

## De eerlijkheid van de Willem-Alexander roeibaan nabij Rotterdam

### Doel

Bij roeien is een wedstrijd baan opgedeeld in verschillende start banen, waar je veelal met zes ploegen naast elkaar start krijgt iedere ploeg zijn eigen baan toegewezen. In de roeiwereld wordt veel gepraat over zogeheten "voordeelbanen", een toegewezen baan waar je voordeel zou kunnen ondervinden van de wind en/of stroming, waar in een andere baan dit voordeel niet is en je zelfs een nadeel zou kunnen hebben. Recent is er een nieuwe roeibaan in Nederland aangelegd, de Willem-Alexander Baan. Deze baan is niet alleen qua faciliteiten een verbetering op de in Amsterdam gelegen Bosbaan, hij zou ook een stuk eerlijker zijn dan laatstgenoemde. Van de Bosbaan wordt gezegd dat het weer zeer veel invloed heeft op de stroming van het water in de verschillende banen, bij de nieuwe Willem Alexander Baan zou dit effect niet zo groot zijn.

De site van de Willem-Alexander baan zegt zelf: "De roeibaan als geheel ligt open onder de hemel, alleen omringd door kleine rietlanden. De wind heeft er weliswaar vrij spel, maar treft alle banen even hard. Het maakt de Willem Alexander Baan tot een eerlijke wedstrijd baan."

Ik ben benieuwd in hoeverre dit correct is. Is de Willem-Alexander baan echt zo eerlijk als dat ze zelf beweren of is de baan in feite, net als de Bosbaan, zeer beïnvloedbaar door de wind. Om deze vraag te kunnen beantwoorden hoe eerlijk de afgelopen drie wedstrijden op de Willem-Alexander baan nou geweest zijn moeten er natuurlijk weersdata en roeiuitslagen zijn.

De weersdata zijn gedownload van <http://www.knmi.nl/klimatologie/uurgegevens/selectie.cgi> voor weersstation Rotterdam waar de richting van de wind de snelheid en de hardste windstoot gekozen zijn als variabelen. Dit voor de volgende drie tijdvakken: 23 & 24 mei 2015 (Westelijke 2015), 2 & 3 mei 2015 (Damen 2015) en 3 & 4 mei 2014 (Damen 2014). Van alle dagen is wind data verzameld van 7:00 tot 20:00 uur aangezien dit vaak het tijdvlak is waarbinnen alle heats op een wedstrijd dag gevaren worden. De windrichting is weergegeven in graden, de windsnelheid en hardste windstoot in 0.1 m/s.

De roeiuitslagen, van de Westelijke 2015, de Damen 2015 en de Damen 2014, zijn gescraped van time-team.nl. Hier zijn drie verschillende scrapers voor gebruikt omdat de resultaten sites van een specifieke wedstrijd net iets anders in elkaar zitten. Gevolg hier van is dat elke wedstrijd zijn eigen databestand heeft in plaats van één databestand met alle resultaten van de drie wedstrijden. De drie scrape-implementaties zijn gemaakt in python en maken gebruik van patter.web. Het python script krijgt de URL van de webpagina weer waar alle heats in een tabel staan opgenoemd. Met behulp van deze URL wordt de DOM bereikt en uit de eerder beschreven tabel wordt steeds de URL gehaald voor die specifieke heat en alle URL's worden opgeslagen in een lijst.

Vanuit deze lijst wordt elke URL afgegaan om naar de pagina te gaan van een specifieke heat waar opnieuw de dom wordt gebruikt om een tabel te bereiken waar alle informatie van de heat staat opgeslagen. De resultaten van een heat worden allemaal opgeslagen in een lijst zoals te zien is in figuur 1. De vorm van de resultaten wordt later besproken. De heat worden in een grotere lijst gezet die het hele weekend representeert en vervolgens wordt de volgende URL uit de lijst gepakt om het zelfde te doen voor de volgende heat. De uiteindelijke lijst van het hele weekend wordt als een JSON bestand opgeslagen.

De resultaten staan in JSON bestanden waarin in een lijst staat welke resultaten er het hele weekend neergezet zijn. Elk element in de lijst heeft informatie over een ploeg in een heat. Van de ploeg staat weergegeven welke dag er gestart is, hoe laat, in welke heat, de naam van de ploeg

(verenigingsnaam), in welke baan de ploeg lag en een dictionary van resultaten waar per 500m de positie en de tijd aan toe zijn gevoegd. Een voorbeeld van een element in het JSON bestand is te zien in figuur 1.

```
[{"za", "08:00", "B 1x voorwedstrijd 1", "Okeanos", "7", {"results": {
  "1000m": {"position": "1", "time": "03:54,91"},
  "1500m": {"position": "1", "time": "05:58,87"},
  "finish": {"position": "1", "time": "08:04,32"},
  "500m": {"position": "1", "time": "01:53,80"}
}}]
```

**Figuur 1:** een voorbeeld van hoe een element in het json bestand er uit ziet. Het hele json bestand is een lijst die van alle startende ploegen in een weekend zo'n element bevat.

## Vertaling

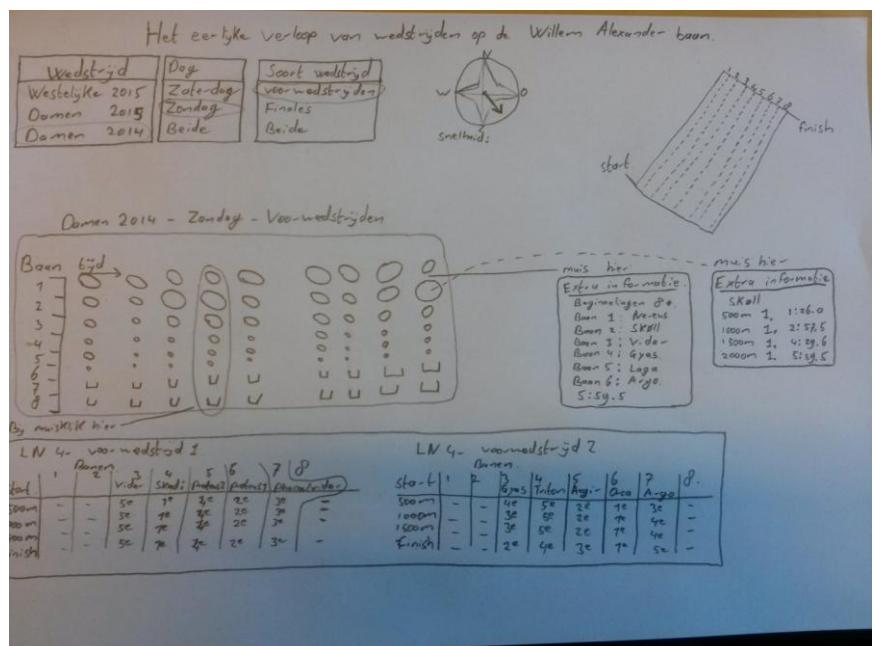
Hoe kom je van een grote hoeveelheid uitslagen naar een visualisatie waar je snel kan zien of de wedstrijden eerlijk zijn verlopen? Een spreidingsdiagram plot waar de banen zijn uitgezet over de tijd, op deze manier zie je van elke heat een stip wanneer er een boot in die baan gestart is. Als de grote of de kleur van zo'n stip afhangt van de behaalde positie is het makkelijk om te zien per heat in welke baan er gewonnen is.

De weersdata is gedownload in een csv bestand met bovenaan een tekst welke uitlegt hoe alle data zijn weergegeven, dit stuk tekst is verwijderd. In de csv's stond ook een kolom met de code van het weerstation Rotterdam, aangezien alle data van dit zelfde weerstation komt is het niet relevant om deze kolom te behouden en is dan ook handmatig verwijderd. Vervolgens zijn de csv bestanden omgezet in json bestanden om ze op de zelfde manier als de resultaten in te lezen.

## Ontwerp

In figuur 2 is een uitgebreide schets te zien van de bedachte visualisatie. Er staat statisch een baankaart met de ligging aangegeven met daarnaast een windroos. In een drie-stappen-keuze-menu moet de gebruiker aangeven welk wedstrijd weekend en welke dag hij uitslagen wil zien en of dit van de hele dag moet zijn, alleen de voorwedstrijden of alleen de finales.

Vervolgens gebeuren er twee hoofd stappen, er wordt wind informatie weergegeven in de windroos rechts boven, met een extra tekst boxje. Het tweede ding is dat er een spreidingsdiagram wordt getekend met de uitslagen van de heats uitgezet over de banen, waarin een groot bolletje de winnende ploeg weergeeft en hoe kleiner het bolletje hoe lager de behaalde positie van de ploeg in die heat. Als de gebruiker met de muis over de spreidingsdiagram plot navigeert verschijnt er een tooltip met extra informatie over de betreffende heat. Deze geeft aan welke heat het is, wat de winnende tijd is geweest van die heat en welke ploegen er startte en in welke baan. Tegelijkertijd wordt er gekeken in welke tijdvlak die heat valt en

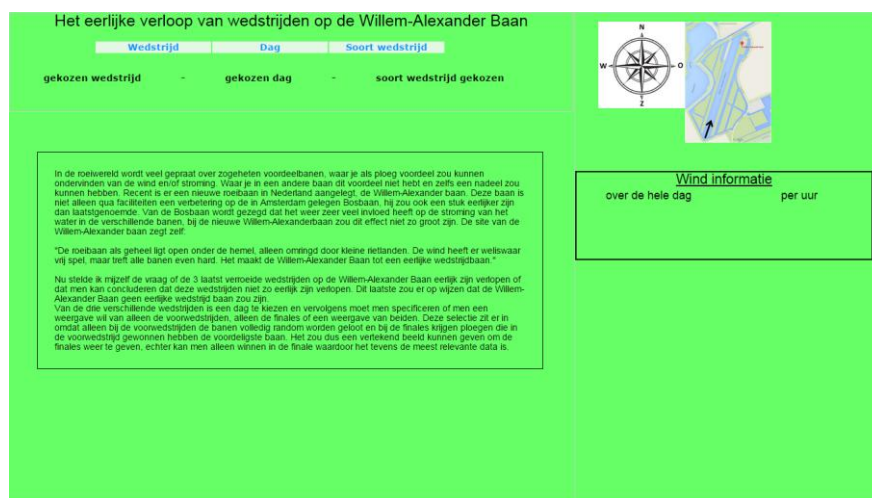


**Figuur 2:** Uitgebreide schets van de bedachte visualisatie. Rechtsboven een windroos en statische baankaart. Links boven een keuze menu, daar onder een spreidingsdiagram en rechts daarvan twee tooltips met extra informatie. Alleen zichtbaar als de muis op de juiste plaats staat. Onderaan is een extra svg geschetst waarin van alle heats in een veld een overzicht geschetst zou worden van het verloop van de races in dat veld.

wordt de bijbehorende wind-data er bij gezocht waar een extra pijl in de windroos wordt getekend om de richting op dat tijdstip weer te geven. In het boxje met informatie over de wind wordt nu ook een uur gemiddelde weergegeven.

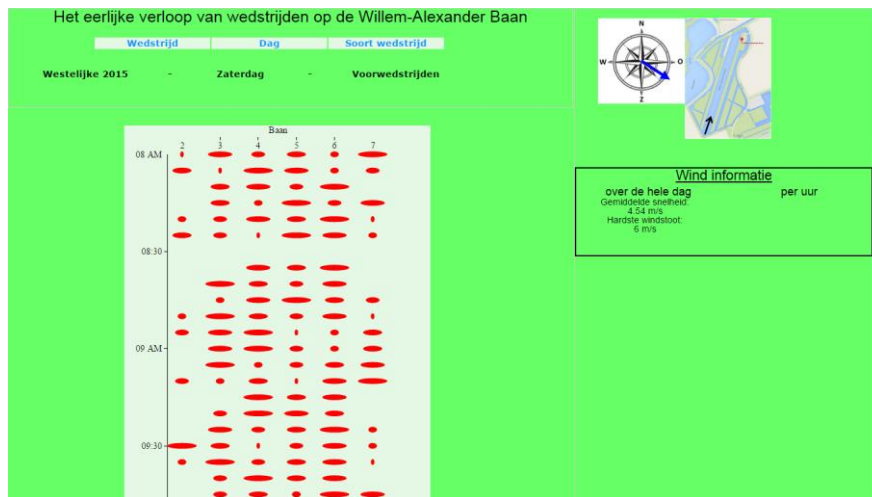
Gaat de gebruiker met de muis over een data punt in de grafiek verschijnt er een extra tooltip welke wederom weergeeft welke heat het is, welke ploeg je geselecteerd hebt en per 500m wat de positie en de tijd van die ploeg was.

De figuren 3 tot en met 7 laten zien hoe de visualisatie uiteindelijk is geworden. In figuur 3 is het begin scherm te zien. Er is nog geen enkele selectie gemaakt en er staat een groot tekstblok waarin men kan lezen over het doel van de visualisatie. Wanneer een selectie is gemaakt voor een wedstrijd, een dag en aangegeven is of men de voorwedstrijden, de finales of beiden wil zien verdwijnt het tekstblok en wordt vervangen voor een SVG waar de spreidingsdiagram in te zien is tevens worden de wind informatie en de windrichting toegevoegd (figuur 4).

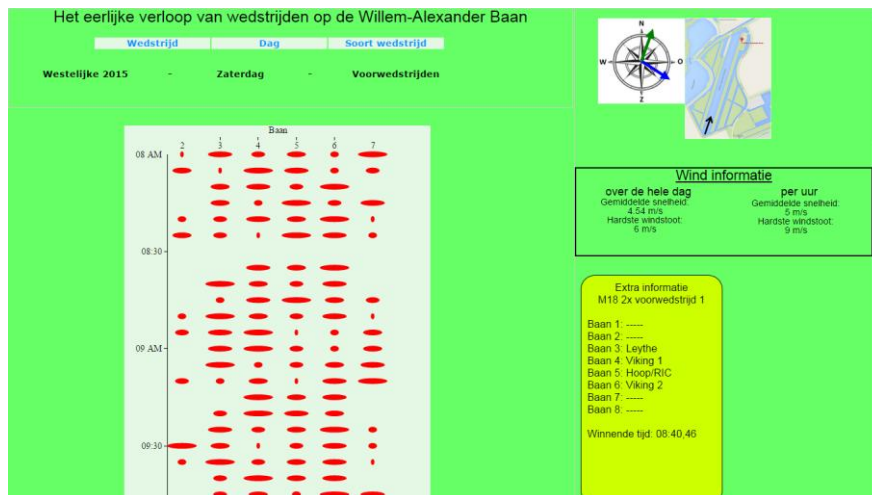


**Figuur 3:** De begin pagina met een tekst die uitlegt wat men kan met de pagina en waarom dit relevant zou zijn. Een lege windroos en wind-informatie box zijn te zien omdat er geen wedstrijd en dag gekozen zijn.

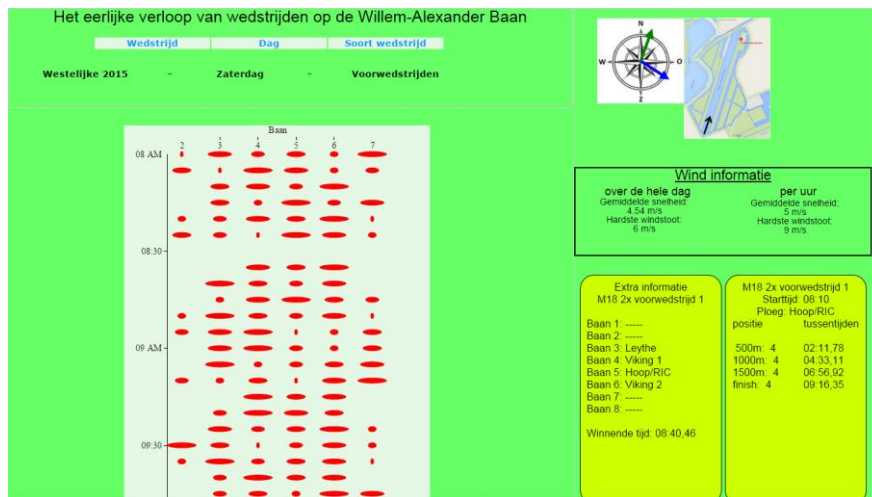
In figuur 5 is te zien hoe dat de gebruiker met de muis over een heat heen gaat. We zien een extra pijl in het kompas voor het uur van de dag waarin de heat is verroeid en extra informatie over de wind op dat uur. Tevens verschijnt er een box die weergeeft welke heat het is met de baanverdeling. In figuur 6 zien we de zelfde selectie maar nu staat de muis op het datapunt in baan 6, naast de informatie die al weer werd gegeven verschijnt er een extra informatie box welke aangeeft welke ploeg de muis aanwijst en per 500 meter op welke positie de ploeg lag en met wat voor tijd. De wind informatie voor dat uur blijft zichtbaar. Bij het verlaten van de grafiek zien we de beide informatie boxen en de wind informatie van dat uur verdwijnen. Als men naar beneden scrollt ziet men de conclusie die getrokken kan worden uit de visualisatie (figuur 7).



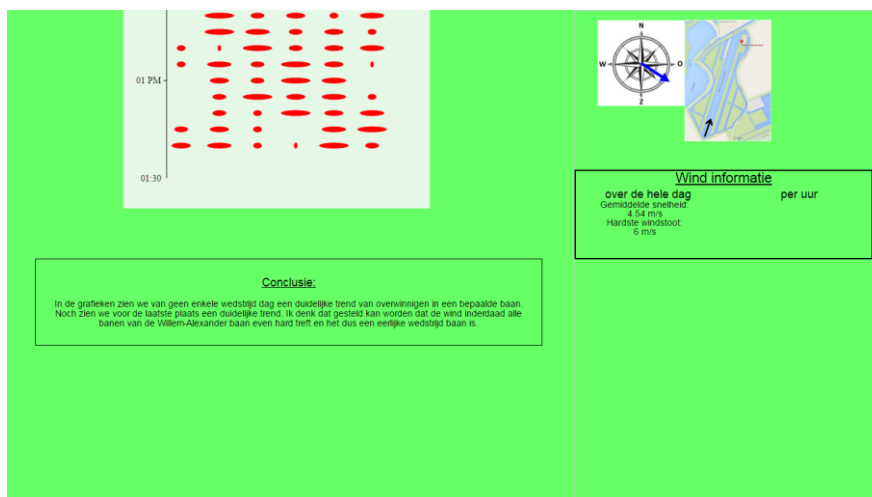
Figuur 4: Een selectie is gemaakt voor de Westelijke 2015 en de zaterdag is gekozen. Rechtsboven in de windroos is een pijl weergegeven die de windrichting weergeeft. In de box windinformatie is nu informatie verschenen van de gemiddelden van die dag. De resultaten van de wedstrijden zijn te zien in de spreidingsdiagram, een grote ellips is de winnaar van die heat, de kleinste ellips is de laatst gefinishte.



Figuur 5: De muis gaat over de spreidingsdiagram en een tooltip verschijnt met welke heat het is, de winnende tijd en welke ploeg in welke baan gestart is. In de windroos is een extra pijl verschenen met de windrichting in het uur dat de heat gestart is. In de wind-informatie-box zijn nu ook gemiddelden van dat uur weergegeven.



**Figuur 6:** De muis staat nu ook op een data punt en laat naast de tooltip over de heat nu ook een tooltip zien welke voor dat geselecteerde data punt specifiek is. Hier staat nog een keer welke heat het is, hoe laat deze gestart is en vervolgens per 500 meter interval op welke positie de ploeg lag en met wat voor tijd.



**Figuur 7:** Helemaal onderaan de pagina is een conclusie te lezen.

## Uitvoering

Om de visualisatie pas op te bouwen als alle data ingeladen is werd gebruik gemaakt van Queue om alle json bestanden in te laden. Vervolgens is met D3 de visualisatie opgebouwd en is het spreidingsdiagram gemaakt in een SVG.

Op de pagina wordt door de gebruiker een wedstrijd, dag en soort wedstrijd geselecteerd waarna door het script van de beschikbare data het juiste deel geselecteerd wordt voor de visualisatie. Vervolgens wordt de hele data selectie afgelopen om de visualisatie op te bouwen.

## Validatie

In de uitgebreide schets is te zien dat er naast het spreidingsdiagram nog een extra visualisatie getekend zou worden wanneer er op het spreidingsdiagram geklikt wordt. Deze visualisatie zou bestaan uit meerdere gelijksoortige grafieken, voor elke heat in het veld één. Dus is er geklikt op een heat uit het veld SA2- zouden er een aantal grafieken gemaakt worden, elk van een heat uit het SA2-veld. In elke grafiek zou dan een overzicht komen van hoe de heat is verlopen. In welke baan is welke ploeg gestart en dan van elke 500 meter de positie in het veld. Hier in zou een idee kunnen ontstaan van een trend als een bepaalde 500 meter meer wordt beïnvloed door de wind. Dit als bijvoorbeeld

in alle heats van een veld te zien is dat in de laatste 500 meter van de race baan 2 en 3 opeens plaatsen verliezen ten opzichte van de rest. Er zou geen echte conclusie uit getrokken kunnen worden omdat er slechts één veld per keer zichtbaar zou zijn maar een trend zou er mee aan het licht kunnen komen. Uiteindelijk is door tijdgebrek deze extra visualisatie niet bijgevoegd. Het oorspronkelijke idee was om het spreidingsdiagram horizontaal weer te geven en de extra visualisatie daar onder toe te voegen. Dit zou uiteindelijk niet werken omdat het spreidingsdiagram nu vertikaal staat, een betere oplossing zou zijn om een vlak over het spreidingsdiagram heen te voegen waar de extra visualisatie in getekend zou worden en bij een klik op de extra visualisatie dat dit vlak weer verdwijnt en het spreidingsdiagram weer te zien zou zijn.

Aangezien er twee verschillende tooltips zichtbaar zijn wanneer de gebruiker met de muis over een datapunt gaat is het mogelijk om te kijken of deze tooltips de correcte data weergeven. De functies die achter de tooltips zitten krijgen net op een andere manier de data binnen maar wanneer ze beide overeenkomstige dingen weergeven gaat het dus goed. De tooltip van de heats maakt gebruik van de muiscoördinaten, zet dat om in een tijd en zoekt daar een bijbehorend datapunt bij. Vervolgens wordt de index van dat datapunt doorgegeven aan de functie die de tooltip bijwerkt. In de bijwerk-functie wordt de doorgegeven index gebruikt en wordt er gekeken of de data punten die er 5 voor of achter liggen dezelfde heat code hebben. Er is gekozen voor 5 punten er voor en er na omdat in een heat maximaal 6 ploegen starten dus meer was niet nodig en bij minder zou je ploegen kunnen missen. Mochten de punten dezelfde heat code hebben worden ze toegevoegd aan een lijst welke uiteindelijk wordt gebruikt om de tooltip bij te werken.

Voor de tooltip van de ploeg werkt het net anders, daar wordt gekeken op welk data punt de muis staat en wordt direct de data van dat punt gezocht en gebruikt. Dit heeft echter wel tot gevolg dat men met de muis precies op het data punt moet staan omdat het anders de selectie niet werkt. Een verbetering zou dan ook zijn dat wanneer men met de muis in de buurt komt van een data punt dat dan de extra tooltip al verschijnt. Dit zou kunnen door er een extra ellips overheen te leggen welke niet zichtbaar is maar wel iets groter dan de originele. Op deze manier lijkt het of een datapunt geselecteerd is voordat de muis op het zichtbare gedeelte staat.

Wat nog meer verbeterd zou kunnen worden is de plaatsing van de conclusie box. Deze is onderaan de pagina altijd zichtbaar en voor alle wedstrijden en subselecties van data het zelfde. Dit omdat ik geen patroon heb ontdekt in de uitslagen en voor alle subselecties de zelfde conclusie heb getrokken, dat de Willem-Alexander Baan een eerlijke baan is, leek het mij niet nodig om een conclusie box te maken waar de inhoud aangepast kan worden. Noch leek het mij nodig om de conclusie enkel te tonen wanneer hier om gevraagd wordt, bijvoorbeeld bij een keer klikken op de grafiek. Dit is echter wel iets wat geïntegreerd zou kunnen worden. De verschillende opties voor de conclusie zouden bijvoorbeeld in een lijst opgeslagen kunnen worden en wanneer op een grafiek van een subselectie wordt geklikt, welke een patroon laat zien, de juiste conclusie wordt getoond. Na het klikken op de conclusie zou deze weer kunnen verdwijnen zodat de gebruiker verder kan rond kijken of bij andere data subselecties ook een patroon zichtbaar is.

De windroos is nu makkelijk een afbeelding van een kompas, in een svg-element, waar een pijl overheen wordt getekend. Het is natuurlijk mooier als dit kompas wordt vervangen door een windroos volledig getekend als svg waardoor het beter één geheel lijkt. De pijl die de windrichting per uur aangeeft wordt gedraaid met een transitie en zou dus mooi als een wijzer moeten werken. Echter zit hier een gebrekkige implementatie achter waardoor de wijzer niet mooi om het middelpunt van het kompas draait. In plaats daarvan verspringt de wijzer tijdens de transitie, de eindstand van de transitie is wel goed maar de overgang is niet zo mooi en soepel als dat wenselijk is. Het is zeker een verbeter punt om de transitie uit te voeren vanuit de basis van de pijl zodat hij echt als een wijzer draait vanuit het centrum van het kompas.