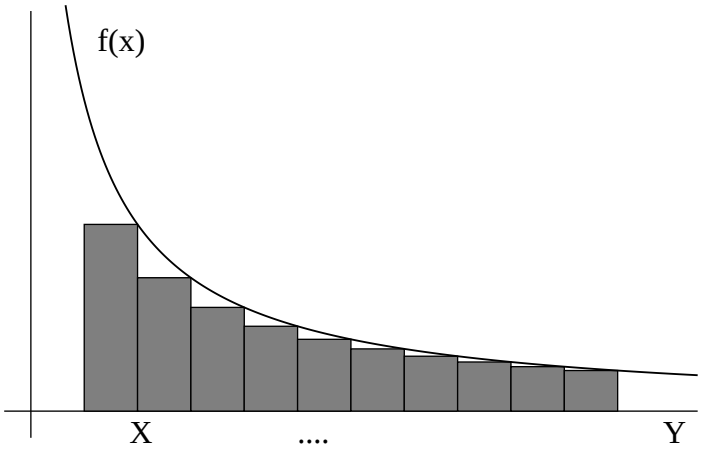
Stelling 7.1 *Zij $f(x)$ een monotoon dalende functie op de positieve reële getallen. Dan geldt voor elk tweetal positief gehele getallen $X < Y$ dat*

$$\int_X^Y f(x)dx \leq \sum_{X \leq n < Y} f(n) \leq f(X) + \int_X^Y f(x)dx.$$

Het bewijs van deze stelling is het makkelijkst aan de hand van een plaatje uit te leggen. De sommatie $\sum_{X \leq n < Y} f(n)$ is gelijk aan het oppervlak van de gearceerde figuur in onderstaand plaatje.



Hieruit volgt direct dat onze sommatie groter is dan de integraal $\int_X^Y f(x)dx$. De bovengrens volgt uit het volgende plaatje.



Het gearceerde gebied bestaat nu uit de eerste kolom met hoogte $f(X)$ en verder een gebied dat onder de grafiek van $f(x)$ ligt. De bovengrens voor onze sommatie volgt nu direct.
Passen we deze Stelling toe op $f(x) = 1/x$ dan vinden we,

$$\sum_{X \leq n < Y} \frac{1}{n} \geq \int_X^Y \frac{1}{x} dx = \log(Y/X)$$

en

$$\sum_{X \leq n < Y} \frac{1}{n} \leq \frac{1}{X} + \int_X^Y \frac{1}{x} dx = \frac{1}{X} + \log(Y/X)$$

.

Het integraalcriterium toegepast op $f(x) = 1/x^2$ geeft

$$\sum_{X \leq n < Y} \frac{1}{n} \leq \frac{1}{X^2} + \int_X^Y \frac{1}{x^2} dx = \frac{1}{X^2} + \frac{1}{X} - \frac{1}{Y}.$$

Het laatste getal is weer kleiner dan $\frac{1}{X^2} + \frac{1}{X}$.