

Wiskunde achter het Spoorboekje

Lex Schrijver

Centrum voor Wiskunde en Informatica (CWI)
en Universiteit van Amsterdam

1 Spoorboekje 2007

In december 2006 werd bij de Nederlandse Spoorwegen het *Spoorboekje 2007* ingevoerd. Hiermee kreeg de dienstregeling van NS een geheel nieuwe structuur, met veel nieuwe verbindingen en overstapmogelijkheden, maar ook met wat ruimere rij- en stoptijden en enkele verbroken rechtstreekse verbindingen.

Tot 2006 werden nieuwe dienstregelingen bij NS gemaakt door het ‘met de hand’ aanpassen van de dienstregeling die in 1970 was ingevoerd (‘Spoorslag 70’). Nieuwe treinen werden tussengevoegd door de tijden van bestaande treinen wat op te schuiven. De dienstregelingplanners bij de NS beseften echter dat zij computeralgoritmen nodig hadden om de capaciteit van het spoorwegnetwerk beter te benutten. Hiervoor riep NS de hulp in van het Centrum voor Wiskunde en Informatica (CWI) in Amsterdam.

Hoewel Nederland geen groot land is, heeft het wel een aantal kenmerken die het uitzonderlijk maken voor spoorwegplanning. Nederland is heel dicht bevolkt, en daardoor is ook het spoorwegnetwerk tamelijk dicht, met veel korte verbindingen, waarop veel treinen rijden die verschillende aansluitingen hebben, onderweg en aan de eindpunten. Ook is er slechts zeer beperkte ruimte voor uitbreiding van de capaciteit van het spoorwegnetwerk, waardoor het nog veel knelpunten kent.

Door deze uitzonderlijke positie was er in ander landen geen algoritmiëk beschikbaar die in Nederland gebruikt kon worden, en moesten nieuwe algoritmen worden ontwikkeld. Deze maken gebruik van methoden uit de discrete wiskunde, en in dit artikel zullen wij daar iets van uitleggen.



2 Het Basis Uur Patroon

De basis van de Nederlandse spoorwegdienstregeling is het *Basis Uur Patroon* (BUP): de dienstregeling herhaalt zich elke 60 minuten. Aan deze basis worden extra spitsreinen toegevoegd, en in de avond en nacht en in het weekend zijn er andere afwijkingen, maar het Basis Uur Patroon is het onderliggende patroon. Het betekent bijvoorbeeld dat elk uur, om 26 minuten over het hele uur, een stoptrein van Zwolle naar Emmen vertrekt, van 's morgens vroeg tot 's avonds laat. Makkelijk te onthouden dus, deze 'kloktijden'.

Zwolle	↻ 9 56	10 26	↻ 10 56	11 26	↻ 11 56	12 26	↻ 12 56	13 26	↻ 13 56	14 26
Dalfsen	10 05	10 34	11 05	11 34	12 05	12 34	13 05	13 34	14 05	14 34
Ommen	10 15	10 44	11 15	11 44	12 15	12 44	13 15	13 44	14 15	14 44
Mariënberg		10 51		11 51		12 51		13 51		14 51
Hardenberg	10 26	10 57	11 26	11 57	12 26	12 57	13 26	13 57	14 26	14 57
Gramsbergen		11 01		12 01		13 01		14 01		15 01
Coevorden	10 36	11 07	11 36	12 07	12 36	13 07	13 36	14 07	14 36	15 07
Dalen		11 11		12 11		13 11		14 11		15 11
Nieuw Amsterdam	10 44	11 18	11 44	12 18	12 44	13 18	13 44	14 18	14 44	15 18
Emmen Bargeres		11 23		12 23		13 23		14 23		15 23
Emmen	A ↻ 10 51	11 28	↻ 11 51	12 28	↻ 12 51	13 28	↻ 13 51	14 28	↻ 14 51	15 28

Elke uurlijkse vertrek- en aankomsttijd wordt gegeven door een getal uit de verzameling

$$\{0, 1, 2, \dots, 59\}.$$

Het is handig om deze getallen te zien als getallen modulo 60, waarmee je dan kunt optellen en aftrekken: als de uitkomst boven de 60 is trek je er 60 van af, en als de uitkomst negatief is tel je er 60 bij op. Dus

$$45 + 8 = 53, 55 + 8 = 3, 45 - 8 = 37, 5 - 8 = 57.$$

Het vaststellen van de tijden in het BUP komt dan neer op het toekennen van getallen modulo 60 aan 'onbekenden'. Wat zijn dit en waaraan moeten ze voldoen?

3 De uurlijkse treinen

Bekijk allereerst, als voorbeeld, een uurlijkse trein van Amsterdam naar Amersfoort, de 'treinserie' 15. Deze stopt onderweg in Hilversum.

treinnummer	4637	IC 1539
Amsterdam Centraal	11 20	11 27
Amsterdam Muiderpoort	11 26	
Diemen	11 30	
Hoofddorp		
Schiphol	A	
Schiphol	A	
Amsterdam Zuid		
Amsterdam Zuid		
Amsterdam RAI		
Duivendrecht		
Duivendrecht		
Diemen Zuid		
Weesp	11 36	
Weesp		
Naarden-Bussum		
Bussum Zuid		
Hilversum Noord		
Hilversum	A	11 48
Hilversum		
Baarn		11 49
Amersfoort	A	12 02

Voor de vast te stellen vertrek- en aankomsttijden voeren we ‘onbekenden’ in:

$$x_{15,Asd,V}, x_{15,Hvs,A}, x_{15,Hvs,V}, x_{15,Amf,A}.$$

Hierin zijn Asd, Hvs en Amf de NS-afkortingen voor Amsterdam, Hilversum en Amersfoort, en V en A staan voor vertrek en aankomst.

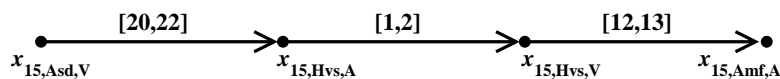
Voor het maken van de dienstregeling, moeten we tijden vaststellen voor deze onbekenden. Als eis wordt gesteld dat de rijtijd van Amsterdam naar Hilversum minimaal 20 en maximaal 22 minuten moet zijn. In Hilversum dient de trein minimaal 1 en maximaal 2 minuten te stoppen.

Van Hilversum naar Amersfoort is de rijtijd minimaal 12 en maximaal 13 minuten. Deze eisen kunnen worden beschreven als:

$$\begin{aligned} x_{15,Hvs,A} - x_{15,Asd,V} &\in [20, 22], \\ x_{15,Hvs,V} - x_{15,Hvs,A} &\in [1, 2], \\ x_{15,Amf,A} - x_{15,Hvs,V} &\in [12, 13]. \end{aligned}$$

Getallen modulo 60 die voldoen aan deze eisen geven voor deze treinserie de uurlijkse vertrek- en aankomsttijden. Dus, zoals in de lopende dienstregeling, $x_{15,Asd,V} = 27$, $x_{15,Hvs,A} = 48$, $x_{15,Hvs,V} = 49$ en $x_{15,Amf,A} = 2$ is toegestaan als oplossing van deze eisen, want $2 - 49$ zit in $[12, 13]$ modulo 60.

We kunnen de eisen aan deze treintijden grafisch als volgt weergeven:



Elke vast te stellen vertrek- of aankomsttijd wordt dus voorgesteld door een punt, en elk van de eisen wordt aangegeven met een pijl waarlangs het voorgeschreven interval staat.

Dit zijn de eisen voor één uurlijkse treinserie, en ook voor elke andere uurlijkse treinserie kunnen we de eisen aan de tijden als boven beschrijven. Zo is er ook een treinserie 15 terug van Amersfoort naar Amsterdam, waarbij we in de naamgeving van de variabelen even moeten uitkijken dat geen

dubbele variabelen ontstaan, bijvoorbeeld voor de tijden in Hilversum. De teruggaande trein zouden we met 15t in plaats van 15 kunnen aangeven.

Alle treinen tezamen geven dan een grote gerichte graaf, met zo'n 1800 punten en 1600 pijlen, bestaande uit ongeveer 200, los van elkaar liggende, gerichte paden, elk corresponderend met een treinserie heen of terug.

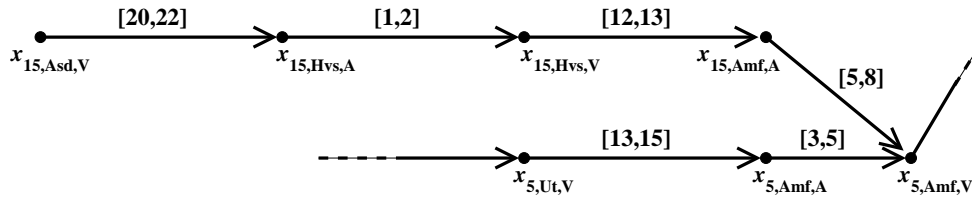
4 Aansluitingen

Als de bovenbeschreven voorwaarden alle eisen zouden zijn, zou het niet zo moeilijk zijn de tijden vast te stellen, maar er zijn meer condities. Ten eerste de condities die ontstaan doordat we overstapmogelijkheden tussen verschillende treinseries willen creëren. Zo is er een uurlijkse treinserie 5 van Rotterdam via Amersfoort naar Leeuwarden, waarop de treinserie 15 uit Amsterdam in Amersfoort dient aan te sluiten. Hierdoor ontstaat een eis voor twee verschillende treinseries:

$$x_{5,Amf,V} - x_{15,Amf,A} \in [5, 8].$$

Dus de trein naar Leeuwarden moet minimaal 5 en maximaal 8 minuten na aankomst van de trein uit Amsterdam van Amersfoort vertrekken. (De overstaptijd houdt rekening met de perronsituatie in Amersfoort.)

Grafisch kan dit als volgt worden weergegeven:



Dit als voorbeeld van één gewenste overstapmogelijkheid; over het hele land zijn er zo'n 600. Alle aansluitingsvoorwaarden tezamen geven dus veel meer samenhang aan de graaf.

5 Voorkomen van conflicten

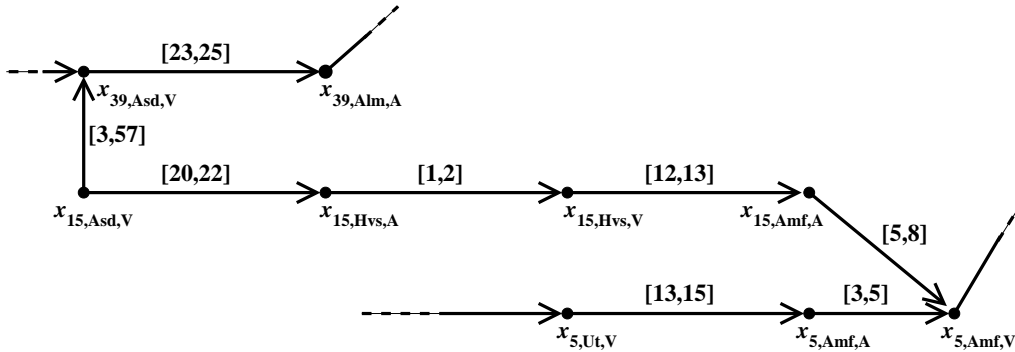
Alle gewenste aansluitingen, over het hele land, geven een al wat ingewikkelder patroon van eisen aan de tijden, maar dat is nog niet alles. Ook moeten zgn. 'conflicten' worden voorkomen: twee verschillende treinen mogen niet

gelijktijdig van hetzelfde spoor gebruik maken, bijvoorbeeld op een kruising of op enkelspoor, maar ook langs een traject: om te zorgen dat er voldoende ruimte tussen twee treinen wordt gepland, wordt een minimale ‘opvolgtijd’ van 3 minuten geëist, dit ook om kleine vertragingen op te kunnen vangen. (De seinen moeten zorgen voor de uiteindelijke veiligheid.)

Als voorbeeld de treinserie 39 van Hoofddorp via Amsterdam naar Lelystad. Deze heeft op het traject Amsterdam-Weesp een traject gemeenschappelijk met de bovengenoemde treinserie 15 van Amsterdam naar Amersfoort. Op dit traject moeten de treinen een minimale afstand van 3 minuten in tijd hebben. Dit kan als volgt worden aangegeven:

$$x_{39,Asd,V} - x_{15,Asd,V} \in [3, 57].$$

Door deze eis te stellen wordt dus voorkomen dat het verschil in de vertrektijden vanuit Amsterdam gelijk is aan 58, 59, 0, 1 of 2 (modulo 60) — dan zouden de treinen te dicht op elkaar zitten.

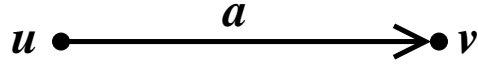


Toevoeging van deze eisen (die alle op zich een vrij ruim interval hebben) zorgt voor de grootste toevoeging van pijlen aan de graaf, meer dan 8000.

6 Graafkleuring

Al deze eisen aan rij-, stop-, overstap- en opvolgtijden tezamen vormen alle voorwaarden waaraan de vast te stellen tijden moeten voldoen. Ze geven een heel verknoopt patroon, maar ze hebben wel een eenvoudige structuur. Ze zijn in feite wiskundig alle van hetzelfde type: elke voorwaarde schrijft voor dat het verschil van twee variabelen in een voorgeschreven interval zit (modulo 60). Ze passen alle in het model van de gerichte graaf.

Als gebruikelijk geven we de verzameling punten van de graaf aan met V ('vertices') en de verzameling pijlen met A ('arrows' of 'arcs'). Een pijl a die van punt u naar punt v loopt wordt aangegeven met (u, v) .



Met $[l_a, u_a]$ geven we het langs pijl a staande interval aan. (De letters l en u staan voor 'lower bound' en 'upper bound'.)

De taak is nu om getallen modulo 60 bij de punten neer te zetten zo dat voor elke pijl a , het verschil van het getal bij de kop en het getal bij de staart in het interval $[l_a, u_a]$ zit. Dat wil zeggen, we moeten een functie $t : V \rightarrow \{0, \dots, 59\}$ vinden zo dat

$$t(v) - t(u) \in [l_a, u_a] \pmod{60} \text{ voor elke pijl } a = (u, v). \quad (1)$$

Dit doet denken aan het aloude kaartkleuringsprobleem, waarbij landen worden voorgesteld door punten, en een lijn tussen twee punten aangeeft dat de twee corresponderende landen aan elkaar grenzen. Kaartkleuring met 4 kleuren komt dan neer op het neerzetten van getallen modulo 4 bij de punten, zo dat voor elke lijn het verschil van de getallen bij de uiteinden in het interval $[1, 3]$ (modulo 4) zit.

7 Reductie van de zoekruimte

Naïef zou je nu natuurlijk kunnen denken: er zijn maar eindig veel mogelijkheden om getallen modulo 60 aan n punten toe te wijzen (nl. 60^n), en dus zou de computer alle mogelijkheden kunnen afdraaien en kijken of daar een oplossing tussen zit die aan alle eisen voldoet.

Voor $n = 1800$ duurt dit natuurlijk veel en veel te lang, en je doet ook te veel: als je bij alle tijden van een oplossing 1 minuut optelt, krijg je weer een oplossing, want de verschillen veranderen daardoor niet. Dus je zou één vast te stellen tijd vast op 0 kunnen zetten. Hierdoor reduceert de 'zoekruimte' met een factor 60, maar dat is uiteraard niet voldoende.

De zoekruimte kan flink worden gereduceerd als volgt. We verlaten het 'modulo 60 model', en zien de intervallen als intervallen van gehele getallen. (We zetten dan bijvoorbeeld $[58, 3]$ om naar $[58, 63]$.) Als $t : V \rightarrow \mathbb{Z}$ een oplossing is, dan bestaat er een functie $k : A \rightarrow \mathbb{Z}$ zo dat

$$t(v) - t(u) \in [l_a + 60k(a), u_a + 60k(a)] \text{ voor elke pijl } a = (u, v). \quad (2)$$

De ‘modulo 60’ voorwaarde in (1) wordt hier dus omzeild door een geheel veelvoud van 60 toe te voegen.

We kunnen de zoekprocedure nu omdraaien: we zoeken naar een functie $k : A \rightarrow \mathbb{Z}$ zodanig dat er een functie $t : V \rightarrow \mathbb{Z}$ bestaat die voldoet aan (2). Nu is het mogelijk om voor vaste $k : A \rightarrow \mathbb{Z}$ snel te bepalen of een dergelijke functie $t : V \rightarrow \mathbb{Z}$ bestaat. Dit kan met behulp van een aanpassing van het zgn. *Bellman-Ford kortste pad algoritme*: zet eerst $t(v) := 0$ voor elk punt v . Doe daarna het volgende:

voor elke pijl $a = (u, v)$ met $t(v) - t(u) \notin [l_a + 60k(a), u_a + 60k(a)]$: (3)

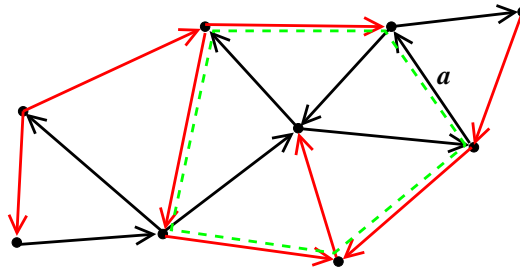
- als $t(v) - t(u) < l_a + 60k(a)$, d.w.z. $t(v) < t(u) + l_a + 60k(a)$, verhoog $t(v)$ tot $t(v) := t(u) + l_a + 60k(a)$,
- als $t(v) - t(u) > u_a + 60k(a)$, d.w.z. $t(u) < t(v) - u_a - 60k(a)$, verhoog $t(u)$ tot $t(u) := t(v) - u_a - 60k(a)$.

Herhaal (3) $|V|$ keer. Als t hierna aan alle eisen voldoet heb je een oplossing voor (2). Als de uiteindelijke t niet aan alle eisen voldoet, dan kan bewezen worden dat (2) helemaal geen oplossing heeft (voor deze $k : A \rightarrow \mathbb{Z}$), en dan moeten we k anders kiezen.

Voor vaste $k : A \rightarrow \mathbb{Z}$ vergt dit $|V| \cdot |A|$ stappen, en je kunt het met wat eenvoudige ingrepen versnellen. Voor het NS-probleem kost het een fractie van een seconde — er is geen exponentieel snel groeiende zoekboom meer nodig.

Het probleem is nu dus gereduceerd tot: hoe vinden we $k : A \rightarrow \mathbb{Z}$ waarvoor (2) een oplossing $t : V \rightarrow \mathbb{Z}$ heeft? In eerste instantie zijn er oneindig veel mogelijkheden om k te kiezen.

We kunnen dit aantal echter eindig maken door een willekeurige opspannende boom B te kiezen. (We nemen aan dat de graaf samenhangend is — dat is bij de NS inderdaad het geval.) Dan kunnen we aannemen dat $k(a) = 0$ als a in B zit. Vervolgens blijven er nog maar eindig veel



In rood een opspannende boom, en de groene streepjeslijn geeft de met pijl a gevormde ‘fundamentele cykel’ aan.

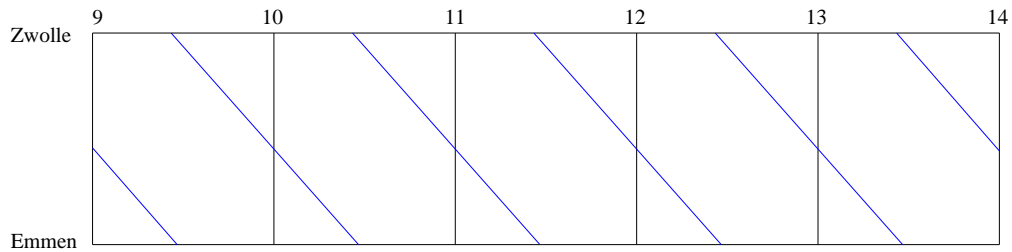
mogelijkheden over voor de waarden van $k(a)$ als a niet in B zit. Dit kunnen we zien door naar de cykel te kijken die a met B maakt (de ‘fundamentele cykel’), en te zien dat de intervallen langs de pijlen in dit circuit een onder- en een bovengrens voor $k(a)$ impliceren.

Het aantal mogelijkheden voor $k(a)$ is afhankelijk van de ‘ruimte’ die deze cykel toelaat. Soms is $k(a)$ hierdoor volledig vastgelegd, soms is er keus uit slechts 2 mogelijkheden, soms ook is het aantal mogelijkheden iets groter. Hiermee wordt het zoekproces flink gereduceerd.

8 Symmetrische dienstregeling

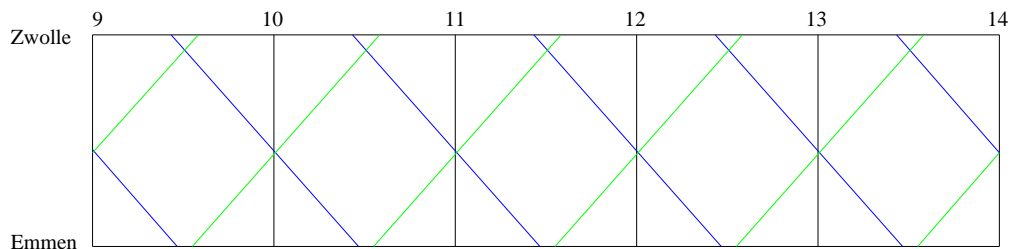
Een uurlijks patroon van treinen geeft een interessante symmetrie, die handig is bij het onthouden van treintijden. Deze symmetrie kan worden uitgelegd met de zgn. *dienstregelinggrafiek*, waarmee dienstregelingen vaak worden weergegeven.

De uurlijkse stoptreinen Zwolle-Emmen (zoals gegeven in § 2) geven bijvoorbeeld de volgende grafiek (van 9.00 uur tot 14.00 uur):

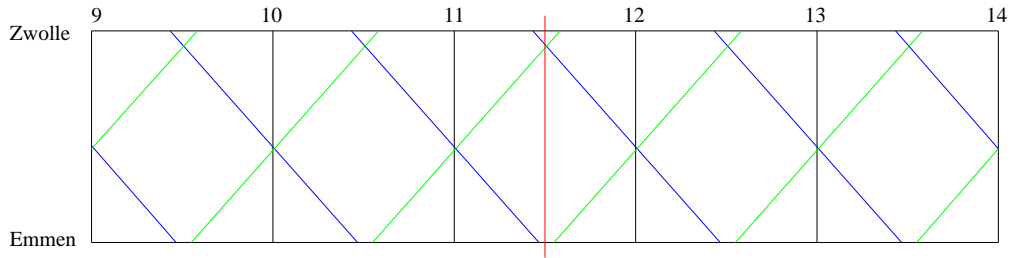


De X-as correspondeert dus met de tijd, en de Y-as met het traject Zwolle-Emmen.

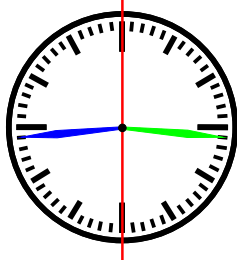
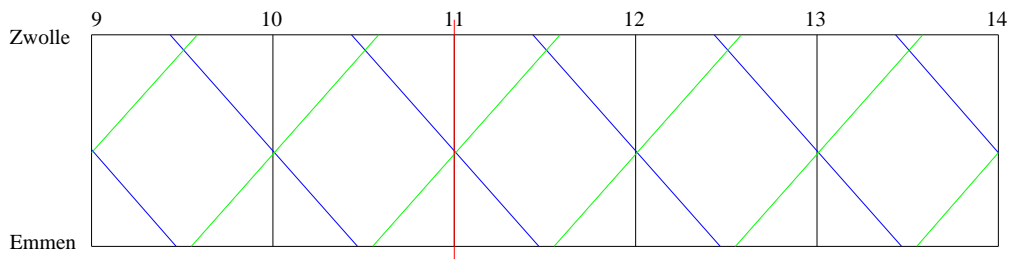
We kunnen ook de uurlijkse stoptreinen terug in de grafiek aangeven:



Je ziet nu een interessante symmetrie in het plaatje: als je spiegelt om de rode as in



dan gaan de treinen heen en terug in elkaar over. In feite is er elk half uur een symmetrie-as, ook bijvoorbeeld de volgende rode as:



Deze symmetrie-as ligt op .00 en .30. Wanneer je nu deze symmetrie-as kent, en je weet ook dat de stoptrein Zwolle-Emmen om .44 in Ommen aankomt, dan weet je ook dat de stoptrein Emmen-Zwolle om .16 uit Ommen vertrekt.

Iedere treinserie heeft dus een symmetrie-as. Nu is het interessante dat deze as voor alle treinseries, over het hele land, dezelfde is! Dit komt doordat in de dienstregeling aansluitingen tussen treinseries in de ene richting ook in de andere richting bestaan, met (vrijwel) dezelfde overstaptijden. Dus als treinserie 15 heen in Amersfoort aansluit op treinserie 5 heen, dan sluit treinserie 5 terug in Amersfoort aan op treinserie 15 terug. Hierdoor moeten de treinseries 5 en 15 dezelfde symmetrie-as hebben. Door de verknoping van aansluitingen tussen verschillende treinseries door het hele land, impliceert dit dat er één symmetrie-as is die geldt voor alle treinseries. Een symmetrie-as plant zich als het ware voort over het hele net.

Overigens is de landelijke symmetrie-as niet .00/.30, maar .59/.29, en daarop zijn weer flink wat uitzonderingen, o.a. doordat sommige trajecten of bruggen enkelsporig zijn, en gelijkvloerse splitsingen vaak ook asymmetrische verschuivingen in de treintijden nodig maken.

9 Ten slotte

De meeste treinseries in Nederland hebben een frequentie van 30 minuten in plaats van 60. Door overal 60 te vervangen door 30 is het model natuurlijk eenvoudig om te zetten. Toch hebben nog steeds enkele series een uurlijkse frequentie, en ook goederentreinen hebben vaak een uurlijkse dienstregeling (althans: ruimte in de dienstregeling). Om de capaciteit van het spoorweg-netwerk goed te benutten wordt bij NS nog steeds van een uurlijks patroon uitgegaan. Frequenties van 30 minuten kunnen ook in het model als eis worden opgenomen, bijvoorbeeld door te eisen

$$x_{15,Asd,V} - x_{115,Asd,V} \in [30, 30].$$

Hierin vertrekt treinserie 115 precies 30 minuten na treinserie 15.

Ten slotte: NS wil niet zomaar een dienstregeling, maar graag een optimale ... Hiermee raken we een open zenuw. Een commentaar in *NRC Handelsblad* noemde de NS-dienstregeling “de enige vorm van hogere wiskunde die heftige emoties oproept in het land”. Want wat bepaalt optimaliteit? Het reizigersconfort, de bedrijfswinst, de personeelsroosters, de materieelomloop, of de punctualiteit? Elk op zich is al lastig te kwantificeren, maar als dat al zou lukken, hoe zwaar moeten elk van deze aspecten onderling tegen elkaar worden afgewogen?

Binnen het bovenbeschreven model kunnen bepaalde aspecten worden geoptimaliseerd, zoals stop- en overstaptijden. Ook zijn er speciale algoritmen voor de materieelomloop en het personeelsrooster ontwikkeld. Maar één geïntegreerd systeem, dat alle onderdelen simultaan optimaliseert, is op dit moment nog een paar stations te ver. De wiskunde achter het Spoorboekje is voorlopig nog niet af!