Radboud Universiteit

Bachelor Scriptie

De vergelijking van Schröder

Voor functies met vaste punten

Auteur:
Jasper de Klein

Supervisor:
Michael Müger

June 25, 2015

VOORWOORD

Toen ik in 2012 aan mijn studie wiskunde begon, wist ik niet wat ik precies kon verwachten. Meteen werd me duidelijk dat wiskunde veel breder was dan ik ooit had gedacht. Ik kwam al snel in aanraking met vakken uit verschillende vakgebieden binnen de wiskunde en natuurlijk ontstond er een voorkeur voor bepaalde vakgebieden.

Halverwege mijn eerste jaar kwam ik voor het eerst echt in aanraking met Analyse. Ik kon me in eerste instantie niet zo goed voorstellen wat het doel ervan was, maar nadat ik er een tijdje mee bezig was, begon het me steeds meer te fascineren. Tegelijkertijd volgde ik ook het vak *Dynamische Systemen*, wat ik nog interessanter vond.

Op een gegeven moment kwam daar het punt aan dat ik mijn scriptie moest gaan schrijven. Ik had nog geen duidelijke idee waarover ik mijn scriptie precies wilde schrijven, maar over het vakgebied ontstond weinig twijfel. Óf Analyse óf Dynamische systemen. Na kort overleg kwam mijn supervisor Michael Müger met een stelling die Carl Ludwig Siegel in 1942 bewees in zijn artikel *Iteration of analytic functions* gepubliceerd in Annals of Mathematics.

Het betreft een stelling uit de complexe analyse die toentertijd veel stof deed opwaaien en waar inmiddels meerdere bewijzen voor zijn. Natuurlijk Siegels eigen bewijs, maar onder andere ook nog een bewijs dat hij samen met een van zijn studenten Jürgen Kurt Moser uitbracht in *Lectures on Celestial Mechanics*. Waar het ene bewijzen puur analytisch is, bevat het andere wat raakvlakken met combinatoriek en dynamische systemen.

Het leek mij interessant om beide bewijzen te bestuderen. Van het bewijs van Siegel en Moser heb ik een herschrijving bestudeerd in het boek *Complex Dynamics* van Lennart Carleson en Theodore Gamelin, waarbij ik af en toe stukken heb nagelezen in bovengenoemd boek van Siegel en Moser. Het bewijs dat Siegel in eerste instantie gaf, heb ik bestudeerd aan de hand van het artikel *A Brief but Historic Article of Siegel* van Rodrigo Pérez.

Beide bewijzen waren ontzettend beknopt opgeschreven en mijn doel was om in mijn scriptie de bewijzen toegankelijk te maken voor studenten met slechts een basiskennis van Complexe analyse. Ik heb met veel plezier aan mijn scriptie gewerkt en wil Michael Müger dan ook bedanken voor het leuke voorstel en de hulp die hij mij geboden heeft.

Jasper de Klein

CONTENTS

SYMBOLEN EN NOTATIES

λ	multiplier bij z_0	$f'(z_0))$ met z_0 een vast punt van f
θ	hoek van λ	in radialen

Notatie 1. We schrijven $f^1 = f$ en $f^n = f^{n-1} \circ f$, dus

$$f^n(z) = \underbrace{f\big(f(\cdots f}_{\text{n}}(z))\big)$$

Notatie 2. We schrijven $f^{(n)}$ voor de n^{de} afgeleide van f.

Notatie 3. We schrijven $\triangle(0, r)$ voor $\{z \in \mathbb{C} : |z| < r\}$

1. VASTE PUNTEN, CONJUGERENDE AFBEELDINGEN EN ANALYTISCHE FUNCTIES

1.1 Vaste punten

We gaan functies $f(z) : \mathbb{C} \supseteq \mathcal{U} \to \mathcal{V} \subseteq \mathbb{C}$ met een vast punt bekijken. Hierbij definiëren we als volgt:

Definitie 1.1.1. We noemen z_0 een *vast punt* van f als $f(z_0) = z_0$. Daarnaast noemen we $\lambda = f'(z_0)$ de *multiplier* van z_0.

Als $|\lambda| = 0$ heet z_0 een *superaantrekkend* vast punt.

Als $0 < |\lambda| < 1$ heet z_0 een *aantrekkend* vast punt.

Als $|\lambda| = 1$ heet z_0 een *neutraal* vast punt.

Als $|\lambda| > 1$ heet z_0 een *afstotend* vast punt.

Voorbeeld 1.1.2. De functie $f(x) = \sqrt{x}$ heeft vast punt $x_0 = 1$ met multiplier $\lambda = \frac{1}{2}$, dus x_0 is een aantrekkend vast punt. Uit een simpele berekening volgt $d\big(x_0, f^n(x)\big) < d\big(x_0, f^{n-1}(x)\big) \forall x \in \mathbb{R}_+$, wat de naam verklaart.

Voorbeeld 1.1.3. De functie $g(x) = x^2$ heeft vast punt $x_0 = 1$ met multiplier $\lambda = 2$, dus x_0 is een afstotend vast punt. Een simpele berekening verklaart de naam afstotend, want $d\big(x_0, g^n(x)\big) > d\big(x_0, g^{n-1}(x)\big) \forall x \in \mathbb{R}_+$.

Opmerking 1.1.4. Als x_0 een neutraal vast punt is, is het gedrag van punten uit de omgeving van van x_0 niet te voorspellen zonder extra informatie (te halen uit hogere orde afgeleide), zie **Voorbeeld 1.1.5**

Voorbeeld 1.1.5. De functie $h(x) = x + x^2$ heeft vast punt $x_0 = 0$ met multiplier $\lambda = 1$ en is dus een neutraal vast punt. Het is eenvoudig in te zien dat $d\big(x_0, f^n(x)\big) < d\big(x_0, f^{n-1}(x)\big) \forall x \in (-1, 0)$ en
$d\big(x_0, f^n(x)\big) > d\big(x_0, f^{n-1}(x)\big) \forall x \in (0, 1)$. Het vaste punt x_0 is dus links aantrekkend en rechts afstotend.

Definitie 1.1.6. Laat z_0 een neutraal vast punt zijn. We noemen z_0 een *rationaal neutraal vast punt* als $\lambda^n = 1$ voor zekere $n \in \mathbb{N}$ en een *irrationaal neutraal vast punt* als $\lambda \neq 1$ voor alle $n \in \mathbb{N}$.

Propositie 1.1.7. *Laat z_0 een neutraal vast punt zijn. Dan is z_0 een rationaal vast punt $\Longleftrightarrow \lambda = e^{2\pi i\theta}$ met $\theta \in \mathbb{Q}$ en z_0 is een irrationaal vast punt $\Longleftrightarrow \lambda = e^{2\pi i\theta}$ met $\theta \in \mathbb{R}\backslash\mathbb{Q}$*

Bewijs. $|\lambda = 1| \Longleftrightarrow \lambda = e^{2\pi i\theta}$, dus $\lambda^n = 1 \Longleftrightarrow e^{2\pi i\theta n} = e^{2\pi ik}$ met $n, k \in \mathbb{N}$. Dus $\theta n = k$, ofwel $\theta = n/k \in \mathbb{Q}$. Hieruit volgt meteen dat $\lambda^n \neq 1 \forall n \in \mathbb{N} \Longleftrightarrow \theta \in \mathbb{R}\backslash\mathbb{Q}$ ◻

1.2 Conjugerende afbeeldingen

We gaan kijken of we deze functies kunnen conjugeren aan een andere functie. Hierbij definiëren we als volgt:

Definitie 1.2.1. De functie $f : \mathcal{U} \to \mathcal{U}$ is *geconjugeerd* aan $g : \mathcal{V} \to \mathcal{V}$ als er een inverteerbare afbeelding $\varphi : \mathcal{U} \to \mathcal{V}$ bestaat, zodat

$$g = \varphi \circ f \circ \varphi^{-1}$$

ofwel

$$g\big(\varphi(z)\big) = \varphi\big(f(z)\big). \tag{1.1}$$

Vergelijking (1.1) wordt ook wel de *Schröders vergelijking* genoemd.

Propositie 1.2.2. *Als f door φ geconjugeerd wordt aan g, dan conjugeerd φ ook de iteraties f^n aan g^n voor alle $n \in \mathbb{N}$. Ook de inverse functies van f en g zijn geconjugeerd door φ. Verder worden vaste punten van f door φ afgebeeld op vaste punten van g,.*

Bewijs. Het bewijs volgt eenvoudig door uitschrijven. We bekijken

$$\begin{aligned} g^n &= \varphi \circ f \circ \varphi^{-1} \circ \varphi \circ f \circ \varphi^{-1} \circ \cdots \circ \varphi \circ f \circ \varphi^{-1} \\ &= \varphi \circ f \circ f \cdots \circ f \circ \varphi^{-1} \\ &= \varphi \circ f^n \circ \varphi^{-1}. \end{aligned}$$

En

$$\begin{aligned} g^{-1} &= (\varphi \circ f \circ \varphi^{-1})^{-1} \\ &= (\varphi^{-1})^{-1} \circ f^{-1} \circ \varphi^{-1} \\ &= \varphi \circ f^{-1} \circ \varphi^{-1}. \end{aligned}$$

Laat z_0 een vast punt zijn van f. De functies f en g zijn geconjugeerd door φ, dus

$$\begin{aligned} \varphi\big(f(z)\big) &= g\big(\varphi(z)\big) \\ \varphi\big(f(z_0)\big) &= g\big(\varphi(z_0)\big) \\ \varphi(z_0) &= g\big(\varphi(z_0)\big). \end{aligned}$$

We zien dat $\varphi(z_0)$ een vast punt van g is. □

1.3 Analytische functie

Definitie 1.3.1. Een complexe functie $f : \mathcal{U} \to \mathbb{C}$ is analytisch in het domein $\mathcal{U}$ als f in ieder punt uit zijn domein continu afgeleide heeft. Hieruit volgt dat f in ieder punt gelijk aan zijn Taylorreeksontwikkeling, ofwel:

$$f(z) = \sum_{n=0}^{\infty} \frac{f^{(n)}(z_0)}{n!}(z - z_0)^n, z_0 \in \mathcal{U}$$

Opmerking 1.3.2. Als f een analytische functie is, is f in ieder punt van zijn domein te schrijven als machtreeks. Laat z_0 een vast punt van f met multiplier λ zijn, dan:

$$f(z) = z_0 + \lambda(z - z_0) + \sum_{n=2}^{\infty} a_n (z - z_0)^n \tag{1.2}$$

2. AANTREKKENDE EN AFSTOTENDE VASTE PUNTEN

2.1 Aantrekkende vaste punten

Definitie 2.1.1. Laat f een functie zijn. Een *schaling* van f is een vermenigvuldiging van f met $c \in \mathbb{C}\backslash\{0\}$.

Lemma 2.1.2. *Als $f(z)$ een analytische functie is met aantrekkend vast punt z_0, dus multiplier $0 < |\lambda| < 1$, dan is een conjugatie φ van $f(z)$ aan $g(z) = \lambda(z - z_0)$ uniek tot op een schaling.*

Bewijs. Eerst bewijzen we dat de samenstelling van een conjugatie van f aan g en een conjugatie van g aan zichzelf een conjugatie van f aan g geeft. Laat φ een conjugatie van f aan g zijn en ψ een conjugatie van g aan zichzelf. Dan zien we

$$\begin{aligned} f(z) &= \varphi \circ g(z) \circ \varphi^{-1} \\ &= \varphi \circ \left(\psi \circ g(z) \circ \psi^{-1}\right) \circ \varphi^{-1} \\ &= \varphi \circ \psi \circ g(z) \circ (\varphi \circ \psi)^{-1}. \end{aligned}$$

Dus $\varphi \circ \psi$ is een conjugatie van f aan g.

Het is nu voldoende om te bewijzen dat iedere conjugatie ψ van $g(z) = \lambda(z - z_0)$ aan zichzelf een schaling is. Stel namelijk dat er een conjugatie ψ_1 van $g(z)$ aan zichzelf bestaat die geen schaling is. Dan zou de samenstelling van φ met ψ_1 een conjugatie van $f(z)$ aan $g(z)$ geven, die geen schaling van φ is en bestaat er dus een nog een andere conjugatie.

Merk op dat als de conjugatie ψ_1 wel een schaling is, dat $\varphi \circ \psi_1$ dan een schaling van φ is. Als alle conjugaties ψ van $g(z)$ aan zichzelf schalingen zijn, dan is de conjugatie φ dus uniek op schaling na.

Stel nu dat $\psi = a_1 z + a_2 z^2 + \ldots$ is een conjugatie van $g(z)$ aan zichzelf, ofwel
$\psi(\lambda z) = \lambda\psi(z)$. We bekijken $\psi(\lambda z)$ en $\lambda\psi(z)$

$$\begin{aligned}\psi(\lambda z) &= a_1(\lambda z) + a_2(\lambda z)^2 + a_3(\lambda z)^3 + \ldots \\ &= a_1\lambda z + a_2\lambda^2 z^2 + a_3\lambda^3 z^3 + \ldots\end{aligned}$$

$$\begin{aligned}\lambda\psi(z) &= \lambda(a_1 z + a_2 z^2 + a_3 z^3 + \ldots) \\ &= a_1\lambda z + a_2\lambda z^2 + a_3\lambda z^3 + \ldots.\end{aligned}$$

Als we de coëfficiënten vergelijken, zien we $a_1\lambda = a_1\lambda$ en

$$\left.\begin{aligned} a_n\lambda^n = \lambda a_n,\ \forall n \geq 2 \\ \lambda \neq 0 \text{ en } \lambda^n \neq 1 \\ \lambda^{n+1} \neq \lambda \end{aligned}\right\} a_n = 0,\ \forall n \geq 2.$$

Dus $\psi(z) = a_1 z$ en hieruit volgt dat $\varphi(z)$ uniek is tot op een schaling na. $\square$

Lemma 2.1.3. *Zij $\{f_n(z)\}$ een rij functies met*

$$|f_{n+1}(z) - f_n(z)| \leq C^n|z|^2, \quad C < 1, |z| \leq \delta$$

Dan convergeert $\{f_n(z)\}$ uniform voor $|z| \leq \delta$.

Bewijs. Zij $i > j$. We bekijken

$$f_i(z) - f_j(z) = \sum_{k=j}^{i-1} \big(f_{k+1}(z) - f_k(z)\big)$$

dus

$$\begin{aligned} |f_i(z) - f_j(z)| &\leq \sum_{k=j}^{i-1} |f_{k+1}(z) - f_k(z)| \\ &\leq \sum_{k=j}^{i-1} C^k |z|^2 \leq \sum_{k=j}^{i-1} C^k \delta^2 \\ &\leq \sum_{k=j}^{\infty} C^k \delta^2 \\ &= \frac{\delta^2}{1-C} C^j. \end{aligned}$$

Dus de afstand tussen f_i en f_j wordt willekeurig klein als $j \to \infty$, ongeacht van z, dus $\{f_n(z)\}$ convergeert uniform naar $f(z)$.

Stelling 2.1.4. *Stel $f(z)$ is een analytische functie met aantrekkend vast punt z_0, dan bestaat er een afbeelding $\varphi(z)$ van een omgeving van z_0 naar een omgeving van de oorsprong die $f(z)$ conjugeert aan de lineaire functie $g(z) = \lambda(z - z_0)$. Deze conjugerende functie is uniek tot op schaling na.*

Bewijs. Zonder verlies van algemeenheid mogen we aannemen dat $z_0 = 0$, want als dit niet zo is, kunnen we een translatie toepassen op$f(z)$ zodat 0 het vaste punt wordt.

We gaan nu een functie φ construeren zodat $\varphi\big(f(z)\big) = \lambda\varphi(z)$. We definiëren

$$\varphi_n(z) = \lambda^{-n} f^n(z).$$

Dan voldoet φ_n aan

$$\varphi_n \circ f = \lambda^{-n} f^{n+1} = \lambda\varphi_{n+1},$$

dus als $\varphi_n \to \varphi$, dan $\varphi \circ f = \lambda\varphi$, ofwel $\varphi \circ f \circ \varphi^{-1} = \lambda z$. Dan is φ de gezochte conjugatie.

Om de convergentie aan te tonen, merk op dat voor $\delta_0 > 0$ klein en zekere $C \in \mathbb{R}$ en $|z| \leq \delta_0$

$$\begin{aligned} |f(z) - \lambda z| &\leq C|z|^2 \\ |f(z)| &\leq |\lambda||z| + C|z|^2 \\ &\leq (|\lambda| + C\delta_0)|z|. \end{aligned}$$

We kunnen C zo kiezen dat $|\lambda| + C\delta < 1$, want $|\lambda| < 1$, dus dan volgt met inductie dat

$$f^n(z)| \leq (|\lambda| + C\delta_0)^n|z|, \qquad |z| \leq \delta.$$

Kiezen we $\delta_1 \leq \delta_0$ zo klein dat $p = \dfrac{(|\lambda| + C\delta_1)^2}{|\lambda|} < 1$, krijgen we

$$\begin{aligned} |\varphi_{n+1}(z) - \varphi_n(z)| &= \left| \frac{f^n\big(f(z)\big) - \lambda f^n(z)}{\lambda^{n+1}} \right| \\ &\leq \frac{C|f^n(z)|^2}{|\lambda|^{n+1}} \\ &\leq \frac{p^n C|z|^2}{|\lambda|} \end{aligned}$$

Dus met **Lemma 2.1.3** volgt dat $\varphi_n(z)$ uniform convergeert voor $|z| \leq \delta_1$ en de conjugatie φ bestaat. Uniciteit, op schaling na, volgt uit **Lemma 2.1.2**. □

2.2 Afstotende vaste punten

Stelling 2.2.1. *Stel $f(z)$ is een analytische functie met afstotend vast punt z_0, dan bestaat er een afbeelding $\varphi(z)$ van een omgeving van z_0 naar een* *omgeving van de oorsprong die $f(z)$ conjugeert aan de lineaire functie $g(z) = \lambda(z - z_0)$. Deze conjugerende functie is uniek tot op schaling na.*

Bewijs. Dit volgt direct uit **Stelling 2.1.4**.
Stel namelijk $f(z) = z_0 + \lambda(z - z_0) + \ldots$ met vast punt z_0 en $\lambda > 1$. Uit $f'(z_0) \neq 0$ volgt dat f locaal inverteerbaar is. Bekijken we nu $f^{-1}(z) = z_0 + \frac{1}{\lambda}(z - z_0) + \ldots$ dan zien we dat z_0 nog steeds een vast punt is met multiplier $\frac{1}{\lambda} < 1$. Met **Stelling 2.1.4** volgt dat er een φ bestaat die $f^{-1}(z)$ conjugeert aan $g^{-1}(z) = \frac{1}{\lambda}(z - z_0)$, omdat z_0 nu een aantrekkend vast punt is. En uit **Propositie 1.2.2** volgt dat φ ook $f(z)$ aan $g(z)$ conjugeert. $\square$

3. IRRATIONALE NEUTRALE VASTE PUNTEN

3.1 Diophantiene getallen

Definitie 3.1.1. Een reëel getal θ heet *Diophantien* als het slecht benaderbaar is door rationale getallen, in de zin dat er een $\mu, c > 0$ bestaan, zodat $\forall p, q \in \mathbb{Z}, q \neq 0$

$$\left|\theta - \frac{p}{q}\right| \geq \frac{c}{q^{\mu}}.$$

De kleinste μ waarvoor bovenstaande conditie waar is voor een getal θ, noemen we de Diophantische orde van θ.

Lemma 3.1.2. *Voor $\lambda = e^{2\pi i\theta} \in \mathbb{C}$ is de conditie dat θ Diophantien is van orde μ equivalent aan*

$$|\lambda^n - 1| \geq \frac{C}{n^{\mu-1}}.$$

Bewijs. Zij $\lambda = e^{2\pi i\theta}$ met θ Diophantien, dus

$$\begin{aligned} |\lambda^n - 1| &= |e^{2\pi i\theta n} - 1| \\ &= |e^{\pi i\theta n} - e^{-\pi i\theta n}| \\ &= 2|\sin(\pi\theta n)|. \end{aligned}$$

De eerste gelijkheid hierbij is duidelijk. De tweede gelijkheid geldt omdat de absolute waarde niet verandert door te vermenigvuldigen met $|e^{-\pi i\theta}|$. Voor de laatste gelijkheid maken we gebruik van $\sin(x) = \dfrac{e^{ix} - e^{-ix}}{2i}$.

Laat m het dichtstbijzijnde gehele getal zijn van $n\theta$, dus $|n\theta - m| < 1/2$. Hier geldt een strikte ongelijkheid, omdat θ irrationaal is. Omdat $\sin(x)$ periodiek is, geldt dat

$$|\lambda^n - 1| = 2\sin(\pi|n\theta - m|).$$

Merk verder op dat de grafiek van $\sin(\pi x)$ tussen de lijnen $y_1 = 2x$ en $y_2 = \pi x$ voor $0 \leq x \leq 1/2$, dus

$$4|n\theta - m| \leq |\lambda^n - 1| < 2\pi|n\theta - m|.$$

Omdat θ Diophantien is, dus $\left|\theta - \dfrac{m}{n}\right| \geq \dfrac{C_0}{n^\mu}$, volgt met de eerste ongelijkheid

$$\begin{aligned} \frac{4C_0}{n^\mu} &\leq \frac{|\lambda^n - 1|}{n} \\ \frac{C}{n^{\mu-1}} &\leq |\lambda^n - 1|. \end{aligned}$$

Andersom, als we weten dat $\frac{C}{n^{\mu-1}} \leq |\lambda^n - 1|$ met $n \neq 0$, zien we

$$\frac{C}{n^\mu} \leq \frac{|\lambda^n - 1|}{n}.$$

We zagen al dat

$$\begin{aligned} |\lambda^n - 1| &= 2|\sin(\pi n\theta)| \\ &= 2|\sin(\pi n\theta n - m)| \\ &\leq 2\pi|n\theta - m|. \end{aligned}$$

De tweede gelijkheid geldt omdat $|\sin(\pi x)|$ periodiek is met periode één en

$m \in \mathbb{Z}$. De ongelijkheid geldt omdat $\sin|\pi x| \leq |\pi x|$. We hebben dus

$$\frac{C_0}{n^\mu} \leq \frac{2\pi|n\theta - m|}{n}$$
$$\frac{C_0}{2\pi n^\mu} \leq |\theta - \frac{m}{n}|$$
$$\frac{C}{n^\mu} \leq |\theta - \frac{m}{n}|$$

voor alle $n, m \in \mathbb{Z}, n \neq 0$. □

Opmerking 3.1.3. Voor vaste $\mu > 2$ voldoen bijna alle reële getallen θ. Laat $E = \{\theta \in [0,1] \mid |\theta - p/q| < q^{-\mu}$ oneindig vaak$\}$, dus de verzameling van alle getallen (tussen 0 en 1) die *niet* Diophantien zijn. We kunnen de maat van E op de volgende manier schatten;

$$|E| \leq \sum_{q=n}^{\infty} 2 \cdot q^{-\mu} \cdot q = \mathcal{O}(n^{2-\mu}) \to 0$$

3.2 Irrationale neutrale vaste punten

Lemma 3.2.1. *Als f een holomorfe fucntie is op een open verzameling Ω die een schijf D van straal r om z_0 en zijn rand C bevat, dan*

$$|f^{(n)}(z_0)| \leq \frac{n!\,\|f\|_C}{r^n}$$

waar $\|f\|_C = \sup_{z \in C} |f(z)|$.

Lemma 3.2.2. *Laat f holomorfe functie zijn op een open verzameling Ω. Als D een schijf is rond z_0, waarvan de afsluiting bevat is in Ω, dan heeft f*

in z_0 *de machtreeksontwikkeling*

$$f(z) = \sum_{n=0}^{\infty} a_n (z - z_0)^n$$

voor alle $z \in D$*. De coëfficiënten* a_n *worden gegeven door*

$$a_n = \frac{f^{(n)}(z_0)}{n!}$$

voor alle $n \geq 0$.

Voor beide lemma's kan een bewijs worden gevonden in het boek *II Complex Analysis* van E. M. Stein en R. Shakarchi.

Propositie 3.2.3. *Laat* f *een holomorfe functie zijn op een open verzameling* Ω*. Als* D *een schijf is rond* z_0*, waarvan de afsluiting bevat is in* Ω*, dan geldt voor de coëfficiënten van de machtreeks:*

$$|a_n| \leq \frac{\|f\|_C}{r^n}$$

waar $\|f\|_C = \sup_{z \in C} |f(z)|$

Bewijs. Dit volgt direct uit bovenstaande lemma's. □

Lemma 3.2.4. *Laat* f *een holomorfe functie zijn op een open schijf* D *en* $v, w \in D$*, dan*

$$|f(v) - f(w)| \leq \sup_{z \in D} |f'(z)| \cdot |v - w|.$$

Bewijs. De verzameling D is convex, dus er bestaat een lijn tussen de punten v en w. Als we de afgeleide van f integreren langs deze lijn, weten we.

$$f(v) - f(w) = \int_v^w f'(z) \, \mathrm{d}z.$$

Hieruit volgt dat

$$|f(v) - f(w)| \leq \sup_{z\in D} |f'(z)| \cdot |v - w|.$$

$\square$

Stelling 3.2.5. *Laat θ Diophantien zijn en f een analytische functie zijn met vast punt in $z_0 = 0$ en multiplier $\lambda = e^{2\pi i\theta}$. Dan bestaat er een oplossing voor $\varphi\big(f(z)\big) = \lambda\varphi(z)$, ofwel f kan in een omgeving van 0 geconjugeerd worden aan een vermenigvuldiging met λ.*

Bewijs. We willen een oplossing voor de Schröder vergelijking $\varphi\big(f(z)\big) = \lambda\varphi(z)$ genormaliseerd door $\varphi'(0) = 1$. Nemen we $h = \varphi^{-1}$, dan krijgen we

$$f\big(h(z)\big) = h(\lambda z), \quad h'(0) = 1. \tag{3.1}$$

We definiëren $\hat{f}$ en $\hat{h}$ door $f(z) = \lambda z + \hat{f}(z)$ en $h(z) = z + \hat{h}(z)$. Nu kunnen we, gebruikmakend van $\hat{f}$ en $\hat{h}$, de vergelijking omschrijven:

$$\begin{aligned} h(\lambda z) &= f\big(h(z)\big) \\ h(\lambda z) - \lambda h(z) &= f\big(h(z)\big) - \lambda h(z) \\ \big(\lambda z + \hat{h}(\lambda z)\big) - \big(\lambda z + \lambda\hat{h}(z)\big) &= \big(\lambda h(z) + \hat{f}(h(z))\big) - \lambda h(z) \\ \hat{h}(\lambda z) - \lambda\hat{h}(z) &= \hat{f}\big(h(z)\big). \end{aligned} \tag{3.2}$$

We gebruiken nu KAM theorie (genoemd naar A.N. Kolmogorov, V.I. Arnold, J. Moser) in een simpel geval. We bekijken coördinaat veranderingen ψ rond $z = 0$, genormaliseerd zodat $\psi(z) = z + \hat{\psi}(z)$ met $\hat{\psi} = \mathcal{O}(z^2)$. We vinden niet direct een oplossing voor (3.1), maar we maken ψ zo dat

$$\psi^{-1} \circ f \circ \psi = g(z) = \lambda z + \hat{g}(z)$$

met $\hat{g}$ kleiner dan $\hat{f}$ in zekere zin. Merk hierbij op dat $\hat{g} = 0$ de oplossing geeft. We blijven dit proces herhalen met f vervangen door g en ψ gedefiniëerd op een iets kleinere schijf.
Schrijf nu $f = \sum_{n=2}^{\infty} b_n z^n$, als we in plaats van (3.2) de lineaire variant nemen, dus we vervangen $h(z)$ aan de rechterkant door z, krijgen we:

$$\hat{\psi}(\lambda z) - \lambda \hat{\psi}(z) = \hat{f}(z).$$

Dit kunnen we eenvoudig oplossen voor $\hat{\psi}$

$$\hat{\psi}(z) = \sum_{j=2}^{\infty} \frac{b_j}{\lambda^j - \lambda} z^j.$$

We gaan nu $\hat{g}$ en $\hat{g}'$ afschatten onder de volgende aannamen

$$\frac{1}{|\lambda^n - 1|} \leq c_0 \frac{n^\mu}{\mu!},$$

$$|f'(z)| < \delta \text{ voor } z \in \triangle(0, r).$$

De eerste ongelijkheid geldt omdat θ Diophantien is met (vaste) Diophantische orde μ, zie hiervoor ook **Lemma 3.1.2**.
Gebruikmakend van de twee parameters δ, r willen we een afschatting maken van $g = \psi^{-1} \circ f \circ \psi$. Eerst schatten we $\hat{\psi}$ op een kleinere schijf $\triangle\big(0, r(1-\eta)\big)$ voor zekere $0 < \eta < 1/5$. Met behulp van de **Propositie 3.2.3** voor de machtreekscoëfficiënten van $\hat{f}'$, hebben we

$$|b_j| \leq \frac{\delta}{j r^{j-1}}.$$

Dus voor $z \in \triangle\big(0, r(1-\eta)\big)$ zien we

$$
\begin{aligned}
|\hat{\psi}'| &\leq \sum_{j=2}^{\infty} \frac{jb_j}{|\lambda^j - \lambda|} r^{j-1}(1-\eta)^{j-1} \\
&\leq \sum_{j=2}^{\infty} \frac{j\delta}{|\lambda^j - \lambda| j r^{j-1}} r^{j-1}(1-\eta)^{j-1} && \text{want } |b_j| \leq \frac{\delta}{jr^{j-1}} \\
&\leq \sum_{j=2}^{\infty} c_0 \delta \frac{j^\mu}{\mu!}(1-\eta)^{j-1} && \text{want } \frac{1}{|\lambda||\lambda^j - 1|} \leq c_0 \frac{(j-1)^\mu}{\mu!} \leq c_0 \frac{j^\mu}{\mu!} \\
&\leq \frac{c_0\delta}{\mu!} \sum_{j=1}^{\infty} j(j+1)\cdots(j+\mu-1)(1-\eta)^{j-1} \\
&= \frac{c_0\delta}{\eta^{\mu+1}}
\end{aligned}
$$

De laatste gelijkheid is een speciaal geval van de binomiaalreeks, $(1+x)^n = \sum_{k=0}^{n} \binom{n}{k} x^k$. Als we aannemen dat $c_0\delta < \eta^{\mu+2}$, hebben we

$$|\hat{\psi}'| \leq \eta. \tag{3.3}$$

Omdat

$$
\begin{aligned}
\psi(z) &= z + \hat{\psi}(z) && \text{volgt dat} \\
|\psi(z)| &\leq |z| + |\eta z| && \text{voor } |z| \leq r(1-4\eta) \\
&\leq r(1-4\eta) + r\eta \\
&= r(1-3\eta).
\end{aligned}
$$

Dus $\psi(z)$ beeldt $\triangle\big(0, r(1-4\eta)\big)$ af op $\triangle\big(0, r(1-3\eta)\big)$. Verder impliceert

$|\hat{f}| \leq \delta$ dat $\hat{f}(z) \leq \delta z$, zoals hierboven. Dus als $\delta < \eta$

$$\begin{aligned} f(z) &= \lambda z + \hat{f}(z) \\ |f(z)| &= |\lambda z| + |\hat{f}(z)| \\ &\leq |z| + |\eta z| \qquad\qquad \text{dus voor } |z| \leq r(1-3\eta) \text{ is} \\ &\leq r(1-3\eta) + \eta r \\ &= r(1-2\eta). \end{aligned}$$

Dus f beeld $\triangle\big(0, r(1-3\eta)\big)$ af op $\triangle\big(0, r(1-2\eta)\big)$. We claimen verder dat ψ^{-1} de verzameling $\triangle\big(0, r(1-2\eta)\big)$ op $\triangle\big(0, r(1-\eta)\big)$ afbeeldt.

Om dit te bewijzen laten we zien dat dat voor $z \in \triangle\big(0, r(1-2\eta)\big)$ de vergelijking $\zeta + \psi(\zeta) = z$ een unieke oplossing heeft voor $\zeta \in \triangle\big(0, r(1-\eta)\big)$. We construeren ζ als $\lim\limits_{n\to\infty} \zeta_n$ met $\zeta_0 = 0$ en

$$\zeta_{n+1} + \psi(\zeta_n) = z$$

Bovenstaande definitie geeft ons een rij ζ_n van analytische functies van z die, wegens $|\hat{\psi}'| < \eta$, voldoet aan

$$|\zeta_{n+1} - \zeta_n| = |\psi(\zeta_n) - \psi(\zeta_{n-1})| \leq \eta|\zeta_n - \zeta_{n-1}|.$$

Dus als $\zeta_i < r(1-\eta)$ voor $i = 0, 1, \ldots, n$

$$|\zeta_{n+1} - \zeta_n| \leq \eta^n |\zeta_1 - \zeta_0| = \eta^n |z|.$$

Voor $|z| < r(1-2\eta)$ hebben we dus

$$|\zeta_{n+1}| \leq \sum_{j=1}^{n+1} |\zeta_j - \zeta_{j-1}| < \frac{|z|}{1-\eta} < \frac{1-2\eta}{1-\eta} r < (1-\eta)r.$$

Dit laat zien dat de functies ζ_n gedefiniëerd en analytisch zijn voor $|z| < (1-2\eta)r$. Omdat $\eta < 1$ convergeert de rij ζ_n voor $|z| < (1-2\eta)r$ naar $\zeta = \zeta(z)$, de gewenste inverse functie van $\zeta + \psi(\zeta)$.

Bekijken we nu $g = \psi^{-1} \circ f \circ \psi$, dan zien we dat ψ de verzameling $\triangle\big(0, r(1-4\eta)\big)$ afbeeldt op $\triangle\big(0, r(1-3\eta)\big)$. De functie f beeldt vervolgens de verzameling $\triangle\big(0, r(1-3\eta)\big)$ af op $\triangle\big(0, r(1-2\eta)\big)$ en tenslotte beeldt ψ^{-1} de verzameling $\triangle\big(0, r(1-2\eta)\big)$ af op $\triangle\big(0, r(1-\eta)\big)$. Dus g beeldt $\triangle\big(0, r(1-4\eta)\big)$ op $\triangle\big(0, r(1-\eta)\big)$.

We gaan nu $\hat{g}$ schatten. We weten

$$\begin{aligned}
\psi^{-1} \circ f \circ \psi(z) &= g(z) = \lambda z + \hat{g}(z) \\
f \circ \psi(z) &= \psi\big(\lambda z + \hat{g}(z)\big) \\
f\big(\psi(z)\big) &= \lambda z + \hat{g}(z) + \hat{\psi}\big(\lambda z + \hat{g}(z)\big) \\
f\big(z + \hat{\psi}(z)\big) &= \lambda z + \hat{g}(z) + \hat{\psi}\big(\lambda z + \hat{g}(z)\big) \\
\lambda z + \lambda\hat{\psi}(z) + \hat{f}\big(z + \hat{\psi}(z)\big) &= \lambda z + \hat{g}(z) + \hat{\psi}\big(\lambda z + \hat{g}(z)\big) \\
\hat{g}(z) &= \lambda\hat{\psi}(z) + \hat{f}\big(z + \hat{\psi}(z)\big) - \hat{\psi}\big(\lambda z + \hat{g}(z)\big) \\
\hat{g}(z) &= \hat{\psi}(\lambda z) - \hat{\psi}\big(\lambda z + \hat{g}(z)\big) + \hat{f}\big(z + \hat{\psi}(z)\big) - \hat{f}(z)
\end{aligned}$$

We willen $\hat{g}$ schatten in $\triangle\big(0, r(1-4\eta)\big)$. Laat C het maximum van $|\hat{g}|$ zijn over deze schijf, dan

$$\begin{aligned}
C &\leq \sup \Big|\hat{\psi}(\lambda z) - \hat{\psi}\big(\lambda z + \hat{g}(z)\big) + \hat{f}\big(z + \hat{\psi}(z)\big) - \hat{f}(z)\Big| \\
C &\leq \sup \Big|\hat{\psi}(\lambda z) - \hat{\psi}\big(\lambda z + \hat{g}(z)\big)\Big| + \sup \Big|\hat{f}\big(z + \hat{\psi}(z)\big) - \hat{f}(z)\Big|.
\end{aligned}$$

Bekijken we nu $\Big|\hat{\psi}(\lambda z) - \hat{\psi}\big(\lambda z + \hat{g}(z)\big)\Big|$, dan zien we met **Lemma 3.2.4**

$$\begin{aligned}
|\hat{\psi}(\lambda z) - \hat{\psi}\big(\lambda z + \hat{g}(z)\big)| &\leq \sup(|\hat{\psi}'(z)|) \cdot |\lambda z - \big(\lambda z - \hat{g}(z)\big)| \\
&= \sup(|\hat{\psi}'(z)|) \cdot |\hat{g}(z)| \leq \eta C.
\end{aligned}$$

De laatste ongelijkheid geldt, vanwege (3.3) en $|\hat{g}(z)| \leq C|$.

Op een zelfde manier zien we dat voor $\left|\hat{f}\big(z + \hat{\psi}(z)\big) - \hat{f}(z)\right|$ geldt:

$$\begin{aligned} |\hat{f}\big(z + \hat{\psi}(z)\big) - f(z)| &\leq \sup(|f'(c)|) \cdot |\hat{\psi}(z)| \\ &\leq \delta \frac{c_0 \delta}{\eta^{\mu+1}} r \\ &= \frac{c_0 \delta^2 r}{\eta^{\mu+1}}. \end{aligned}$$

Dit geeft ons

$$C \leq \eta C + \frac{c_0 \delta^2 r}{\eta^{\mu+1}} \quad \text{ofwel} \quad C \leq \frac{c_0 \delta^2 r}{\eta^{\mu+1}} \cdot \frac{1}{1-\eta}.$$

Met behulp van de Cauchy afschatting, zie **Lemma 3.2.1** vinden we

$$|\hat{g}'| \leq \frac{c_0 \delta^2 r}{\eta^{\mu+2}} \frac{1}{1-\eta}, \quad z \in \triangle\big(0, r(1-5\eta)\big).$$

Laten we even kort samenvatten wat we hebben gedaan. We zijn begonnen met een functie f die voldoet aan $|\hat{f}'| \leq \delta$ op $\triangle(0,r)$. Deze hebben we vervangen door een functie g met $|\hat{g}'| \leq c_0 \delta^2 r \eta^{-(\mu+2}(1-\eta)^{-1}$ op $\triangle\big(0, r(1-5\eta)\big)$. Hiervoor moesten we aannemen dat

$$0 < \eta < 1/5, \quad c_0 \delta < \eta^{\mu} + 2, \quad \delta < \eta.$$

Als we $c_1 > 0$ klein genoeg nemen en $\eta < c_1$ eisen, voldoen we aan de eerste voorwaarde en de laatste voorwaarde volgt dan uit de tweede.

Stel nu dat we n_0, δ_0 hebben gekozen die voldoen aan de voorwaarde. We definiëren de volgende rijen:

$$\begin{aligned} r_{n+1} &= r_n(1-5\eta_n), \\ \eta_{n+1} &= \eta_n/2, \\ \delta_{n+1} &= c_0\delta_n^2\eta_n 2^{-(\mu+2)}. \end{aligned}$$

De vereiste voorwaarde $c_0\delta_n < \eta_n^{\mu+2}$ is nu gemakkelijk te controleren met inductie. We hebben η_0, δ_0 zo gekozen dat aan de voorwaarde is voldaan. Stel nu dat ook voor n geldt, dan

$$c_0\delta_{n+1} = \frac{c_0^2\delta_n^2}{\eta_n^{\mu+2}2^{\mu+2}} \leq \frac{\eta_n^{2\mu+4}}{\eta_n^{\mu+2}2^{\mu+2}} = \eta_{n+1}^{\mu+2}.$$

Dus het geldt ook voor $n+1$ en daarmee voor alle $n \in \mathbb{N}$. We hebben ook rijen $\{\psi_n\}$ en $\{g_n\}$ geconstruëerd met $g_0 = f$ en $g_n = \psi_n^{-1} \circ g_{n-1} \circ \psi_n$, ofwel

$$g_n = \psi_n^{-1} \circ \psi_{n-1}^{-1} \circ \ldots \circ \psi_1^{-1} \circ f \circ \psi_1 \circ \psi_2 \ldots \circ \psi_n.$$

Laat $R = r_0 \prod_n (1-5\eta) > 0$, dan $|\hat{g}_n'| \leq \delta_n r_n/(1-\eta_n) \to 0$ op $\triangle(0,R)$, dus $g_n \to \lambda z$ op de schijf. En $\{\psi_1 \circ \psi_2 \ldots \circ \psi_n\}$ convergeert dus naar een afbeelding h die f conjugeert aan λz, zoals gewenst. □

3.3 Alternatief bewijs

We bekijken opnieuw **Stelling 3.2.5**. Voor de stelling bestaan meerdere bewijzen, waarvan er hier nog een besproken wordt. Hiervoor moeten we wel nog een definitie invoeren.

Definitie 3.3.1. Een functie $f = \sum_{r=0}^{\infty} a_r z^r$ heet *genormaliseerd* als $|a_r| \leq 1 \ \forall\ r \in \mathbb{N}$.

Lemma 3.3.2. *Een holomorfe functie f met vast punt z_0 en multiplier λ kan* *worden geconjugeerd aan zijn genormaliseerde vorm, waarbij de multiplier gelijk blijft.*

Bewijs. Laat $f = \sum_{r=0}^{\infty} a_r z^r$, omdat f holomorf is, is deze machtreeks convergent. Er bestaat dus een kleinste $c > 0$ zodat $|a_r| \leq c^{r-1}$. We definiëren $g(z) = cf(z/c)$, dus de conjugerende functie is $\varphi(z) = cz$. We zien

$$\begin{aligned} g(z) &= c\sum_{r=0}^{\infty} a_r \left(\frac{z}{c}\right)^r \\ &= \sum_{r=0}^{\infty} ca_r \frac{z^r}{c^r} = \sum_{r=0}^{\infty} \frac{a_r}{c^{r-1}} z^r, \end{aligned}$$

dus $\sum_{r=0}^{\infty} b_r z^r$ met $|b_r| = \left|\frac{a_r}{c^{r-1}}\right| \leq 1$ en $b_1 = \frac{a_1}{c^0} = \frac{\lambda}{1} = \lambda$. Dus we kunnen f conjugeren aan zijn genormaliseerde vorm waarbij de multiplier gelijk blijft. □

Lemma 3.3.3. *Laat $\lambda = e^2\pi i\theta$ met θ Diophantien van orde $\mu+1$. Dan geldt*

$$|\lambda^n - 1|^{-1} < (2n)^{\mu}.$$

Bewijs. We nemen aan dat $\log(|\lambda^n - 1|)$, ofwel er bestaat een constante $v > 0$ zodat voor n groot genoeg geldt

$$|\log(|\lambda^n - 1|)| \leq v\log(n).$$

Merk op dat $\log(|\lambda^n - 1|)$ hooguit $\log(2)$ wordt en dat het dus een grens is voor $-\log(|\lambda^n - 1|)$ als $\lambda \approx 1$. We zien dus

$$-\log(|\lambda^n - 1|) \leq v\log(n).$$

Nemen we nu aan beide kanten de exponent, dan krijgen we

$$|\lambda^n - 1|^{-1} \leq n^v$$

voor n groot genoeg, zeg $n > N$. We willen dat het geldt voor alle n. Dit kan door K de grootste van één en $\max_{n \leq N} \{(n^v|\lambda^n - 1|)^{-1}\}$ te kiezen. Dan krijgen we $|\lambda^n - 1|^{-1} \leq Kn^v$. Nemen we vervolgens $\mu = v + \log_2(K)$, dan krijgen we

$$\lambda^n - 1|^{-1} < (2n)^{\mu}. \tag{3.4}$$

Uit **Lemma 3.1.2** weten we al dat $|\lambda^n - 1|^{-1} \leq Kn^v$ equivalent is aan θ is Diophantien van orde $\mu + 1$. □

Stelling 3.3.4. *Laat θ Diophantien zijn van orde $\mu + 1$ en f een analytische functie zijn met vast punt in $z_0 = 0$ en multiplier $\lambda = e^{2\pi i\theta}$. Dan bestaat er een oplossing voor $\varphi\big(f(z)\big) = \lambda\varphi(z)$, ofwel f kan in een omgeving van 0 geconjugeerd worden aan een vermenigvuldiging met λ.*

Bewijs. We zoeken een oplossing voor

$$\psi\big(f(z)\big) = \lambda\psi(z). \tag{3.5}$$

Nemen we $\varphi = \psi^{-1}$, dan krijgen we $\varphi(\lambda z) = f\big(\varphi(z)\big)$. Laat $f(z) = \sum_{r=1}^{\infty} a_r z^r$ met $a_1 = \lambda$ en $\varphi(z) = \sum_{k=1}^{\infty} c_k z^k$ met $c_1 = 1$, omdat we mogen aannemen dat φ genormaliseerd is. Vergelijking (3.5) wordt dan:

$$\begin{aligned}
\sum_{k=1}^{\infty} c_k(\lambda z)^k &= \sum_{r=1}^{\infty} a_r \left(\sum_{j=1}^{\infty} c_j z^j \right)^r \\
\lambda z + \sum_{k=2}^{\infty} c_k(\lambda z)^k &= \lambda \sum_{j=1}^{\infty} c_j z^j + \sum_{r=2}^{\infty} a_r \left(\sum_{j=1}^{\infty} c_j z^j \right)^r \\
&= \lambda z + \lambda \sum_{j=2}^{\infty} c_j z^j + \sum_{r=2}^{\infty} a_r \left(\sum_{j=1}^{\infty} c_j z^j \right)^r \\
\sum_{k=2}^{\infty} c_k \lambda^k z^k - \lambda \sum_{j=2}^{\infty} c_j z^j &= \sum_{r=2}^{\infty} a_r \left(\sum_{j=1}^{\infty} c_j z^j \right)^r \\
\sum_{k=2}^{\infty} c_k(\lambda^k - \lambda) z^k &= \sum_{r=2}^{\infty} a_r \left(\sum_{j=1}^{\infty} c_j z^j \right)^r . \qquad (3.6)
\end{aligned}$$

Omdat we hebben aangenomen dat $c_1 = 1$, geeft vergelijking (3.6) nu een expliciete recursieve beschrijving van $\{c_k\}$. De termen $c_k(\lambda^k - \lambda)z^k$ zijn namelijk de som van alle z^k-eentermen aan de rechterkant. De uitdrukking aan de rechterkant produceert z^k-eentermen precies als $2 \leq r \leq k$. Deze eentermen hebben de vorm $a_r \cdot (c_{j_1} z^{j_1}) \cdot \ldots \cdot (c_{j_r} z^{j_r})$, waarbij de machten van z optellen tot k. Dus voor $k \geq 2$ krijgen we

$$c_k = \left(\frac{1}{\lambda^k - \lambda} \right) \left(\sum_{r=2}^{\infty} \sum_{j_1 + \ldots + j_r = k} a_r \cdot c_{j_1} \ldots \cdot c_{l_r} \right). \qquad (3.7)$$

Dit lijkt een oplossing te geven voor het probleem, maar afhankelijk van de waarde van λ kan de absolute waarde van

$$d_k := \frac{1}{|\lambda^{k+1} - \lambda|} = \frac{1}{|\lambda^k - 1|}$$

ontzettend vaak ontzettend groot worden. Hierdoor is de machtreeks niet altijd convergent. We noemen d_k een SD-term, waarbij SD staat voor "Small

Denominator", ofwel "kleine noemer".

We gaan nu de waarden van de rij $\{c_k\}$ afschatten. Merk eerst op dat de absolute waarden $|c_k|$ begrensd worden door de reële rij $\{\widehat{c_k}\}$, gedefinieerd door $\widehat{c_1} = 1$ en

$$\widehat{c_k} = d_{k-1}\left(\sum_{r=2}^{\infty} \sum_{j_1+\ldots+j_r=k} \widehat{c_{j_1}} \cdot \ldots \cdot \widehat{c_{l_r}}\right). \tag{3.8}$$

Zie ook vergelijking (3.7) en merk op dat we aan mogen nemen dat $|a_r| \leq 1$ wegens **Lemma 3.3.2**. De recursieve definitie verhult de structuur van deze getallen enigszins. De eerste vier zien er als volgt uit:

$$\begin{aligned}
\widehat{c_1} &= 1, \\
\widehat{c_2} &= d_1([\widehat{c_1}\widehat{c_1}]) = d_1, \\
\widehat{c_3} &= d_2([\widehat{c_1}\widehat{c_2} + \widehat{c_2}\widehat{c_1}] + [\widehat{c_1}\widehat{c_1}\widehat{c_1}]) = 2d_2d_1 + d_2, \\
\widehat{c_4} &= d_3([\widehat{c_1}\widehat{c_3} + \widehat{c_2}\widehat{c_2} + \widehat{c_3}\widehat{c_1}] + [\widehat{c_1}\widehat{c_1}\widehat{c_2} + \widehat{c_1}\widehat{c_2}\widehat{c_1} + \widehat{c_2}\widehat{c_1}\widehat{c_1}] + [\widehat{c_1}\widehat{c_1}\widehat{c_1}\widehat{c_1}]) \\
&= 4d_3d_2d_1 + 2d_3d_2 + d_3d_1d_1 + 3d_3d_1 + d_3.
\end{aligned}$$

De precieze samenstelling van iedere $\widehat{c_k}$ is niet van belang, maar het is duidelijk dat deze de som is van meerdere uitdrukkingen. Ieder van deze uitdrukkingen is het product van, niet per se verschillende, SD-termen.

Laat nu τ_k het aantal termen in de uitdrukking voor $\widehat{c_k}$ zijn, dus $\tau_4 = 4+2+1+3+1 = 11$. Laat verder σ_k het de maximale waarde van de termen in de uitdrukking voor $\widehat{c_k}$ zijn. De waarde hiervan hangt af van λ, maar merk op dat (als we d_k buiten beschouwing laten) het grootste product van SD-termen in de uitdrukking voor $\widehat{c_k}$ een product is van $\widehat{c_{j_1}} \cdot \ldots \cdot \widehat{c_{j_r}}$. Het moet dus het product zijn van de grootste producten in iedere $\widehat{c_{j_1}}, \ldots, \widehat{c_{j_r}}$ of wel een product van $\sigma_{j_1}, \ldots, \sigma_{j_r}$. Dus σ_k is gegeven door $\sigma_1 = 1$ en

$$\sigma_k = d_{k-1} \cdot \max_{\substack{j_1+\ldots+j_r=k \\ 2\leq r\leq k}} \{\sigma_{j_1} \cdot \ldots \cdot \sigma_{j_r}\}.$$

Het is duidelijk dat $|c_k| \leq \widehat{c_k} \leq \sigma_k \tau_k$. We gaan nu een τ_k en σ_k afschatten, waarna het bewijs compleet is.
De getallen τ_k staan bekend als de *Schröder numbers* en in dit geval de *Small Schröder numbers*. De genererende functie $y(x) = \sum \tau_l x^l$ voldoet aan de functionaal vergelijking

$$y = x + \sum_{r=2}^{\infty} y^r.$$

Dit is eenvoudig te verifiëren:

$$\begin{aligned} y &= x + \sum_{r=2}^{\infty} y^r \qquad \text{en we vullen in } y = \sum \tau_l x^l \\ y &= x + \sum_{r=2}^{\infty} \left(\sum \tau_l x^l \right)^r = \sum_{i=1}^{\infty} a_i x^i. \end{aligned}$$

Als we hierin a_m gaan bepalen, moet we eerst kijken wanneer er een term x^m ontstaat in $\sum_{r=2}^{\infty} (\sum \tau_l x^l)^r$. Dit is precies als $2 \leq r \leq m$ en deze termen zijn van de vorm $\tau_{l_1} \cdot \ldots \cdot \tau_{l_r}$ met $l_1 + \ldots + l_r = m$, zodat de machten optellen tot m. Dus we zien

$$a_m = \sum_{r=2}^{\infty} \sum_{l_1 + \ldots + l_r = m} \tau_{l_1} \cdot \ldots \cdot \tau_{l_r}.$$

En als we dit vergelijken met (3.8) zien we dat dit overeenkomt met de definitie van τ_m.

De functionaal vergelijking kunnen we omschrijven naar een meer inzichtelijke functie:

$$\begin{aligned}
y &= x + \sum_{r=2}^{\infty} y^r \\
&= x + \sum_{r=2}^{\infty} y^2 \cdot y^{r-2} \\
&= x + y^2 \sum_{r=0}^{\infty} y^r = x + \frac{y^2}{1-y} \\
y - y^2 &= x - xy + y^2 \\
2y^2 - xy - y &= -x \\
16y^2 - 8xy - 8y &= -8x \\
16y^2 - 8xy - 8y + 2x^2 + 2x + 1 &= -8x + x^2 + 2x + 1 \\
(4y - x - 1)^2 &= (\sqrt{x^2 - 6x + 1})^2 \\
y_1 = \frac{1 + x + \sqrt{x^2 - 6x + 1}}{4} &\vee y_2 = \frac{1 + x - \sqrt{x^2 - 6x + 1}}{4}.
\end{aligned}$$

Het blijkt dat je y_2 de juiste functie is. Merk op dat τ_k natuurlijke getallen zijn voor alle k. Dit betekent dat alle coëfficiënten van de machtreeksontwikkeling van y, rond 0, positief en geheel moeten zijn. Hiertoe bekijken we eerst de Taylorontwikkeling van $\sqrt{x^2 - 6x + 1}$ in 0.

We vinden $\sqrt{x^2 - 6x + 1} = 1 - 3x - 4x^2 - 12x^3 - 44x^4 + \ldots$. Hieruit volgt

$$\begin{aligned}
y_1 &= \frac{1}{4}(1 + x + 1 - 3x - 4x^2 - 12x^3 - 44x^4 + \ldots) \\
&= \frac{1}{4}(2 - 2x - 4x^2 - 12x^3 - 44x^4 + \ldots)
\end{aligned}$$

en

$$\begin{aligned}
y_2 &= \frac{1}{4}\big(1 + x - (1 - 3x - 4x^2 - 12x^3 - 44x^4 + \ldots)\big) \\
&= \frac{1}{4}(4x + 4x^2 + 12x^3 + 44x^4 + \ldots).
\end{aligned}$$

Het is duidelijk dat de machtreeks van y_1 negatieve (en gebroken) coëfficiënten geeft, terwijl de machtreeks van y_2 alleen natuurlijke getallen oplevert en dus de juiste genererende functie is.

De convergentiestraal van deze functie is de absolute waarde van het kleinste nulpunt van $x^2 - 6x + 1$, omdat hier een singulariteit optreed. De nulpunten zijn $3 - \sqrt{8}$ en $3 + \sqrt{8}$, dus de convergentiestraal $R = 3 - \sqrt{8}$. We weten dat

$$\begin{aligned} \limsup \tau_n^{1/n} &= R^{-1} \\ \text{dus} \qquad \limsup \tau_n^{1/n} &= \frac{1}{3 - \sqrt{8}} = 3 + \sqrt{8} \\ \tau_n &\leq (3 + \sqrt{8})^n. \end{aligned}$$

Hiermee hebben we τ_k dus afgeschat.

Nu moeten we allen nog σ_k afschatten, maar hier is nog wat werk voor nodig. We hadden al gezien dat σ_k recursief gedefiniëerd wordt door

$$\sigma_k = d_{k-1} \cdot \max_{\substack{j_1 + \ldots + j_r = k \\ 2 \leq r \leq k}} \{\sigma_{j_1} \cdot \ldots \cdot \sigma_{j_r}\}$$

en een exponentiële bovengrens heeft als de SD-termen voldoen aan **Lemma 3.3.3**. Hierbij maken we gebruik van de *simpele afschatting* $d_k \leq (2k)^\mu$. Dit is echter geen efficiënte grens, omdat iedere d_k een product is van $\mathcal{O}(k)$ SD-termen. Siegel zag in dat, wanneer een SD-term groot wordt, het meerdere stappen duurt voordat een SD-term weer groot wordt. Hierdoor kunnen we een betere afschatting voor d_k vinden, die we *subtiele afschatting* zullen noemen. Bekijk

$$\begin{aligned} \lambda^q(\lambda^{p-q} - 1) &= \lambda^p - \lambda^q \\ &= (\lambda^p - 1) - (\lambda^q - 1), \end{aligned}$$

omdat $|\lambda^q| = 1$, volgt met de driehoeksongelijkheid dat

$$|\lambda^{p-q} - 1| \leq |\lambda^p - 1| + |\lambda^q - 1|$$

in SD-notatie $$d_{p-q}^{-1} \leq d_p^{-1} + d_q^{-1} \leq 2(\min\{d_p, d_q\})^{-1}.$$

Dus we zien $2(\min\{d_p, d_q\}) \leq d_{p-q}$ en als we hier de simpele afschatting op toepassen, krijgen we

$$\min\{d_p, d_q\} \leq 2^{\mu+1}(p-q)^{\mu}.$$

Deze afschatting is duidelijk efficiënter dan $\min\{d_p, d_q\} \leq \min\{(2p)^{\mu}, (2q)^{\mu}\}$.

Met behulp van **Lemma 3.3.5** (Deze is in verband met gebruikte notatie onder dit bewijs te vinden) vinden we hiermee een bovengrens voor het product $\prod_{p=0}^{r} d_{k_p}$. Namelijk $\prod_{p=0}^{r} d_{k_p} < N^{r+1} \cdot k_0^{\mu} \prod_{p=1}^{r} (k_{p-1} - k_p)^{\mu}$, waarbij de $r+1$ indices als volgt geordend zijn $k_0 > \ldots > k_r \geq 1$ en $N = 2^{2\mu+1}$ met v de orde van het Diophantine getal θ.

Hiermee kunnen we een exponentiële bovengrens voor σ_k vinden. We beginnen met de afschatting $\sigma_k \leq \dfrac{AC^k}{k^B}$ en gaan A, B en C zoeken die hieraan voldoen. Merk op dat $\sigma_1 = d_0 = 1 \leq AC$, waardoor de keus $A = C^{-1}$ logisch lijkt. Dit betekent dat we op willen gaan lossen

$$\sigma_k \leq \frac{C^{-1}C^k}{k^B} = \frac{C^{k-1}}{k^B} \tag{3.9}$$

voor alle k. We kunnen een oplossing hiervoor altijd aanpassen zodat

$$B \overset{(a)}{>} 0 \text{ en } C \overset{(b)}{\geq} 2^B$$

door C groter te kiezen. Deze extra condities hebben het voordeel dat

$$\begin{aligned}
\sigma_{j_1} \cdot \sigma_{j_2} &\leq \frac{C^{j_1-1}}{j_1^B} \frac{C^{j_2-1}}{j_2^B} = \frac{C^{j_1+j_2-2}}{j_1^B \cdot j_2^B} \\
&= \frac{C^{j_1+j_2-2}}{(j_1 \cdot j_2)^B} \cdot \frac{(j_1+j_2)^B}{(j_1+j_2)^B} \\
&= \frac{C^{j_1+j_2-2}}{(j_1+j_2)^B} \cdot \frac{(j_1+j_2)^B}{(j_1 \cdot j_2)^B} \\
&= \frac{C^{j_1+j_2-2}}{(j_1+j_2)^B} \cdot \left(\frac{j_1}{j_1 \cdot j_2} + \frac{j_2}{j_1 \cdot j_2}\right)^B \\
&= \frac{C^{j_1+j_2-1}}{(j_1+j_2)^B} \cdot C^{-1} \cdot \left(\frac{1}{j_2} + \frac{1}{j_1}\right)^B \\
&\overset{(a)}{\leq} \frac{C^{j_1+j_2-1}}{(j_1+j_2)^B} \cdot C^{-1} 2^B \\
&\overset{(b)}{\leq} \frac{C^{j_1+j_2-1}}{(j_1+j_2)^B}.
\end{aligned}$$

En in het algemeen voor $j_1 + \ldots + j_l = J$

$$\sigma_{j_1} \cdot \ldots \cdot \sigma_{j_l} \leq \frac{C^{J-1}}{J^B}. \tag{3.10}$$

Stel nu dat we B en C hebben die voldoen aan (a), (b) en $\sigma_k \frac{C^{k-1}}{k^B}$ voor alle $1 \leq k < k_0$. Dan kunnen we σ_{k_0} afschatten met het volgende argument:

In de ontbinding van σ_{k_0}

$$\sigma_{k_0} = d_{k_0-1} \cdot \max_{\substack{j_1+\ldots+j_r=k_0 \\ 2 \leq r \leq k_0}} \{\sigma_{j_1} \cdot \ldots \cdot \sigma_{j_r}\}$$

zien we dat er hoogstens een index k_1 kan zijn met $k_0/2 < k_1 < k_0$. Als dat zo is, schrijven we $\sigma_{k_0} = d_{k_0-1} \cdot \sigma_{k_1} \cdot \Delta_1$ waarbij Δ_1 het product is van de resterende σ_j. Als er geen index groter is dan $k_0/2$ zijn we klaar met deze stap.

Merk op dat er in de ontbinding van k_1 nog steeds een index k_2 kan zijn met $k_0/2 < k_2 < k_1 < k_0$. Als dit zo is schrijven we $\sigma_{k_1} = d_{k_1-1} \cdot \sigma_{k_2} \cdot \Delta_2$. Dit kunnen we herhalen tot we een σ_{k_r} vinden met een ontbinding waarin geen indices voorkomen groter dan $k_0/2$. We vinden dan het volgende:

$$\begin{aligned} \sigma_{k_0} &= d_{k_0-1} \cdot \sigma_{k_1} \cdot \Delta_1 \\ \sigma_{k_1} &= d_{k_1-1} \cdot \sigma_{k_2} \cdot \Delta_2 \\ &\vdots \\ \sigma_{k_{r-1}} &= d_{k_{r-1}-1} \cdot \sigma_{k_r} \cdot \Delta_r \\ \sigma_{k_r} &= d_{k_{r-1}} \cdot \sigma_{l_1} \cdot \ldots \cdot \sigma_{l_s} \end{aligned}$$

Merk op dat de Δ_p hierbij een product is van ongenoemde σ_j waarvan de indices allemaal kleiner zijn dan k_0 en optellen tot $k_{p-1} - k_p$. De indices l_q tellen op tot k_r en zijn allemaal hoogstens $k_0/2 < k_0$. Deze condities zijn nodig om **Lemma 3.3.6** (die hierna bewezen zal worden) later toe te kunnen passen.

Eerst schrijven we de bovenstaande ontbindingen als een product, door te substituëren. Dit geeft ons

$$\sigma_{k_0} = \prod_{p=0}^{r} d_{k_p-1} \cdot \prod_{p=1}^{r} \Delta_p \cdot (\sigma_{l_1} \cdot \ldots \cdot \sigma_{l_s}).$$

Nu kunnen we **Lemma 3.3.5** toepassen op het eerste product, ongelijkheid (3.10) op ieder Δ_p en (3.9) op iedere σ_{l_q}. Dan zien we dat

$$\sigma_{k_0} \leq \left[N^{r+1} \cdot k_0^{\mu} \prod_{p=1}^{r} (k_{p-1} - k_p)^{\mu} \right] \cdot \left[\prod_{p=1}^{r} \frac{C^{(k_{p-1}-k_p)-1}}{(k_{p-1} - k_p)^B} \right] \cdot \left[\prod_{q=1}^{s} \frac{C^{l_q - 1}}{l_q^B} \right].$$

Bij het tweede product zien we dat de teller

$$C^{k_0-k_1+k_1-k_2+\ldots+k_{r-1}-k_r-r} = C^{k_0-k_r-r}$$

wordt en voor het laatste product krijgen we de teller

$$C^{l_1+\ldots+l_s-s} = C^{k_r-s}.$$

Verder kunnen de eerste twee producten eenvoudig samengenomen worden. Met het bovenstaande krijgen we dan

$$\sigma_{k_0} \leq (N^{r+1} \cdot C^{k_0-k_r-r+k_r-s}) \cdot k_0^{\mu} \cdot \prod_{p=1}^{r} \frac{(k_{p-1}-k_p)^{\mu}}{(k_{p-1}-k_p)^{B}} \cdot \prod_{q=1}^{s} \frac{1}{l_q^B}$$

$$\sigma_{k_0} \leq (N^{r+1} \cdot C^{k_0-r-s}) \cdot \Big(k_0^{\mu} \prod_{p=1}^{r} (k_{p-1}-k_p)^{\mu-B} \cdot \prod_{q=1}^{s} l_q^{-B}\Big). \qquad (3.11)$$

Herinner dat deze afschatting nog afhankelijk is van het vinden van een geschikte B, C die voldoen aan $B > 0, C \geq 2^B$ en $\sigma_k \leq \frac{c^{k-1}}{k^B}$ voor $1 \leq k < k_0$. We willen nu B en C zo kiezen, zodat de rechterkant van bovenstaande vergelijking begrensd is door $\dfrac{C^{k_0-1}}{k_0^B}$. Hiertoe moeten we een manier vinden om een factor $\dfrac{1}{k_0^B}$ uit

$$\Big(k_0^{\mu} \prod_{p=1}^{r} (k_{p-1}-k_p)^{\mu-B} \cdot \prod_{q=1}^{s} l_q^{-B}\Big)$$

te halen. Hiervoor gebruiken we de volgende hulpvergelijking

$$\Big(\prod_{p=1}^{r} (k_{p-1}-k_p)^{\mu-B} \cdot \prod_{q=1}^{s} l_q^{-B}\Big) \leq \frac{1}{k_0^{\mu+B}} \cdot \varphi = k_0^{-v-B} \cdot \varphi, \qquad (3.12)$$

waarbij φ van r en s mag afhangen, maar niet van k_0.

We kiezen nu

$$B \overset{(c)}{=} 2v.$$

Merk op dat dit geen conflict oplevert met (a). Het blijkt dat bovenstaande methode volstaat en we dus B gelijk kunnen stellen aan $2v$. Samen met de ongelijkheid uit **Lemma 3.3.6**, die we kunnen toepassen door de eerder gestelde restricties op de indices, vinden we:

$$\left(\prod_{p=1}^{r}(k_{p-1}-k_p)^{-v}\cdot\prod_{q=1}^{s}l_q^{-2v}\right)\leq\left(\frac{k_0}{2^{r+s+1}}\right)^{-3v}.$$

Hiermee kunnen we de afschatting die we hebben gevonden in (3.11) concreter maken, waarna we alleen nog een geschikte waarde voor C hoeven te kiezen. Als we bovenstaande ongelijkheid invullen zien we namelijk

$$\begin{aligned}
\sigma_{k_0} &\leq (N^{r+1}\cdot C^{k_0-r-s})\cdot\left(k_0^{\mu}\cdot\left(\frac{k_0}{2^{r+s+1}}\right)^{-3v}\right)\\
&= \left((2^{2\mu+1})^{r+1}\cdot C^{k_0-1}\cdot C^{-r-s+1}\right)\cdot\left(k_0^{\mu}\cdot\left(\frac{2^{r+s-1}}{k_0}\right)^{3\mu}\right)\\
&= \left((2^{2\mu+1})^{r+1}\cdot(2^{3\mu})^{r+s-1}\cdot\frac{1}{C^{r+s-1}}\right)\cdot\frac{C^{k_0-1}}{k_0^{2\mu}}\\
&\leq \left(\frac{2^{2\mu+1}\cdot 2^{3\mu}}{C}\right)^{r+s-1}\cdot\frac{C^{k_0-1}}{k_0^{2\mu}}\\
&= \left(\frac{2^{5\mu+1}}{C}\right)^{r+s-1}\cdot\frac{C^{k_0-1}}{k_0^{2\mu}}
\end{aligned}$$

De laatste ongelijkheid geldt omdat $s\geq 2$. Als we nu $C\geq 2^{5\mu+1}$ kiezen dan $\frac{2^{5\mu+1}}{C}\leq 1$, dus is de laatste uitdrukking kleiner dan $\frac{C^{k_0-1}}{k_0^{2\mu}}$. Deze keuze voor C levert geen problemen op met (b) en (c), dus als we deze substituëren, zien we

$$\sigma_{k_0}\leq\frac{(2^{5\mu+1})^{k_0-1}}{k_0^{2\mu}}$$

voor alle $k_0 \geq 1$ en hebben we dus succesvol de σ_j afgeschat.

Herinner dat $|c_k| \leq \widehat{c_k} \leq \sigma_k \tau_k$, dus hiermee hebben we $|c_k|$ afgeschat en is de stelling bewezen. $\square$

Lemma 3.3.5. *Laat $N = 2^{2\mu+1}$ met v de orde van het Diophantine getal θ. Voor $r+1$ gegeven indices $k_0 > \ldots > k_r \geq 1$, geldt het volgende:*

$$\prod_{p=0}^{r} d_{k_p} < N^{r+1} \cdot k_0^{\mu} \prod_{p=1}^{r} (k_{p-1} - k_p)^{\mu}. \tag{3.13}$$

Bewijs. Het bewijs gaat met inductie. De simpele afschatting voldoet voor het geval $r = 0$. Voor $r = 1$ zien we:

$$\begin{aligned} d_{k_0} \cdot d_{k_1} &\leq (2^{\mu} \max\{k_0^{\mu}, k_1^{\mu}\}) \cdot (2^{\mu+1}|k_0 - k_1|^{\mu}) \\ &\leq 2^{\mu} k_0^{\mu} \cdot 2^{\mu+1}(k_0 - k_1)^{\mu} \\ &= 2^{2\mu+1} k_0^{\mu} \cdot (k_0 - k_1)v \\ &< N^2 \cdot k_0^{\mu}(k_0 - k_1)^{\mu}. \end{aligned}$$

Zij nu $r \geq 2$, dus we hebben minstens drie SD-termen. Laat d_{k_j} de kleinste term zijn. Met de inductiehypothese volgt dat de overige r SD-termen voldoen aan vergelijking (3.13), zonder de index k_j. Als $j = 0$, dus $d_{k_j} = d_{k_0}$ zien we $d_{k_0} < 2^{\mu+1}(k_0 - k_1)^{\mu} < N(k_0 - k_1)^{\mu}$. Dus

$$\begin{aligned} d_{k_0} \cdot \prod_{p=1}^{r} d_{k_p} &< d_{k_0} \cdot N^r \cdot k_1^{\mu} \prod_{p=2}^{r} (k_{p-1} - k_p)^{\mu} \\ &< N(k_0 - k_1)^{\mu} \cdot N^r \cdot k_1^{\mu} \prod_{p=2}^{r} (k_{p-1} - k_p)^{\mu} \\ &< N \cdot N^r \cdot k_0^{\mu}(k_0 - k_1)^{\mu} \prod_{p=2}^{r} (k_{p-1} - k_p)^{\mu} \\ &= N^{r+1} \cdot k_0^{\mu} \prod_{p=1}^{r} (k_{p-1} - k_p)^{\mu}. \end{aligned}$$

Een vergelijkbaar bewijs geld voor het geval $j = r$. Zij nu $0 < j < r$.

De volgens de inductiehypothese bestaande bovengrens voor $\prod_{\substack{p=0\\p\neq j}}^{r} d_{k_p}$ bevat de term $(k_{j-1} - k_{j=1})$. Laat a, b zo zijn dat $\{k_a, k_b\} = \{k_{j-1}, k_{j+1}\}$ met $|k_a - k_j| \geq |k_j - k_b|$. Hieruit volgt dat $(k_{j-1} - k_{j+1}) \leq 2|k_a - k_j|$ en met de subtiele afschatting volgt nog $d_{k_j} \leq 2^{\mu+1}|k_j - k_b|^\mu$. Dus

$$\begin{aligned}\prod_{p=0}^{r} d_{k_p} &\leq N^r \cdot k_0(k_0-k_1)^\mu \cdot \ldots \cdot (k_{j-1}-k_{j+1})^\mu \cdot d_{k_j} \cdot \ldots \cdot (k_{r-1}-k_r)^\mu \\ &\leq N^r \cdot k_0(k_0-k_1)^\mu \cdot \ldots \cdot 2^\mu |k_a - k_j|^\mu \cdot 2^{\mu+1}|k_j - k_b|^\mu \cdot \ldots \cdot (k_{r-1}-k_r)^\mu \\ &= N^{r+1} \cdot k_0(k_0-k_1)^\mu \cdot \ldots \cdot (k_{j-1}-k_j)^\mu (k_j - k_{j+1})^\mu \cdot \ldots \cdot (k_{r-1}-k_r)^\mu.\end{aligned}$$

Hiermee is het bewijs geleverd. □

Lemma 3.3.6. *Gegeven drie gehele getallen $k \geq 2, r \geq 0$ en $s \geq 2$. Als we gehele getallen $x_1, \ldots, x_r$ en $y_1, \ldots, y_s$ uit $\{1 \ \ldots, \lfloor \frac{k}{2} \rfloor\}$ hebben die voldoen aan $\sum_{p=1}^{r} x_p + \sum_{q=1}^{s} y_q = k$ met $\sum_{q=1}^{s} y_q > k/2$, dan*

$$\prod_{p=1}^{r} x_p \cdot \prod_{q=1}^{s} y_q^2 \geq \left(\frac{k}{2^{r+s-1}}\right)^3.$$

Bewijs. Zij $t = r + s \geq 2$. Omdat $2t - 2 \leq 2^{t-1}$ is het voldoende om te bewijzen dat

$$\prod_{p=1}^{r} x_p \cdot \prod_{q=1}^{s} y_q^2 \geq \left(\frac{k}{2t-2}\right)^3.$$

Sommige gevallen zijn duidelijk:

- $t = 2$: Als $t = 2$, dan $r = 0$ en $s = 2$, dus $y_1 = y_2 = k/2$. We zien

$$\prod_{p=1}^{r} x_p \cdot \prod_{q=1}^{s} y_q^2 = \frac{k^4}{4} \geq \frac{k^3}{8}.$$

- $k \leq 2t - 2$: Als $k \leq 2t - 2$ is $\frac{k}{2t-2} \leq 1$, dus geldt de ongelijkheid altijd, want het product is minstens één. Het is eenvoudig in te zien dat $k \leq 2t-2$ voor $k = 2, 3, 4$, als $t \geq 3$.

We moeten de ongelijkheid dus nog bewijzen voor $t \geq 3, k \geq 5$ en $k\ 2t-2$, wat equivalent is met $3 \leq t \leq \left\lceil \frac{k}{2} \right\rceil$.

Voor $\prod_{p=1}^{r} x_p$ krijgen we het kleinste product als $r-1$ termen gelijk zijn aan één en de laatste term gelijk is aan $\sum_{p=1}^{r} x_p - (r-1)$. Dit is dan ook de ondergrens van $\prod_{p=1}^{r} x_p$.

Op dezelfde manier vinden we $\sum_{q=1}^{s} y_q - (s-1)$ als ondergrens voor $\prod_{q=1}^{s} y_q$. Hiervoor kunnen we echter een scherpere ondergrens vinden als $\sum_{q=1}^{s} y_q - (s-1) > \left\lfloor \frac{k}{2} \right\rfloor$. Omdat alle $y_q \leq \left\lfloor \frac{k}{2} \right\rfloor$ krijg je in dit geval het kleinste product als $s-2$ factoren gelijk zijn aan één, een term gelijk is aan $\left\lfloor \frac{k}{2} \right\rfloor$ en de laatste term is wat er overblijft, ofwel $\sum_{q=1}^{s} y_q - (s-2) - \left\lfloor \frac{k}{2} \right\rfloor$. We krijgen dus

$$\prod_{p=1}^{r} x_p \geq \sum_{p=1}^{r} x_p - (r-1),$$

$$\prod_{q=1}^{s} y_q \geq \begin{cases} \sum_{q=1}^{s} y_q - (s-1) & \text{if } \sum_{q=1}^{s} y_q - (s-1) \leq \left\lfloor \frac{k}{2} \right\rfloor, \\ \left(\sum_{q=1}^{s} y_q - (s-2) - \left\lfloor \frac{k}{2} \right\rfloor\right) \cdot \left\lfloor \frac{k}{2} \right\rfloor & \text{if } \sum_{q=1}^{s} y_q - (s-1) \geq \left\lfloor \frac{k}{2} \right\rfloor. \end{cases}$$

Merk op dat de waarden van $\sum_{q=1}^{s} y_q$ tussen $(\left\lfloor \frac{k}{2} \right\rfloor + 1$ en $(k-r)$ liggen. We delen de rest van het bewijs nu in twee gevallen op, maar eerst bekijken we nog de polynoom

$$P(x) = (R-x)(x-S)^2$$

met $R > S$. We zien $P'(x) = (2R + S - 3x)(x - S)$, dus $P'(S) = 0$ en S is een kritiek punt. De tweede afgeleide $P''(x) = 2R + 4S - 6x$, dus $P''(S) = 2R - 2S > 0$. Met de tweede afgeleide test volgt dus dat S een lokaal minimum is en aangezien het derdegraads polynoom is met $a_3 = -1$, volgt dat dit het enige lokale minimum is. Als we nu een interval $I = [a, b]$

hebben met $S < a < b$ geldt

$$\min_{x \in I}\{P(x)\} = \min\{P(a), p(b)\}.$$

We gaan nu het eerste geval bekijken. Dit is het geval dat $(\lfloor \frac{k}{2} \rfloor + 1) \leq \sum_{q=1}^{s} y_q \leq (\lfloor \frac{k}{2} \rfloor + s - 1)$, dus $\prod_{q=1}^{s} y_q \geq \sum_{q=1}^{s} y_q - (s-1)$. We krijgen dus

$$\begin{aligned}\prod_{p=1}^{r} x_p \cdot \prod_{q=1}^{s} y_q^2 &\geq \left(\sum_{p=1}^{r} x_p - (r-1)\right) \cdot \left(\sum_{q=1}^{s} y_q - (s-1)\right)^2 \\ &\geq \left(k - \sum_{q=1}^{s} y_q - r + 1\right) \cdot \left(\sum_{q=1}^{s} y_q - s + 1\right)^2\end{aligned}$$

Laat nu $R = k - r + 1$ en $S = s - 1$, dan krijgen we

$$\prod_{p=1}^{r} x_p \cdot \prod_{q=1}^{s} y_q^2 \geq P(\sum_{q=1}^{s} y_q).$$

Heriner dat $\sum_{p=1}^{r} x_p + \sum_{q=1}^{s} y_q = k$, dus $k \geq r + s$. Ofwel $k - r \geq s$, waaruit volgt dat $R > S$. We weten dus dat het product $\prod_{p=1}^{r} x_p \cdot \prod_{q=1}^{s} y_q^2$ van onderen begrensd is door het minimum van P in het bereik van $\sum_{q=1}^{s} y_q$. Dit bereik is omvat in het interval $I = [\big(\lfloor \frac{k}{2} \rfloor - r + 1\big), \big(\lfloor \frac{k}{2} \rfloor + s - 1\big)]$. Uit $t = r + s \leq \lceil \frac{k}{2} \rceil$ zien we $s - 1 < \lceil \frac{k}{2} \rceil - r$, of wel $S = s - 1 < \lfloor \frac{k}{2} \rfloor - r + 1$. Dus het product $\prod_{p=1}^{r} x_p \cdot \prod_{q=1}^{s} y_q^2$ is begrensd door het minimum van

$$P(\left\lfloor \frac{k}{2} \right\rfloor - r + 1) \text{ en } P(\left\lfloor \frac{k}{2} \right\rfloor + s - 1).$$

Herinnerend dat $t = r + s$ zien we dat

$$P(\left\lfloor \frac{k}{2} \right\rfloor - r + 1) = \left\lceil \frac{k}{2} \right\rceil (\left\lfloor \frac{k}{2} \right\rfloor - t + 2)^2$$

en
$$P(\left\lfloor \frac{k}{2} \right\rfloor + s - 1) = (\left\lceil \frac{k}{2} \right\rceil - t + 2) \left\lfloor \frac{k}{2} \right\rfloor^2.$$

En

$$\begin{aligned}
\left\lceil \frac{k}{2} \right\rceil (\left\lfloor \frac{k}{2} \right\rfloor - t + 2)^2 &\leq \left\lceil \frac{k}{2} \right\rceil \cdot (\left\lfloor \frac{k}{2} \right\rfloor - t + 2) \cdot (\left\lceil \frac{k}{2} \right\rceil - t + 2) \\
&\leq (\left\lfloor \frac{k}{2} \right\rfloor + 1) \cdot (\left\lfloor \frac{k}{2} \right\rfloor - 1) \cdot (\left\lceil \frac{k}{2} \right\rceil - t + 2) \\
&\leq (\left\lceil \frac{k}{2} \right\rceil - t + 2) \left\lfloor \frac{k}{2} \right\rfloor^2,
\end{aligned}$$

waarbij we voor de middelste ongelijkheid gebruiken maken van $t \geq 3$. Voor $\sum_{q=1}^{s} y_q \in [\big(\left\lfloor \frac{k}{2} \right\rfloor - r + 1\big), \big(\left\lfloor \frac{k}{2} \right\rfloor + s - 1\big)]$ hebben we

$$\begin{aligned}
\prod_{p=1}^{r} x_p \cdot \prod_{q=1}^{s} y_q^2 &\geq \left\lceil \frac{k}{2} \right\rceil (\left\lfloor \frac{k}{2} \right\rfloor - t + 2)^2 \\
&\geq \frac{k}{2} \cdot \left(\frac{k-1}{2} - t + 2 \right)^2.
\end{aligned}$$

Het is duidelijk dat $\frac{k-1}{2} - t + 2$ een lineaire functie in t is, terwijl $\frac{k-1}{2t-2}$ een convexe functie is. Voor $t \in \big[3, \left\lceil \frac{k}{2} \right\rceil\big]$ is de lineaire functie groter dan $\frac{k-1}{2t-2}$, dus

$$\begin{aligned}
\prod_{p=1}^{r} x_p \cdot \prod_{q=1}^{s} y_q^2 &\geq \frac{k}{2} \cdot \left(\frac{k-1}{2t-2}\right)^2 \\
&= \frac{k^3}{2k^2} \cdot \frac{(k-1)^2}{(2t-2)^2} \\
&= \frac{k^3}{(2t-2)^3} \cdot \frac{(2t-2)(k-1)^2}{2k^2} \\
&= \left(\frac{k}{2t-1}\right)^3 \cdot \left(\frac{k-1}{k}\right)^2 \cdot (t-1) \\
&\geq 2 \cdot \left(\frac{4}{5}\right)^2 \cdot \left(\frac{k}{2t-2}\right)^3 > \left(\frac{k}{2t-2}\right)^3.
\end{aligned}$$

In de laatste stap gebruiken we dan $t \geq 3$ en $k \geq 5$. Hiermee is het eerste geval bewezen. Nu bekijken we het tweede geval: $\left(\left\lfloor \frac{k}{2} \right\rfloor + s - 1 \leq \sum_{q=1}^{s} y_q\right) \leq (k-r)$. Dus $\prod_{q=1}^{s} y_q \geq \left(\sum_{q=1}^{s} y_q - (s-2) - \left\lfloor \frac{k}{2} \right\rfloor\right) \cdot \left\lfloor \frac{k}{2} \right\rfloor$ en hiermee volgt dat

$$\begin{aligned}
\prod_{p=1}^{r} x_p \cdot \prod_{q=1}^{s} y_q^2 &\geq \left(\sum_{p=1}^{r} x_p - (r-1)\right) \cdot \left(\sum_{q=1}^{s} y_q - (s-2) - \left\lfloor \frac{k}{2} \right\rfloor\right)^2 \cdot \left\lfloor \frac{k}{2} \right\rfloor^2 \\
&\geq \left(k - \sum_{q=1}^{s} y_q - r + 1\right) \cdot \left(\sum_{q=1}^{s} y_q - (s-2) - \left\lfloor \frac{k}{2} \right\rfloor\right)^2 \cdot \left\lfloor \frac{k}{2} \right\rfloor^2
\end{aligned}$$

Laat $R = k - r + 1$ en $S = s + \left\lfloor \frac{k}{2} \right\rfloor - 2$, dan $R > S$, want

$$\begin{aligned}
R - S &= k - r + 1 - \left(s + \left\lfloor \frac{k}{2} \right\rfloor - 2\right) \\
&= k - r - s - \left\lfloor \frac{k}{2} \right\rfloor + 3 \\
&= \left\lceil \frac{k}{2} \right\rceil - t + 3 \geq 3.
\end{aligned}$$

Dus nu hebben we

$$\prod_{p=1}^{r} x_p \cdot \prod_{q=1}^{s} y_q^2 \geq P\left(\sum_{q=1}^{s} y_q\right) \cdot \left\lfloor \frac{k}{2} \right\rfloor^2$$

met $R > S$. Het bereik van $\sum_{q=1}^{s} y_s$ ligt in het interval $I = \left[\left(\left\lfloor \frac{k}{2} \right\rfloor + s - 1\right), (k-r)\right]$, wat duidelijk rechts van S ligt. We weten dus dat $\prod_{p=1}^{r} x_p \cdot \prod_{q=1}^{s} y_q^2$ van onderen begrensd wordt door het minimum van

$$P\left(\left\lfloor \frac{k}{2} \right\rfloor + s - 1\right) \cdot \left\lfloor \frac{k}{2} \right\rfloor^2 \text{ en } P(k-r) \cdot \left\lfloor \frac{k}{2} \right\rfloor^2.$$

Herinnerend dat $t = r + s$ zien we dat

$$P\left(\left\lfloor \frac{k}{2} \right\rfloor + s - 1\right) \cdot \left\lfloor \frac{k}{2} \right\rfloor^2 = \left(\left\lceil \frac{k}{2} \right\rceil - t + 2\right) \cdot \left\lfloor \frac{k}{2} \right\rfloor^2$$

en

$$P(k-r) \cdot \left\lfloor \frac{k}{2} \right\rfloor^2 = \left(\left\lceil \frac{k}{2} \right\rceil - t + 2\right)^2 \cdot \left\lfloor \frac{k}{2} \right\rfloor^2.$$

Het is duidelijk dat de bovenste uitkomst de kleinste is, dus

$$\prod_{p=1}^{r} x_p \cdot \prod_{q=1}^{s} y_q^2 \geq \left(\left\lceil \frac{k}{2} \right\rceil - t + 2\right) \cdot \left\lfloor \frac{k}{2} \right\rfloor^2.$$

We zagen in het bewijs van het eerste geval al dat dit groter is dan $\left\lceil \frac{k}{2} \right\rceil \cdot \left(\left\lfloor \frac{k}{2} \right\rfloor - t + 2\right)^2$. Het bewijs gaat vanaf hier dus op dezelfde als het eind van het eerste geval. $\square$

BIBLIOGRAFIE

Lennart. Carleson en Theodore W. Gamelin, *Complex Dynamics*, Springer-Verlag, 1993

Rodrigo A. Pérez, "A Brief but Historic Article of Siegel", *Notices of the AMS*, Volume 58, Number 4, 2011, p. 558-566

Carl L. Siegel en Jürgen K. Moser, *Lectures on Celestial Mechanics*, Spinger-Verlag, 1971

Elias M. Stein en Rami Shakarchi, *II Complex Analysis*, Princeton University Press, 2003