

DEEL 4: VERKEERSSTROOMTHEORIE EN VERKEERSMANAGEMENT

In dit deel wordt ingegaan op het verkeer zelf: Hoe wordt de weg gebruikt? Congestie speelt hier een belangrijke rol. En als het gebruik van de infrastructuur niet efficiënt is, hoe kunnen we dan het verkeerssysteem beter maken. Verkeersmanagement speelt hier een rol, zowel op kruispuntniveau als op netwerkniveau. Achtereenvolgens komen aan bod:

- Verkeersstroomtheorie: congestie en schokgolven
- Verkeersmanagement: kruispunt, autosnelweg, netwerken

In deze tweede versie is een aantal details verbeterd. De figuur van de uitwerking van voorbeeld 1 File voor een bottleneck (Figuur 1.13) is aangepast (het oplossen van de file) en de voetnoot bij voorbeeld 2.2 is toegevoegd.

DEEL 4: VERKEERSSTROOMTHEORIE EN 1**VERKEERSMANAGEMENT..... 1****1 VERKEERSSTROOMTHEORIE OF: HOE ONTSTAAN FILES? 3**

1.1	Inleiding: individueel en collectief rijgedrag	3
1.2	Basisbegrippen uit de verkeersstroomtheorie	5
1.2.1	Trajectoriën	5
1.2.2	Intensiteit, dichtheid en snelheid	5
1.3	Fundamenteel diagram en capaciteit	7
1.3.1	Kenmerken van het Fundamenteel diagram	8
1.3.2	Vrije snelheid, capaciteit en andere fundamenteel diagramparameters	9
1.3.3	Toepassingen van het fundamenteel diagram	10
1.4	Hoe files ontstaan – een kwalitatieve analyse	12
1.4.1	Files en flessenhalzen.....	13
1.4.2	De file als golf	14
1.5	Schokgolf theorie	15
1.5.1	Analytische afleiding schokgolfsnelheid.....	15
1.5.2	Grafische interpretatie.....	16
1.6	Toepassingen schokgolftheorie	16
1.6.1	Berekening van de filelengte bij een bottleneck – grafische uitwerking	17
1.6.2	Berekening van de filelengte bij een bottleneck – nu met formules	18
1.7	Samenvatting en reality check: grenzen aan schokgolftheorie	19
1.7.1	Verkeer in evenwicht en hysteresis.....	20
1.7.2	Filegolven, spontane fase overgangen en de capaciteitsval.....	21

2 VERKEERSMANAGEMENT 23

2.1	Stedelijk verkeersmanagement: verkeerslichten	23
2.1.1	Basisbegrippen: nummering, conflictmatrices en ontruimingstijden.....	24
2.1.2	Ontwerpen van verkeersregelingen	27
2.1.3	Toepassingen van schokgolftheorie bij geregelde kruispunten	31
2.1.4	Voertuigafhankelijke (verkeers)regelingen: gebruik maken van detectiemiddelen	34
2.2	Voorbeelden van dynamisch verkeersmanagement op snelwegen	35
2.2.1	Regelsystemen en regelcirkels.....	35
2.2.2	Voorbeeld 1: toeritdosering.....	38
2.2.3	Voorbeeld 2: de 80 km/u zones	39
2.2.4	Voorbeeld 3: real-time route advies	43
2.3	Integraal verkeersmanagement in netwerken	47
2.3.1	Het netwerk fundamentele diagram	47
2.3.2	Het “meer is minder” effect: waarom netwerk verkeersmanagement belangrijk is.....	48
2.3.3	Geïntegreerd en gecoördineerd netwerk management	50
2.4	Tenslotte – nogmaals de omgekeerde fles	51

1 VERKEERSSTROOMTHEORIE OF: HOE ONTSTAAN FILES?

De beste manier om te begrijpen hoe files ontstaan en zich ontwikkelen over de ruimte en tijd, is om gedurende een middagspits (om een uur of vier) op de brug over de A13 bij afslag Delft-Noord te gaan staan, en te observeren wat er gebeurt. Op een gegeven moment zie je dan dat benedenstrooms, net voorbij de afslag Delft-Noord, iemand op de rem gaat staan. Vervolgens zie je hoe, als een soort golf, de rode remlichten zich tegen de rijrichting van het verkeer in bewegen. Op de plek waar het eerste voertuig op de rem stapte wordt misschien al weer geaccelereerd, maar het proces is – in elk geval op de A13 – onomkeerbaar, want ook verder benedenstrooms (bij afslag Delft of Delft-Zuid) is iemand op de rem gaan staan en een tweede golf met remlichten beweegt zich in de verte al richting de brug.



FIGUUR 1.1: FILE OP DE A13 BIJ DELFT-NOORD (FOTO: M. MINDERHOUD)

1.1 Inleiding: individueel en collectief rijgedrag

Het feit dat files zich laten beschrijven als golven vormt de basis voor ons begrip van de verkeersafwikkeling en levert ons het wiskundig gereedschap om de collectieve dynamica van verkeersstromen en files te beschrijven. De wiskundige modellen die we daarbij gebruiken worden dan ook *macroscopische* verkeersafwikkelingsmodellen genoemd, modellen die verkeer als stroom beschrijven. De kernvraag is natuurlijk *waarom* binnen zo'n verkeersstroom die ene persoon nu zo nodig op de rem gaat staan. Om die vraag te beantwoorden is ook inzicht nodig in het *individuele* rijgedrag van voertuigen in die verkeersstroom. Hoe reageren voertuigbestuurders op hun voorganger, wanneer remmen of accelereren ze, en wanneer besluit iemand een inhaalbeweging te maken? Theorieën en modellen die dit individuele gedrag beschrijven noemen we *microscopische* verkeerstheorieën en -modellen. Als zo een micromodel dat individuele gedrag heel nauwkeurig en juist beschrijft zal het dezelfde macroscopische resultaten opleveren die we in de praktijk zien: file op de A13 bijvoorbeeld. Sterker nog, we zullen in dit hoofdstuk zien dat micro- en macroscopische variabelen en beschrijvingen van verkeer zeer nauw samenhangen.

We hebben al vaak herhaald in dit college dat in het vakgebied Transport en Planning alles met alles samenhangt. Met behulp van de Wegenercirkel hebben we gezien dat grondgebruik (ruimtelijke ordening) via activiteiten en verplaatsingspatronen leidt tot vervoerwijze-, vertrektijdstip- en routekeuzen. Die leiden op hun beurt via de daadwerkelijk gerealiseerde trips tot reistijden en kosten (files!), die bereikbaarheid van locaties en mobiliteit van personen begrenzen en bepalen. Routekeuze (net als vertrektijdstip of vervoerwijzekeuze) speelt dus *voorafgaand* aan een rit een belangrijke rol. Maar ook als mensen eenmaal onderweg zijn is het nog mogelijk om van route te veranderen, zeker in stedelijke gebieden. Al die individuele routekeuzes (zowel vooraf als onderweg) resulteren (macroscopisch) in een bepaalde verdeling van verkeer over een netwerk. De effecten van niet goed afgestemd routekeuzegedrag kunnen voor de verkeersafwikkeling extreem zijn, zodanig zelfs dat een verkeersnetwerk volledig vast komt te staan. Dit effect, wat we vaak "gridlock" noemen, zullen we in de laatste paragraaf van dit hoofdstuk bespreken. Het fenomeen dat in een verkeersnetwerk de doorstroming (de capaciteit) omlaag gaat zodra het wordt overbelast vormt een belangrijk verschil tussen verkeersnetwerken en bijvoorbeeld waternetwerken (zoals systemen van sloten, polders, boezems en kanalen). Dat zorgt er ook voor dat we verkeersnetwerken op een heel andere manier moeten managen en regelen. Hoe dat kan en aan welke knoppen we dan moeten draaien (van verkeersregeling, routenavigatie, toeritdosering en 80 kilometerzone tot volledig geautomatiseerde wegen met onderling communicerende voertuigen) zullen we in het volgende hoofdstuk bespreken.

Persoonlijke voorkeur, ervaring en gewoontegedrag spelen een belangrijke rol in zowel ons rijgedrag als ons routekeuzegedrag. Maar in beide gevallen is een individu niet volledig vrij in zijn keuzes. Er is ook een collectieve dynamica, die de individuele keuzevrijheid begrenst en zelfs stuurt. Dat zo'n collectieve dynamica tot soms bizarre effecten leidt zien we in de natuur heel vaak. Denk maar eens aan de volgende simpele individuele gedragsregels: "(a) volg je buurman en (b) bij gebrek aan buurman: kies maar een richting". Uitgevoerd door een school vissen of een vlucht trekvogels leiden zulke simpele gedragsregels tot prachtige, zeer complexe collectieve zwem- en vliegpatronen. Ons gedrag in het verkeer is ingewikkelder, omdat we – naast enkele basale regels (bijvoorbeeld accelereren of remmen), ook veel geavanceerder rij- en reisgedrag vertonen zoals (soms onverantwoorde) inhaal- en invoegmanoeuvres of allerlei ad hoc route keuzes. Niettemin zie je ook in de verkeersafwikkeling terugkerende patronen, die goed zijn te reproduceren met soms verrassend eenvoudige gedragsregels en -modellen. Het maakt files tot op zekere hoogte ook voorspelbaar. Die reservering (tot op zekere hoogte) is belangrijk. Het gedrag van automobilisten is erg variabel. Vergelijk het oude omaatje met de bumperklever, de vrachtwagenchauffeur met de Porsche-rijder of de forens met de toerist in een vreemde stad. Die grote verscheidenheid aan rijgedrag zorgt ervoor dat het collectieve resultaat (file of geen file?) alleen gemiddeld kan worden voorspeld met in het beste geval onzekerheidsmarges daaromheen. Je kunt dit vergelijken met het weer. Het is wel mogelijk om te voorspellen dat er 100% kans op neerslag is in een bepaald gebied, maar het is onmogelijk om de precieze tijd en locatie te voorspellen waar en wanneer de bui losbarst. Zo kunnen we heel goed voorspellen dat er file ontstaat voor de Coentunnel. We kunnen zelfs heel goed verklaren hoe de onderliggende dynamica in het rij- en reisgedrag in elkaar steekt. Maar we kunnen nooit de precieze tijd voorspellen wanneer deze file zal ontstaan¹ en hoe deze zich precies zal gedragen over de tijd en ruimte.

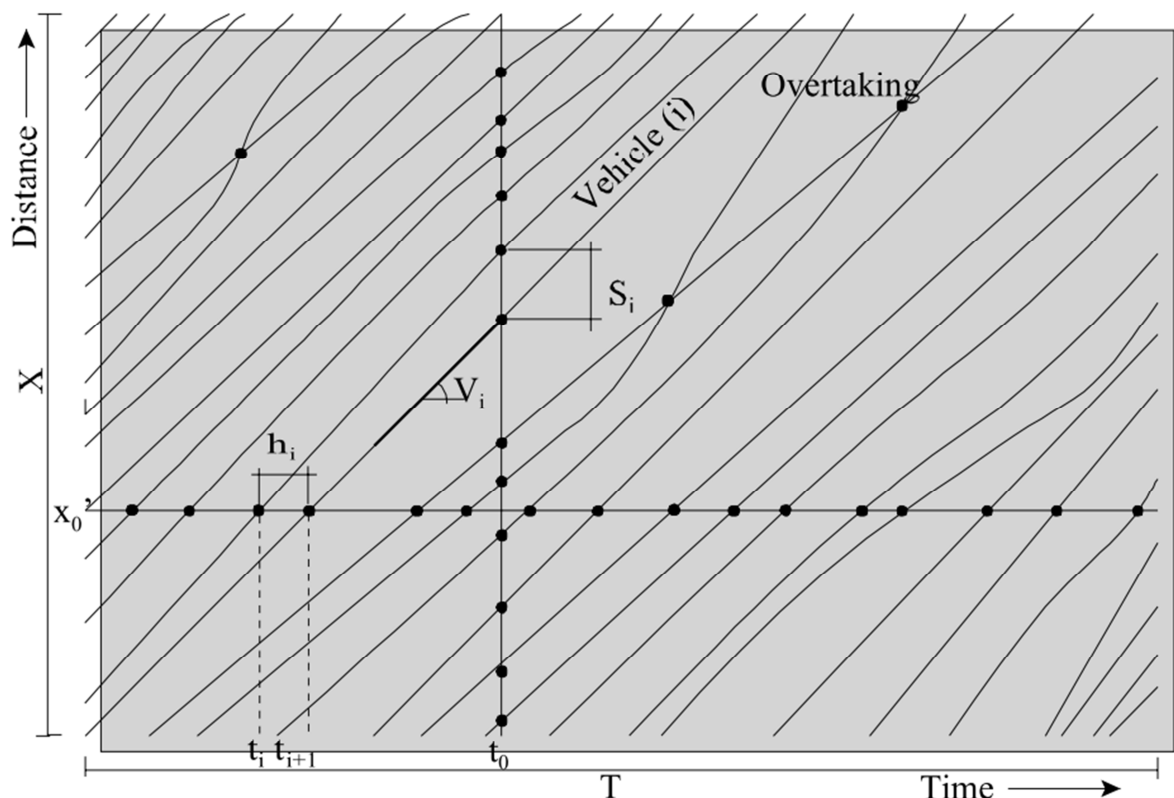
In dit hoofdstuk zullen we kijken naar die onderliggende mechanismen achter de file, zullen we uitleggen hoe ons rij- en routekeuzegedrag samenhangt met het ontstaan of oplossen van files, en zullen we ook de grenzen verkennen van wat we wel en niet kunnen voorspellen. Voordat we dit doen moeten we eerst een aantal basisbegrippen bespreken.

¹ De locatie van files zijn wel voorspelbaar – behalve dan files als gevolg van incidenten of ongelukken.

1.2 Basisbegrippen uit de verkeersstroomtheorie

1.2.1 Trajectoriën

Een trajectorie $x_i(t)$ van een voertuig i beschrijft de positie van een voertuig als functie van de tijd. De trajectorie is de kerngrootheid van de verkeersstroomtheorie en stelt ons in staat *alle relevante microscopische en macroscopische grootheden* af te leiden. Ter illustratie toont Figuur 1.2 diverse trajectoriën. Uit de figuur kan worden opgemaakt wat de volgafstand s_i en de volgtijd h_i van een voertuig i is, wanneer wordt ingehaald, wat de snelheid $v_i = dx_i/dt$ bedraagt, hoe groot de acceleratie $a_i = dv_i/dt$ is, wat de reistijd TT_i van een voertuig is, etc.



FIGUUR 1.2: TRAJECTORIËN, VOLGAFSTANDEN EN VOLGTIJDEN.

Helaas is trajectorie-informatie (nu nog) slechts sporadisch beschikbaar, maar dat gaat in de komende jaren sterk veranderen. Met name bedrijven die routenavigatiesystemen ontwikkelen en verkopen (zoals TomTom) verzamelen grote hoeveelheden voertuigtrajectoriën en ook in het kader van onderzoek worden steeds vaker zeer gedetailleerde trajectoriën verzameld.

1.2.2 Intensiteit, dichtheid en snelheid

De drie belangrijkste *macroscopische verkeerskenmerken* zijn de dichtheid k , de intensiteit q en de gemiddelde snelheid u . Doorgaans wordt de intensiteit q gedefinieerd als "het gemiddeld aantal voertuigen dat gedurende een bepaalde periode een dwarsdoorsnede passeert". Volgens deze definitie is de intensiteit dus een *lokale variabele*. Er geldt:

$$q = \frac{n}{T} = \frac{n}{\sum_{i=1}^n h_i} = \frac{1}{H}$$

VERGELIJKING 1.1

waarin (tussen haakjes de S.I. eenheid):

q (vtg/s)	gemiddelde intensiteit gedurende periode met lengte T
n	aantal voertuigen dat doorsnede passeert gedurende periode met lengte T
h_i (s)	volgtijd voertuig i ten opzicht van voertuig $i-1$
T (s)	lengte periode
H (s)	gemiddelde volgtijd

Daarnaast kennen we de dichtheid k , die gedefinieerd is als "het gemiddeld aantal voertuigen in een bepaald gebied op een tijdstip". De dichtheid is volgens deze definitie een zogenaamde *instantane variabele*. Er geldt:

$$k = \frac{m}{X} = \frac{m}{\sum_{i=1}^m s_i} = \frac{1}{S}$$

VERGELIJKING 1.2

waarin:

k (vtg/m)	gemiddelde dichtheid op traject met lengte X
m	aantal voertuigen op traject met lengte X
s_i (m)	volgafstand voertuig i ten opzicht van voertuig $i-1$
X (m)	lengte traject
S (m)	gemiddelde volgafstand

Intensiteit is dus gelijk aan 1 gedeeld door de gemiddelde volgtijd en dichtheid is gelijk aan 1 gedeeld door de gemiddelde (bruto²) volgafstand. Het is niet zo moeilijk om te bedenken dat als twee voertuigen met een constante snelheid achter elkaar rijden, de volgende relatie moet gelden: volgafstand = volgtijd × snelheid, oftewel

$$s_i = h_i u$$

VERGELIJKING 1.3

Als je veronderstelt dat de verkeersstroom als geheel constant is over een stukje van de ruimte en tijd (homogeen en stationair), dan moet ook gelden intensiteit = dichtheid × gemiddelde snelheid (check de eenheden!), oftewel

$$q = ku$$

VERGELIJKING 1.4

Deze zogenaamde continuïteitsrelatie (gemakkelijk te onthouden met "kuu=kuu") is belangrijk voor zowel de eenvoudige berekeningen aan verkeerstromen die we in dit inleidende vak doen, maar ook voor veel geavanceerdere berekeningen en analyses. Als je de dichtheid en de intensiteit weet, kun je de gemiddelde snelheid met vergelijking 1.4 dus gemakkelijk berekenen ($u=q/k$). Maar je kunt, kijkend naar Figuur 1.2, op nog twee andere manieren een gemiddelde snelheid bepalen: of op een *dwarsdoorsnede voor een periode* (de zogenaamde tijd-gemiddelde of Lokale snelheid),

² Bruto volgafstand is de volgafstand inclusief de lengte van het voertuig zelf

$$u_L = \frac{1}{n} \sum_{i=1}^n v_i$$

of voor een *stuk weg op een bepaald tijdstip* (de zogenaamde ruimtelijk-gemiddelde of **M**omentane snelheid)

$$u_M = \frac{1}{m} \sum_{i=1}^m v_i$$

Deze snelheden zijn, zelfs indien de verkeerscondities niet sterk fluctueren in tijd en ruimte, wezenlijk verschillend van elkaar. Er geldt bijvoorbeeld $u_L \geq u_M$, waarbij het verschil tussen de twee in congestie soms oploopt tot meer dan 25%. In de praktijk wordt het onderscheid echter zelden gemaakt, waardoor ongemerkt zeer aanzienlijke fouten worden gemaakt. Bijvoorbeeld in de verkeersinformatie die je op informatiepanelen boven de weg of op de radio voorgeschoteld krijgt. Welke van deze twee mag je nu gebruiken in vergelijking 1.4? Je kunt bewijzen dat die vergelijking alleen maar opgaat als je de ruimtelijk gemiddelde snelheid u_M gebruikt. Om u_M te bepalen moet je dus ofwel alle voertuigtrajectoriën weten of met een helikopter boven een weg hangen en snapshots maken. Dat is nogal prijzig en bewerkelijk. Daarom wordt die ruimtelijk gemiddelde snelheid vaak geschat met de zogenaamde *harmonisch* gemiddelde snelheid die je wel lokaal kunt berekenen

$$u_M = \frac{q}{k} \approx \left(\frac{1}{n} \sum_{i=1}^n \frac{1}{v_i} \right)^{-1}$$

VERGELIJKING 1.5

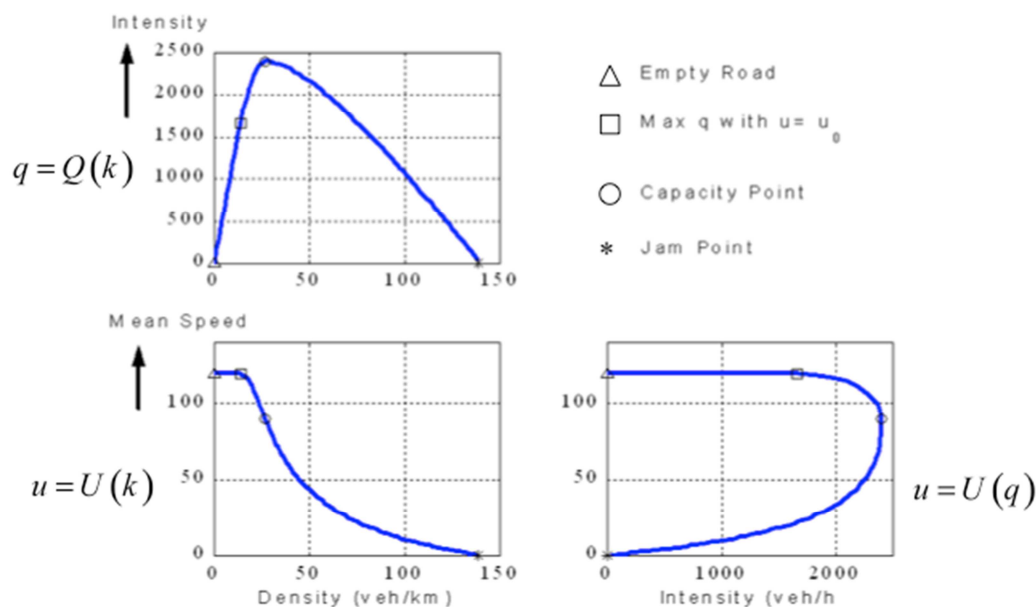
Die harmonisch gemiddelde snelheid is eigenlijk niets anders dan 1 gedeeld door de gemiddelde *traagheid* ($1/v_i$) en is exact gelijk aan de ruimtelijk gemiddelde snelheid als de verkeersstroom werkelijk homogeen en stationair is (maar dat is zelden het geval).

1.3 Fundamenteel diagram en capaciteit

Wat we direct uit vergelijking 1.4 kunnen lezen is dat als voertuigen bij gelijke snelheid dichter op elkaar gaan rijden (k neemt toe), de volgtijd noodzakelijkerwijs kleiner wordt (q wordt groter). Je zou dan geneigd zijn te denken dat q dus altijd groter wordt (hogere doorstroom!) als je de snelheid of de dichtheid (of allebei) maar verder opvoert. Misschien dat Johan Cruijff daarom vermoedde dat je het fileprobleem kunt oplossen door de maximum snelheid naar 200 km/u te verhogen. Het blijkt echter dat er een bepaalde snelheid is waarbij de intensiteit niet meer groter *kan* worden als de dichtheid toeneemt. De reden is dat de gemiddelde volgtijd dan zo klein zou worden dat botsingen onvermijdelijk worden. Dat heeft te maken met de reactietijd die bestuurders nodig hebben om visuele informatie te verwerken en deze om te zetten in actie (bv. op de rem stappen om afstand te houden) en het feit dat (niet-suïcidale) bestuurders om die reden een bepaalde minimale volgtijd zullen aanhouden. Wordt de volgafstand toch kleiner (dichtheid te groot) dan moeten bestuurders hun snelheid wel verminderen. Er is dus een 'optimale' combinatie van snelheid en dichtheid waarbij de gemiddelde volgtijd minimaal is en daarmee de intensiteit maximaal. Die "optimale" snelheid blijkt in Nederland op snelwegen zo rond de 85 à 90 km/u te liggen. Bij die snelheid durven Nederlandse bestuurders gemiddeld nog zo'n 1.6 seconde volgtijd aan te houden wat neer komt op een maximale intensiteit van gemiddeld zo'n 2250 voertuigen per uur per rijstrook. We noemen dit wel de capaciteit C van een weg. De relatie tussen intensiteit en dichtheid (of volgtijd en volgafstand) $q = Q(k)$ noemen we het fundamenteel diagram.

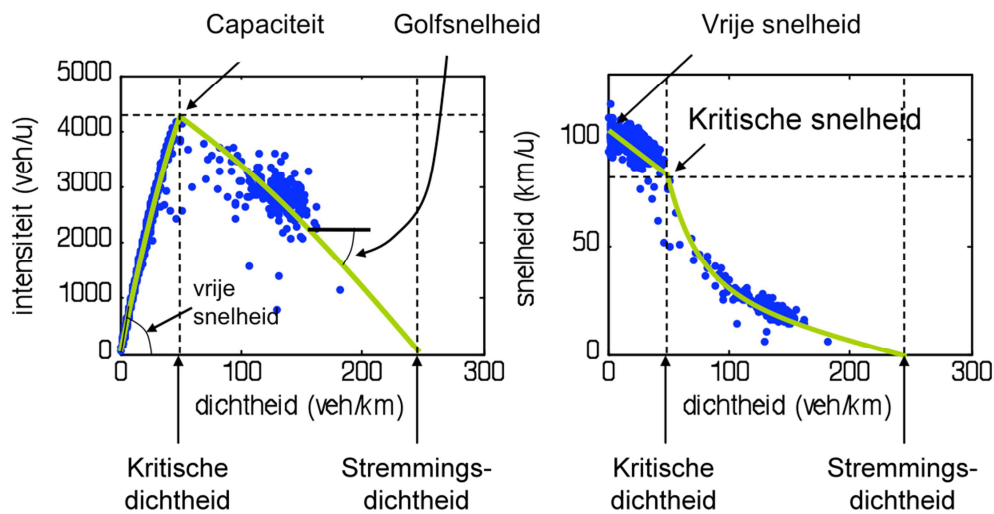
1.3.1 Kenmerken van het Fundamenteel diagram

Het fundamenteel diagram beschrijft een statistische relatie tussen de intensiteit, dichtheid en snelheid. In feite zijn er verschillende verschijningsvormen, maar het verband $q = Q(k)$ tussen de dichtheid k en de intensiteit q wordt in de verkeersstroomtheorie het meest gehanteerd. De andere vormen ($u = U(k)$, $u = U(q)$) kunnen met $q = ku$ eenvoudig worden afgeleid. Figuur 1.3 laat alle drie de verschijningsvormen zien en een aantal kenmerkende parameters. Deze zullen we iets verder beneden nader toelichten aan de hand van Figuur 1.4. Hoewel het snelheids-intensiteits diagram ($u=U(q)$, Figuur 1.3 rechtsonder) in de praktijk erg populair is – onder andere omdat zowel snelheid als intensiteit meestal direct kunnen worden gemeten, wordt deze vorm bij modelberekeningen nauwelijks toegepast. Dat komt omdat $u=U(q)$ geen functie is (geen 1-op-1 afbeelding), en de andere twee wel.



FIGUUR 1.3: DRIE VERSCHIJNINGSVORMEN VAN HET FUNDAMENTEEL DIAGRAM: $Q = Q(k)$, $U = U(k)$ EN $U = U(q)$.

De premisse die aan het fundamenteel diagram (in welke verschijningsvorm dan ook) ten grondslag ligt, is dat *bestuurders zich ongeveer hetzelfde gedragen in vergelijkbare verkeerscondities*. Als het verkeer zich bijvoorbeeld met een bepaalde snelheid u afwikkelt, dan is het aannemelijk dat de automobilisten ongeveer dezelfde volgafstand $S = 1/k$ aanhouden. Dit gedrag – en daarmee de relatie tussen in dit geval snelheid en dichtheid – is natuurlijk wel afhankelijk van zaken als het weer, kenmerken van de weg, de samenstelling van het verkeer (bumperklevers versus oude dames), heersende verkeersregels, etc. Figuur 1.4 toont een typisch voorbeeld van de relatie tussen de intensiteit en de dichtheid en die tussen snelheid en dichtheid voor een tweestrooks autosnelweg.



FIGUUR 1.4: FUNDAMENTEEL DIAGRAM $Q = Q(k)$ EN $U = U(k)$ EN VERKEERSDATA A9 (TER HOOGTE VAN ROTTEPOLDERPLEIN S17), OKTOBER 1994.

In Figuur 1.4 zijn de belangrijkste punten in het fundamenteel diagram weergegeven: de capaciteit C (punt waarbij de maximale intensiteit wordt verwezenlijkt), de kritische dichtheid k_c en kritische snelheid u_c , de stremmingsdichtheid k_{jam} (de maximale dichtheid als de voertuigen stilstaan) en tot slot de vrije snelheid of wenssnelheid u_0 . In de figuur zijn ook enkele empirische waarnemingen opgenomen gemeten met lussen. In de figuur is een duidelijk onderscheid te zien tussen de vrije tak van het fundamenteel diagram ($k < k_c$) en de congestietak ($k > k_c$). Zodra de intensiteit de capaciteit bereikt ontstaat congestie (file), hierover verder beneden meer.

Het is verleidelijk *causaliteit* toe te dichten aan het fundamenteel diagram: vaak wordt gesteld dat de relatie $u = U(k)$ weergeeft dat bij een toenemende dichtheid de snelheid afneemt. Met andere woorden, de snelheid van een voertuig in de verkeersstroom neemt af als de voertuigen dichter op elkaar gaan rijden. Redeneren we vanuit het rijgedrag (vergelijking 1.5), dan ligt het echter meer voor de hand te stellen dat de volgfstand tot de voorligger wordt bepaald door de snelheid van de voorligger: de dichtheid is dan een functie van de snelheid en niet andersom. In de vrije tak van het diagram is het echter het toenemend aantal interacties tussen snelle en langzame voertuigen wat resulteert in een reductie van de gemiddelde snelheid. Hier zouden we dus kunnen zeggen dat de snelheid een functie is van de intensiteit.

1.3.2 Vrije snelheid, capaciteit en andere fundamenteel diagramparameters

Op een lege weg zal elke bestuurder rijden met een snelheid die op die plek en tijd naar eigen inzicht het prettigst aanvoelt. Deze wenssnelheid of vrije snelheid u_0 verschilt heel sterk van persoon tot persoon. Dat heeft te maken met de kenmerken van het voertuig (vrachtwagen versus Ferrari); de plaatselijk geldende verkeersregels (maximum snelheid) en de mate waarin de bestuurder zich daar aan houdt; het reismotief (vliegtuig halen, naar werk of een dagje uit); allerlei lokale omstandigheden (regen, zon of mist) en allerlei persoonsgebonden factoren zoals geslacht, leeftijd, ervaring, fysieke en mentale gesteldheid, enzovoorts. Er is dus een brede verdeling van wenssnelheden die ook nog eens plaats en locatie afhankelijk is. De puntenwolk helemaal aan de linkerkant van de $U(k)$ relatie in Figuur 1.4 (rechts) geeft die spreiding goed weer. Zodra er zich meer verkeer op een weg bevindt kan niet elke bestuurder zich voortbewegen met zijn wenssnelheid – tenzij die toevallig voor iedereen precies even groot is (en zelfs dan kan het niet – dat zien we verder beneden). Er zijn dan twee mogelijkheden: (1) een bestuurder houdt een zekere afstand tot (en *volgt*) degene die voor hem of haar rijdt (met ongeveer dezelfde snelheid); (2) een bestuurder passeert het voertuig. Je hebt daarmee dus twee typen bestuurders in een verkeersstroom: volgers en leiders. Natuurlijk zal een bestuurder regelmatig van rol verwisselen zodra de gelegenheid zich daartoe voordoet.

Capaciteit C is ook een parameter uit het fundamenteel diagram die sterk stochastisch is. Capaciteit is immers gelijk aan *1 gedeeld door de gemiddelde minimale volgtijd* die bestuurders nog aan durven te houden bij de gegeven snelheid, d.w.z.

$$C = \frac{1}{H_{\min}}$$

VERGELIJKING 1.6

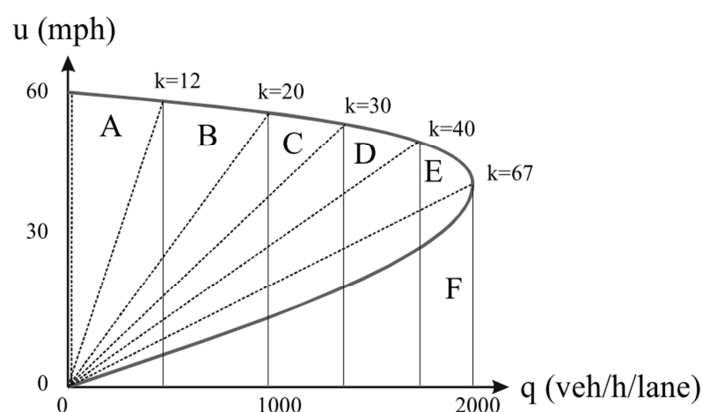
Het is belangrijk om in het achterhoofd te houden dat capaciteit dus het resultaat is van collectief menselijk gedrag, en niet een eigenschap van een stuk asfalt. Dat verklaart ook waarom capaciteit zo sterk wordt beïnvloed door verschillende omstandigheden. In slecht weer rijden (verstandige) mensen gemiddeld met grotere volgtijden en/of lagere snelheden, datzelfde zie je bij wegversmallingen en langs wegwerkzaamheden. Daarbij spelen – behalve “verstand” – vooral factoren mee die de reactietijd van bestuurders onbewust beïnvloeden, zoals een zwaardere rijtaak door slechter zicht of minder comfort, een groter gepercipieerd risico of door onbekendheid met de situatie. De capaciteit verschilt daarnaast natuurlijk voor verschillende wegtypen (snelweg, provinciale weg, stadsweg) met verschillende verkeersregels, geometrie en maximum snelheden. Onderzoek laat tenslotte ook zien dat de capaciteit sterk beïnvloed wordt door sociaal-culturele verschillen. In China is de snelwegcapaciteit bijv. ruim 20% lager dan in de Randstad, wat onder andere te maken heeft de verschillen in rijvaardigheid, rijstrookverdeling en reactietijden.

1.3.3 Toepassingen van het fundamenteel diagram

Voordat we gaan kijken hoe we de basisbegrippen uit deze en de vorige paragraaf kunnen gebruiken om te verklaren hoe files ontstaan, zullen we kort verschillende toepassingen van het fundamenteel diagram de revue laten passeren. De lijst is zeker niet uitputtend, maar geeft wel een goed overzicht van de verschillende applicaties.

Definitie van serviceniveaus

Serviceniveaus worden gebruikt als manier om de kwaliteit van de verkeersafwikkeling te kenmerken, en ook om bij het ontwerpen van infrastructuur het gewenste service niveau te bepalen. Als je een wegverbinding tussen Delft en Schiedam wilt maken met een bepaald service niveau en een bepaalde ontwerpsnelheid, dan kun je het fundamenteel diagram gebruiken om te schatten hoeveel rijstroken die weg zou moeten hebben. Onderstaande figuur geeft een voorbeeld uit de (HCM³, 2001) van deze toepassing.

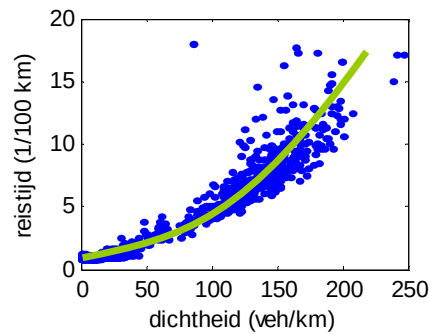


FIGUUR 1.5: DEFINITIE VAN SERVICENIVEAUS (MAY, 1990).

³ De Highway Capacity Manual (HCM) is het handboek voor het ontwerp van wegverkeersinfrastructuur

Afleiden van reistijd- en kostenfuncties

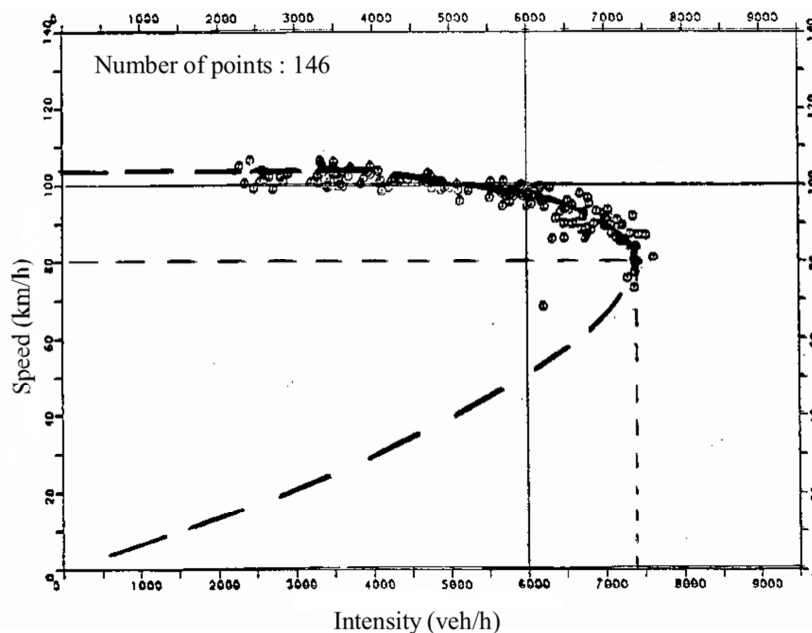
Voor veel toepassingen zijn reistijdfuncties van groot belang. Een reistijdfunctie beschrijft de verwachte reistijd voor een wegsegment als functie van de dichtheid of de intensiteit. Reistijd wordt vaak gehanteerd als maat voor de kosten van in de file staan en met een reistijdfunctie kun je bijvoorbeeld uitrekenen wat de kosten zijn van files⁴. Figuur 1.6 geeft een voorbeeld van reistijdfuncties afgeleid van het fundamenteel diagram. In de figuur wordt de reistijd $TT = 100/u$ per 100 km weergegeven als functie van de dichtheid.



FIGUUR 1.6: REISTIJDFUNCTIE AFGELEID VAN HET FUNDAMENTEEL DIAGRAM.

Toepassing binnen dynamische afwikkelingsmodellen

Veel dynamische macroscopische afwikkelingsmodellen (METANET, MADAM, Integration, FlowSim en FastLane) maken gebruik van het fundamentele diagram om de afwikkeling te voorspellen. We zullen zo meteen zien dat ook we ook bij de allersimpelste manier om naar files te kijken (schokgolftheorie) gebruik maken van het fundamenteel diagram.



FIGUUR 1.7: VOORBEELD VAN HET SCHATTEN VAN DE CAPACITEIT MET BEHULP VAN HET FUNDAMENTEEL DIAGRAM (HOOGENDOORN, 2007).

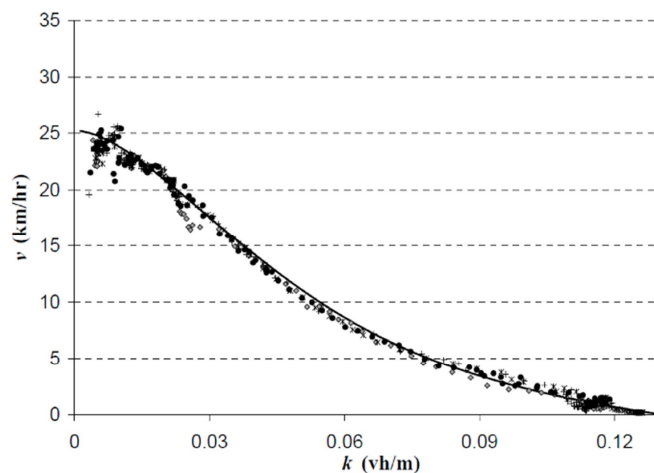
⁴ Zo becijfert het Kennisinstituut voor Mobiliteit dat de totale kosten van in de file staan in 2006 zo rond de €2 miljard bedroegen.

Schatten van de capaciteit

Een volgend voorbeeld van een toepassing van het fundamenteel diagram is het bepalen van de capaciteit. In (Botma e.a.,1998) wordt deze methode toegepast om het effect van regen en verlichting op de capaciteit te bepalen (zie Figuur 1.7).

Alternatieve toepassingen fundamentele relaties

Inmiddels zijn verschillende wetenschappers bezig met 'alternatieve fundamentele diagrammen'. Zo stelde Tu (2008) een fundamenteel diagram voor reistijdbetrouwbaarheid voor en leidde Daganzo (2008) het fundamenteel diagram voor een volledig stedelijk netwerk af (zie Figuur 1.8). De figuur toont de gemiddelde operationele snelheid als functie van de dichtheid. Opvallend is hoe weinig fluctuaties zichtbaar zijn, met name voor hogere dichtheden.

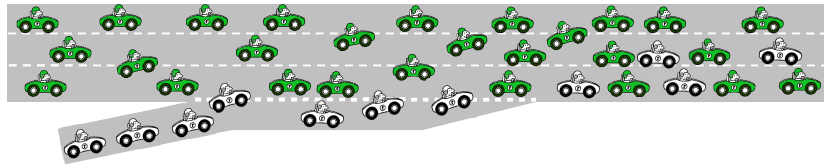


FIGUUR 1.8: FUNDAMENTEEL DIAGRAM VAN SAN FRANCISCO (DAGANZO EN GEROLIMINIS, 2008)

In het laatste geval lijkt het er zelfs op dat de afwijking op het geaggregeerde netwerk niveau (i.p.v. op linkniveau) zich erg goed laat beschrijven door een fundamenteel diagram: diverse tests hebben aangetoond dat de gemiddelde operationele snelheid in congestieve stedelijke gebieden zich goed laat beschrijven als functie van de gemiddelde dichtheid in het gebied, zelfs indien op linkniveau geen sprake lijkt van een duidelijk verband. Voor netwerk management kunnen dergelijke nieuwe inzichten belangrijke hulpmiddelen zijn.

1.4 Hoe files ontstaan – een kwalitatieve analyse

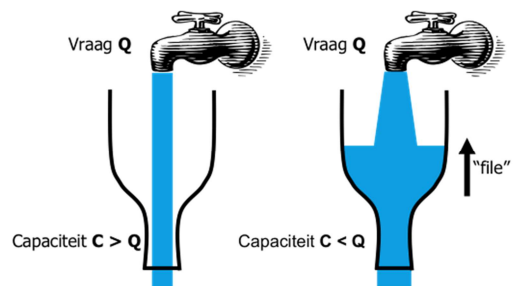
Laten we eerst met behulp van deze begrippen kwalitatief beschrijven en verklaren waarom er file op de A13 ontstaat. Stel, op de hoofdrijbaan is de intensiteit hoog en de voertuigen rijden dicht op elkaar. Stel vervolgens dat de intensiteit op de oprit ook hoog is. Die voertuigen moeten allemaal invoegen op de hoofdrijbaan, waar de ruimte (door kleine volgafstanden) toch al beperkt is. Door in te voegen worden de volgafstanden tijdelijk nog kleiner. Vaak zie je dat bij de invoegstrook zelf voertuigen nog wel bereid zijn om gedurende heel korte tijdsperioden nog kleinere volgafstanden aan te houden, maar dat een klein stukje benedenstreams van de oprit de meeste bestuurders weer proberen terug te keren naar een veiliger volgafstand. Dat kan dus alleen door de snelheid te verlagen, d.w.z. door te remmen. Vervolgens ontstaat er een kettingreactie. Het voertuig dat het remmende voertuig volgt zal nu ook remmen, vaak nog net iets harder om ook zijn volgafstand weer terug te brengen naar een veilige marge. Het voertuig daarachter doet hetzelfde, enzovoorts. Het resultaat is dat de golf remlichten zich snel verplaatst tegen de richting van het verkeer in, terwijl de snelheid van het verkeer zelf sterk naar beneden gaat, soms zelfs tot volledige stilstand. De verkeersafwikkeling "stort in" zoals dat heet.



FIGUUR 1.9: CONGESTIE ONTSTAAT BIJ FLESSENHALZEN, BIJVOORBEELD EEN OPRIT.

1.4.1 Files en flessenhalzen

We kunnen dit proces goed vergelijken met wat er gebeurt als we een fles zonder bodem ondersteboven onder een waterkraan te houden. Zodra we meer water uit de kraan laten stromen (intensiteit) dan er door de flessenhals heen kan (capaciteit) neemt het waterniveau in de fles toe tot deze overstroomt. Er ontstaat als het ware file in de fles, want de snelheid van het water in de fles boven de flessenhals neemt sterk af. Draaien we de kraan nu zo ver dicht dat er minder water in de fles bij komt dan er door de flessenhals weer uit kan, dan zal de file in de fles "oplossen", oftewel, de fles loopt leeg. Blijft er meer water uit de kraan komen dan er door de flessenhals heen kan, dan stroomt onvermijdelijk de fles over. Vertaald naar de verkeerssituatie bij een oprit geldt dus: is de som van de intensiteit op de hoofdrijbaan en een afrit groter dan de capaciteit van de hoofdrijbaan dan ontstaat er file, en die file blijft groeien totdat die som van deze intensiteiten kleiner wordt dan de capaciteit van de hoofdrijbaan. Het onderliggende mechanisme is dat bestuurders een minimale veilige volgtijd (of volgafstand) kiezen, wat bij een te grote intensiteit alleen maar mogelijk is door op de rem te stappen.



FIGUUR 1.10: ANALOGIE TUSSEN FILE IN EEN VERKEERSNETWERK EN 'FILE' IN EEN OMGEKEERDE WATERFLES

De analogie tussen verkeersstromen en hydrodynamica (de wetenschap die zich bezighoudt met de dynamica van vloeistoffen) komt verkeerskundigen goed van pas. Veel van de macroscopische modellen waarmee we beschrijven hoe files zich ontwikkelen zijn gebaseerd op hydrodynamische modellen. Maar voorzichtigheid is geboden, want de analogie gaat maar heel beperkt op:

- Voertuigen houden – in tegenstelling tot watermoleculen – niet van botsen en bewegen zich allemaal in dezelfde richting; watermoleculen bewegen alleen *netto* in een bepaalde richting.
- Verkeer is anisotroop, dat wil zeggen dat de interactie tussen "deeltjes" (voertuigen) richting heeft: voertuigen reageren veel sterker (bijna uitsluitend) op voorliggers.
- Voertuigen en hun bestuurder hebben een eigen willetje, wat veel van de fundamentele variabelen zeer stochastisch maakt: capaciteit is bijvoorbeeld een stochastische variabele die sterk wordt bepaald door rijgedrag, in tegenstelling tot de capaciteit in het waterflesvoorbeeld, die een fysieke eigenschap is van de "infrastructuur" (de fles).
- Voertuigen nemen relatief aan het macroscopische fenomeen (de file) veel ruimte in, ook nog eens afhankelijk van hun snelheid. Watermoleculen zijn verwaarloosbaar klein ten opzichte van de 'file in de fles'.
- Er zijn veel, maar dan ook echt heel veel meer watermoleculen in een fles dan voertuigen in een file. Er zijn zelfs heel veel meer watermoleculen in een fles dan voertuigen op de hele aardbol. Dit betekent dat begrippen als de gemiddelde dichtheid, intensiteit en snelheid in een fles water heel anders moeten worden geïnterpreteerd dan in een verkeerssysteem. Je zou zelfs kunnen beargumenteren dat de continue benadering van verkeer om deze reden fundamenteel onjuist is.

1.4.2 De file als golf

Op de brug bij de A13 zagen we al dat files zich gedragen als golven. Als er benedenstrooms iemand op de rem stapt en er file ontstaat bewegen de remlichten als een soort golf naar achteren (tegen het verkeer in). In de jaren vijftig is gelijktijdig door de heren Lighthill & Witham en Richards de zogenaamde kinematische golftheorie ontwikkeld waarmee de dynamica van verkeersstromen als golfverschijnsel kan worden beschreven. Dat model bestaat uit (slechts) drie eenvoudige vergelijkingen, te weten (1) het al eerder genoemde fundamenteel diagram; (2) de continuïteitsrelatie ($q=ku$), en (3) de zogenaamde "wet van behoud van voertuigen", die vrij vertaald als volgt luidt

$$\text{Aantal voertuigen (straks)} = \text{Aantal voertuigen (nu)} + \text{instroom (nu)} - \text{uitstroom (nu)}$$

Als je in plaats van "Aantal voertuigen" de termen "Hoeveelheid water" invult is het direct duidelijk dat deze formule beschrijft wat er in de omgekeerde fles gebeurt als we de kraan openzetten. De verandering van het aantal voertuigen op een bepaald stukje weg (de dichtheid op die weg) wordt dus bepaald door de hoeveelheid voertuigen die er al waren plus wat er op dat stukje weg netto gezien in- en uitstroomt. Zonder in details te treden over de precieze achtergrond en afleiding, kun je deze behoudswet dan ook als een partiële differentiaalvergelijking opschrijven:

$$\frac{\partial k}{\partial t} = -c(k) \frac{\partial q}{\partial x} \quad \text{VERGELIJKING 1.7}$$

Wat vrij vertaald neerkomt op

Verandering in dichtheid over tijd = - Golfsnelheid × Verandering in Intensiteit over de ruimte
Die golfsnelheid $c(k)$ is ook een functie van de dichtheid, het is namelijk de afgeleide $dQ(k)/dk$ van (en dus de raaklijn aan) het fundamenteel diagram (zie Figuur 1.4 links). In vrije afwikkeling is die afgeleide positief: verstoringen bewegen met het verkeer mee (denk aan een colonne vrachtwagens die plaatselijk voor een iets lagere snelheid en hogere dichtheid zorgt). Als de weg helemaal leeg is en ieder voertuig is vrij om zijn snelheid te kiezen dan geldt $c(k) = u_0$ (golfsnelheid = gemiddelde vrije snelheid), vlak voor de capaciteit wordt bereikt geldt $0 < c(k) \leq u_c$. In file omstandigheden is $dQ(k)/dk$ negatief, dan bewegen verstoringen dus *tegen de richting van het verkeer* (denk maar aan de golf remlichten die tegen het verkeer in beweegt). Vaak wordt voor de congestietak van het fundamentele diagram een lineair verband verondersteld, in welk geval de golfsnelheid in congestie constant wordt en als volgt uit het fundamenteel diagram in Figuur 1.4 (links) kan worden bepaald:

$$c(k) = \frac{C}{k_{jam} - k_c}, k > k_c \quad \text{VERGELIJKING 1.8}$$

Deze 'filegolfsnelheid' bepaalt onder specifieke condities⁵ (die overigens veel voorkomen) met welke snelheid een file groeit. In de volgende paragraaf over schokgolftheorie zullen we nader ingaan op hoe dit precies werkt en wat die condities zijn. Uit de praktijk blijkt in elk geval dat die veronderstelling van een lineair fundamenteel diagram helemaal zo gek niet is. Het blijkt namelijk dat deze filegolfsnelheid inderdaad behoorlijk constant is en op Nederlandse snelwegen ergens tussen de -20 en -15 km/u ligt. Dat betekent dus dat een file in een uurtje tijd 15 tot 20 km kan groeien.

Het is eigenlijk verbazingwekkend hoe aardig deze kinematische golftheorie beschrijft en verklaart wat er op wegennetwerken gebeurt. Maar er zijn natuurlijk ook veel dingen die de kinematische golftheorie niet volledig of zelfs helemaal niet verklaart. Één verschil is bijvoorbeeld dat volgens de theorie de file altijd *voor* de bottleneck ontstaat en niet een paar honderd meter benedenstrooms (zoals in de praktijk vaak gebeurt). Een ander voorbeeld is dat het model ook niet verklaart –

⁵ Deze specifieke condities houden in dat in de file sprake is van stremmingsdichtheid (k_{jam}), met andere woorden, dat voertuigen in de file stilstaan.

althans niet zonder de nodige aanpassingen en verfijningen – dat files zichzelf verergeren. Dit komt door de zogenaamde capaciteitsval, dat wil zeggen, zodra er file ontstaat neemt direct de capaciteit af waardoor de file veel langzamer oplost. Het gaat dan dus van kwaad tot erger, dit zullen we verderop nog iets uitgebreider bespreken. Zo zijn er nog veel meer zaken die zonder geavanceerdere modellen niet goed kunnen worden beschreven en verklaard met kinematische golftheorie. Niettemin, zelfs als we deze theorie nog verder versimpelen (tot we alleen schokgolftheorie overhouden) kunnen we die voor veel praktische toepassingen gebruiken.

1.5 Schokgolftheorie

Een 'schokgolf' is de grens tussen twee verschillende verkeerstoestanden die worden gekenmerkt door verschillende dichtheden, snelheden en/of intensiteiten. Zo'n toestand is daarmee een punt $\{k, q\}$ in het fundamenteel diagram. Schokgolftheorie beschrijft de dynamica van een schokgolf, m.a.w. hoe beweegt de grens tussen (stationaire en homogene) verkeerstoestanden zich over de ruimte als functie van de tijd. De snelheid van een schokgolf is eenvoudig te bepalen.

1.5.1 Analytische afleiding schokgolfsnelheid

Veronderstel dat een homogene verkeersstroom met dichtheid k_1 en snelheid u_1 over een weg beweegt en een eveneens homogene stroom voertuigen met (hogere) dichtheid k_2 en (lagere) snelheid u_2 "tegenkomt" (zie Figuur 1.11). Laten we veronderstellen dat de grens S (de "schok") tussen deze verkeerstoestanden zich beweegt met een snelheid ω . We weten dat de hoeveelheid verkeer dat de schok instroomt gelijk moet zijn aan de hoeveelheid verkeer die er uitstroomt (behoud van voertuigen). Met andere woorden

$$q_s^{in} = q_s^{uit} \quad \text{VERGELIJKING 1.9}$$

De intensiteit van het verkeer dat de schok inrijdt is gelijk aan de dichtheid stroomopwaarts van de schok maal de snelheid *relatief aan de schok*, d.w.z.

$$q_s^{in} = k_1 (u_1 - \omega) \quad \text{VERGELIJKING 1.10}$$

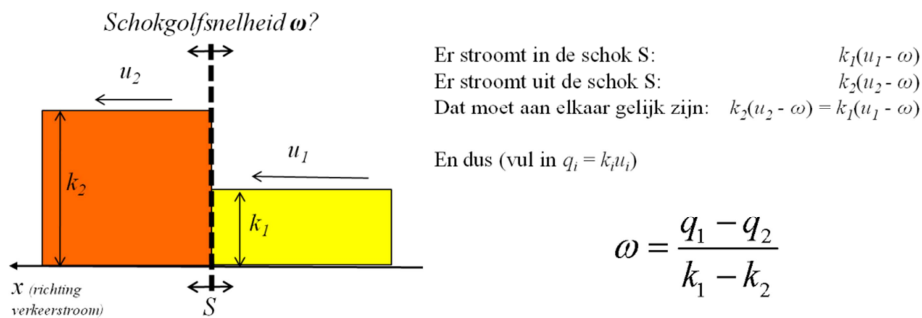
Evenzo geldt voor de intensiteit van verkeer dat de schok uitrijdt

$$q_s^{uit} = k_2 (u_2 - \omega) \quad \text{VERGELIJKING 1.11}$$

Gezien vergelijking 1.9 mogen we 1.10 en 1.11 gelijkstellen. Dit levert dan op voor de schokgolfsnelheid

$$\omega = \frac{q_1 - q_2}{k_1 - k_2} \quad \text{VERGELIJKING 1.12}$$

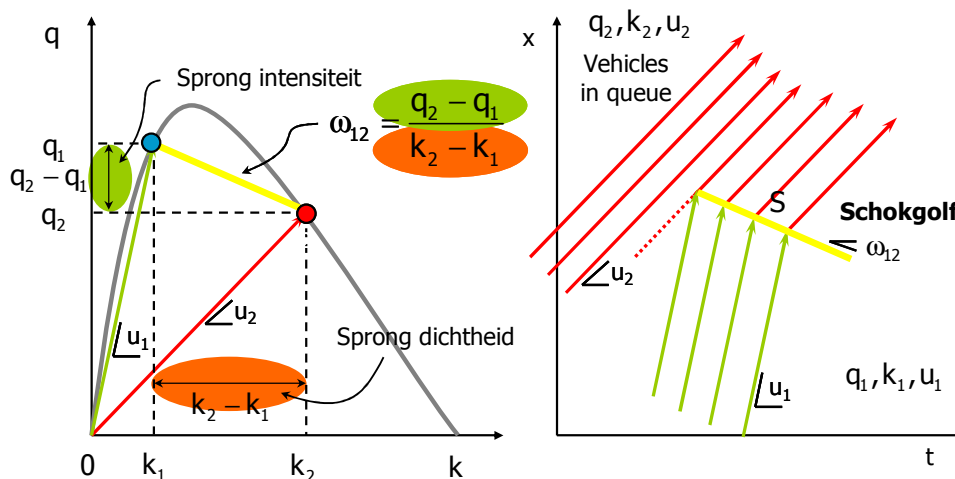
In woorden: de snelheid van een schokgolf tussen twee verkeerstoestanden wordt bepaald door het verschil in intensiteit te delen door het verschil in dichtheid tussen die toestanden.



FIGUUR 1.11: SCHEMATISCHE AFLEIDING SCHOKGOLFSNELHEID.

1.5.2 Grafische interpretatie

Het aardige is dat deze relatie ook een eenvoudige grafische interpretatie heeft. Het is eenvoudig in te zien dat de schokgolfsnelheid in vergelijking 1.12 correspondeert met de lijn tussen twee verkeerstoestanden in het q-k diagram. Figuur 1.12 (links) illustreert dit principe. Figuur 1.12 (rechts) laat vervolgens zien hoe op grond daarvan in een x-t diagram voertuigtrajectoriën en schokgolven (grenzen tussen de verschillende verkeerstoestanden) kunnen worden getekend.



FIGUUR 1.12: GRAFISCHE INTERPRETATIE VAN SNELHEID SCHOKGOLF

1.6 Toepassingen schokgolftheorie

We kunnen schokgolftheorie toepassen om de verkeerssituatie te voorspellen voor tal van situaties. Het recept daarvoor is vrijwel altijd gelijk. Allereerst identificeren we de mogelijke verkeerstoestanden (punten op het fundamenteel diagram). Vervolgens verbinden we de punten in het fundamenteel diagram om de schokgolfsnelheden tussen al die verschillende toestanden te bepalen. Tenslotte kunnen we de toestanden en de schokgolven tussen die toestanden gemakkelijk in een ruimte-tijd diagram tekenen. In het voorbeeld hieronder gebruiken we schokgolftheorie om de filelengte voor een bottleneck op een snelweg te berekenen. In paragraaf 2.1.3 zullen we kijken naar de toepassing van schokgolftheorie op een geregeld kruispunt.

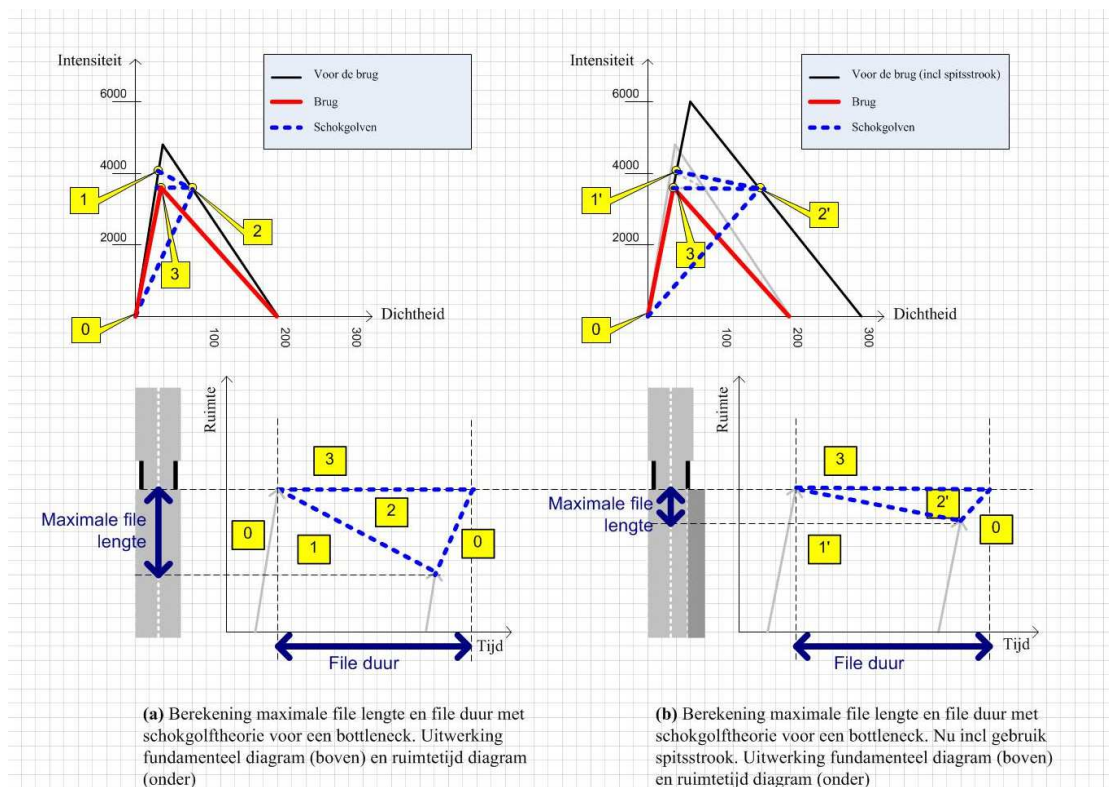
1.6.1 Berekening van de filelengte bij een bottleneck – grafische uitwerking

In een tweestrooksrijbaan van een autosnelweg bevindt zich een brug met smallere rijstroken en kleinere capaciteit dan op het baanlichaam. Men overweegt nu om in de spitsuren de vluchtstrook als spitsstrook te gebruiken, omdat – zo vermoedt men, de dagelijkse file dan minder lang zal zijn en daarmee sneller zal zijn opgelost. Op de brug ontbreekt overigens de vluchtstrook. Tabel 1.1 geeft voor zowel de hoofdrijbaan als de brug de parameterwaarden voor het fundamenteel diagram op beide locaties. We veronderstellen dat dit fundamenteel diagram "driehoekig", d.w.z. een lineair q-k verband in zowel de vrije als de congestietak. We kunnen er van uitgaan dat de intensiteit in de spits 4200 vtg/u bedraagt. Kunnen we op basis van schokgolftheorie nu voorspellen wat het gebruik van de vluchtstrook als spitsstrook daadwerkelijk oplevert? Lost de mogelijk kortere file dan daadwerkelijk sneller op?

	Capaciteit	Dichtheid bij capaciteit	Dichtheid bij stremming
Brug	3600 vtg/u	36 vtg/km	200 vtg/km
Rijbaan met twee stroken	4800 vtg/u	40 vtg/km	200 vtg/km
Idem + vlucht/spitsstrook	6000 vtg/u	60 vtg/km	300 vtg/km

TABEL 1.1: PARAMETERWAARDEN Q-K DIAGRAM BIJ HET VOORBEELD "FILETERUGSLAG OP EEN SNELWEG"

Dit vraagstuk is het eenvoudigst aan te pakken door de verkeerstoestanden simpelweg te tekenen in zowel een fundamenteel diagram als een ruimte-tijd diagram. Figuur 1.13 illustreert deze methode voor zowel de huidige situatie als de situatie met gebruikmaking van de spitsstrook.



FIGUUR 1.13: UITWERKING VOORBEELD 1 "FILE VOOR EEN BOTTLENECK". DE SPITSSTROOK ZORGT WELISWAAR VOOR EEN KORTERE FILE, MAAR HET DUURT EVEN LANG VOOR DEZE OPLOST.

De conclusie zal alvast duidelijk zijn (zie de ruimte-tijd-diagrammen in Figuur 1.13 (a) en (b) onderaan): door de spitsstrook wordt de file inderdaad korter, maar hij lost niet sneller op. De spitsstrook verbreed als het ware het brede gedeelte van de omgekeerde fles (Figuur 1.10), maar omdat de flessenhals zelf even smal is gebleven duurt het even lang om dezelfde hoeveelheid verkeer af te wikkelen.

1.6.2 Berekening van de filelengte bij een bottleneck – nu met formules

We kunnen dit uiteraard ook uitrekenen. Eerst zonder spitsstrook. We weten met behulp van de waarden uit Tabel 1.1 dat de dichtheid van het verkeer stroomopwaarts (linkertak van het fundamentele diagram) en in de file (rechtertak) respectievelijk bedragen:

$$k_1 = \frac{4200}{4800} \times 40 = 35 \text{ vtg/km}$$

$$k_2 = 40 + \frac{4800 - 3600}{4800} \times (200 - 40) = 80 \text{ vtg/km}$$

De intensiteit in de file is gelijk aan de capaciteit van de bottleneck (3600 vtg/u). De file slaat dus terug met een (schokgolf)snelheid

$$\omega_{12} = \frac{4200 - 3600}{35 - 80} = -13.3 \text{ km/u}$$

Op dezelfde wijze kunnen we berekenen dat met spitsstrook de dichtheid voor en in de file bedraagt:

$$k_{1'} = \frac{4200}{6000} \times 60 = 42 \text{ vtg/km}$$

$$k_{2'} = 60 + \frac{6000 - 3600}{6000} \times (300 - 60) = 156 \text{ vtg/km}$$

En dat deze file (inderdaad veel langzamer) groeit met een snelheid

$$\omega_{1'2'} = \frac{4200 - 3600}{42 - 156} = -5.3 \text{ km/u}$$

De lezer kan zelf narekenen dat de andere variabelen en schokgolven de volgende waarden hebben in dit voorbeeld. Zonder spitsstrook vinden we

$$u_1 = \omega_{01} = 4800 / 40 = 120 \text{ km/u};$$

$$u_2 = 3600 / 80 = 45 \text{ km/u}$$

$$u_3 = 3600 / 36 = 100 \text{ km/u}$$

en met spitsstrook hebben we

$$u_{1'} = \omega_{01'} = 6000 / 60 = 100 \text{ km/u}$$

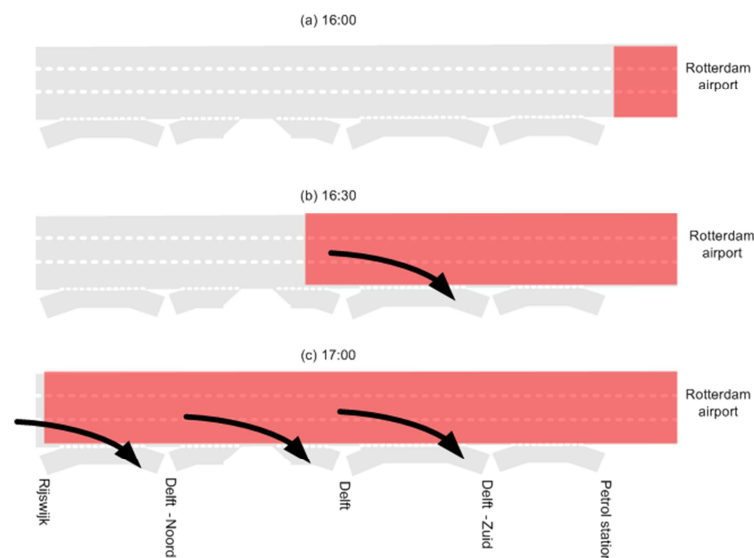
$$u_{2'} = 3600 / 156 = 23 \text{ km/u}$$

$$u_3 = 3600 / 36 = 100 \text{ km/u}$$

Zonder spitsstrook staat er dus na een uur ruim 13 km file en met spitsstrook iets meer dan 5 km. Dit is op zich al reden genoeg om de spitsstrook open te stellen. Het is zeer aannemelijk dat er stroomopwaarts van de brug andere opritten en afritten bestaan die geblokkeerd kunnen raken door de (lange) file. Dit noem je **fileterugslag**. Fileterugslag is een vervelend fenomeen, want

fileterugslag zorgt ervoor dat reizigers in files terechtkomen die niet gerelateerd zijn aan hun eigen route – ze “betalen” als het ware voor iets waar ze niet om hebben gevraagd. Bovendien zullen vanwege die extra reizigers de files nu ook nog eens sneller groeien. Een dubbel negatief effect dus. Dit is goed te zien in Figuur 1.14. De file die daar weergegeven is ontstaat bij Rotterdam en slaat zodanig ver terug dat ook reizigers met andere bestemmingen (bijvoorbeeld Delft) reistijd verliezen door deze file.

Er zijn natuurlijk ook nadelen verbonden aan het openstellen van de vluchtstrook als extra bufferruimte (spitsstrook). Een voorbeeld is de moeilijkere toegang voor hulpdiensten in het geval er een ongeluk of incident plaatsvindt, waardoor de ontruiming van een incident langer duurt en de kans op secundaire incidenten toeneemt.



FIGUUR 1.14: VOORBEELD FILETERUGSLAG OP DE A13 VAN DEN HAAG NAAR ROTTERDAM.

1.7 Samenvatting en reality check: grenzen aan schokgolftheorie

De drie belangrijkste *macroscopische verkeerskenmerken* zijn de dichtheid k , de intensiteit q en de gemiddelde snelheid u . Die eerste twee macroscopische grootheden zijn nauw gerelateerd aan microscopische grootheden: $q = 1/H$ (de gemiddelde volgtijd) en $k = 1/S$ (de gemiddelde volgafstand). Het verband tussen q , k , en u is gemakkelijk te onthouden: $q = ku$ (“ $kuu = ku$ ”). Omdat er een ondergrens is aan de volgtijd die bestuurders op elkaar durven aan te houden bij een bepaalde snelheid is er een bovengrens aan de hoeveelheid voertuigen die per tijdseenheid kunnen passeren, dit noemen we de capaciteit. Het blijkt dat de grootste capaciteit (tussen de 2200 en 2400 voertuigen per uur per rijstrook) wordt bereikt bij een gemiddelde snelheid van zo rond de 85-90 km/u. De (evenwichts)relatie tussen intensiteit en dichtheid, die als maximale waarde dus de capaciteit heeft, noemen we wel fundamenteel diagram (FD). Vanwege $q = ku$ zijn er drie vormen van het FD: $q = Q(k)$, $u = U(k)$ en $u = U(q)$, die alle drie exact dezelfde informatie bevatten.

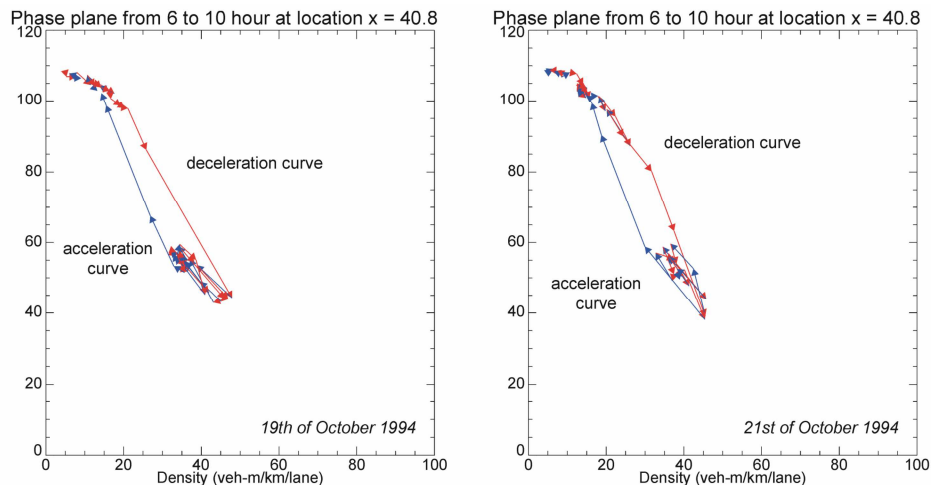
Met behulp van het FD kunnen we ook heel grof de dynamica van verkeer beschrijven en voorspellen. De theorie die we daarbij gebruiken heet schokgolftheorie. Met schokgolftheorie kunnen we voorspellen hoe de grens tussen twee stationaire verkeerstoestanden (evenwichtstoestanden die we in het FD kunnen vinden) zich ontwikkelt over de ruimte en tijd. De belangrijkste formule daarbij is die van de schokgolfsnelheid, die dus bepaald wordt door het verschil in intensiteit te delen door het verschil in dichtheid:

$$\omega_{12} = \frac{q_1 - q_2}{k_1 - k_2}$$

Het grootste pluspunt van schokgolftheorie is dat het een zeer eenvoudige theorie is waarmee het mogelijk is de verkeerscondities in tijd en ruimte te voorspellen. De voorspellingen komen in grote lijnen overeen met wat we in de werkelijkheid zullen waarnemen. We kunnen schokgolftheorie nog een klein stukje realistischer maken door toepassing van kinematische golf theorie (ook wel: eerste orde theorie). Die theorie is gebaseerd op één dynamische vergelijking: de wet van behoud van voertuigen. Maar zelfs dan zijn er veel fenomenen die we in de werkelijkheid waarnemen maar die niet door deze eerste orde benadering kunnen worden beschreven en voorspeld. Om gevoel te krijgen voor de verschillen tussen eenvoudige verkeersstroommodellen (gebaseerd op schokgolf + kinematische golftheorie) laten we in de laatste paragrafen nog wat voorbeelden zien van fenomenen die niet door middel van deze modellen kunnen worden verklaard / voorspeld.

1.7.1 Verkeer in evenwicht en hysteresis

Schokgolftheorie en eerste orde modellen stellen dat verkeer altijd in evenwicht is. Met andere woorden, er geldt $u = U(k)$. Dit blijkt echter lang niet altijd het geval, bijvoorbeeld omdat verkeer eindige acceleratie- en deceleratiemogelijkheden heeft. Andere belangrijke aspecten zijn de vertraagde reactie van weggebruikers en het feit dat automobilisten in sommige gevallen anticiperen op stroomafwaartse verkeerscondities. Acceleratie en deceleratie blijken daarbij *niet symmetrisch* – mede als gevolg van anticipatie en vertraagde reactie – wat ook wel *hysteresis* wordt genoemd. Figuur 1.15 laat duidelijk het verschil zien tussen acceleratie en deceleratie in het fundamenteel diagram (snelheid als functie van dichtheid). Acceleratie en deceleratie zijn niet elkaars “gespiegelde evenbeeld”, maar verschillen fundamenteel. Bij het accelereren nemen voertuigen beduidend meer plaats in, en zijn de dichtheden dus lager.



FIGUUR 1.15: HYSTERESE EFFECTEN A9

Om dergelijke processen te beschrijven moeten we afwijken van de premisse dat het verkeer altijd in evenwicht is ($u = U(k)$). Het Payne model gaat bijvoorbeeld uit van de volgende eenvoudige gedragsrelatie:

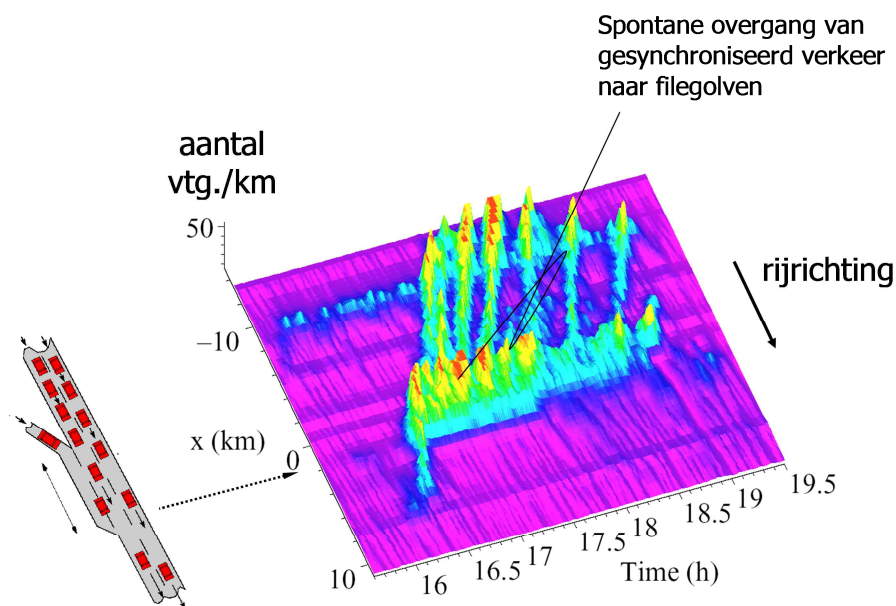
$$u(t + \tau, x) = U(k(t, x + \Delta x))$$

Deze relatie stelt feitelijk dat de snelheid van het verkeer *zich vertraagd aanpast* aan de heersende verkeerscondities stroomafwaarts (anticipatie) van de huidige locatie.

1.7.2 Filegolven, spontane fase overgangen en de capaciteitsval

We hebben gezien dat schokgolftheorie en 1^e orde modellen voorspellen dat congestie ontstaat op het moment dat de verkeersvraag ter hoogte van een bottleneck hoger is dan de (rest)capaciteit. De file ontstaat direct ter hoogte van de toerit. In de praktijk zien we echter vaak dat de file niet ter hoogte van de bottleneck ontstaat, maar zo'n 500 m stroomafwaarts (Figuur 1.16). Dit duidt erop dat automobilisten tijdelijk erg korte volgtijden accepteren (in geval van invoegend verkeer) en pas na enige tijd deze volgtijd aanpassen tot hun gewenste volgtijd. Om dat voor elkaar te krijgen stappen ze op de rem en er ontstaat file.

Daarnaast zien we in Figuur 1.16 ook dat er 'spontane toestandsovergangen' plaatsvinden. In het gebied waar sprake is van reguliere congestie (ook wel 'gesynchroniseerd' verkeer) kunnen kleine verstoringen uitgroeien tot een (soms volledige) stremming, een filegolf, waarbinnen het verkeer nagenoeg tot stilstand is gekomen. Deze filegolven planten zich met een nagenoeg constante snelheid van -18 km/u voort.



FIGUUR 1.16: ONTSTAAN VAN FILE STROOMAFWAARTS VAN DE BOTTLENECK EN HET ONTSTAAN VAN FILEGOLVEN IN EEN GEBIED VAN GESYNCHRONISEERD VERKEER

Kinematische golftheorie en schokgolftheorie zijn niet in staat deze fenomenen te voorspellen. Wel is het mogelijk om, wanneer filegolven eenmaal zijn ontstaan, de wijze waarop de golf zich over het traject verplaatst met deze eerste orde benaderingen te voorspellen. Hiervoor moet dan wel een belangrijke aanpassing aan het fundamenteel diagram worden gedaan: deze moet de zogenaamde *capaciteitsval* bevatten. In stedelijk verkeer is in elk geval intuïtief te verklaren, immers, voertuigen houden waarschijnlijk langere volgtijden aan als ze weggrijden vanuit stilstand (bij een verkeerslicht - zie Figuur 2.5), dan wanneer ze met hoge snelheid achter elkaar rijden. Het blijkt dat ook op snelwegen de capaciteit in vrije afwikkeling structureel hoger is dan die in congestie. Grofweg geldt: hoe groter het snelheidsverschil tussen vrije afwikkeling en congestie, hoe groter de capaciteitsval. De oorzaken van de capaciteitsval zijn divers en nog altijd onderwerp van studie. De kern van de zaak is dat bestuurders in congestie fundamenteel anders rijden dan in vrije omstandigheden, onder andere grotere reactietijden hebben en (dienstengevolge) grotere volgtijden aanhouden. Figuur 1.15 illustreert dit door bijvoorbeeld te laten zien dat acceleratie en deceleratie uit en in een file fundamenteel van elkaar verschillen. We zullen dit hier niet verder bespreken en ons beperken tot een korte slotopmerking over de macroscopische consequenties van de capaciteitsval.

De capaciteitsval is een unieke en bepaald onprettige eigenschap van verkeersstromen. Zodra er sprake is van congestie neemt de capaciteit af, meestal met zo'n 10-15%, maar er zijn ook gevallen onderzocht waarbij de capaciteit meer dan 30-40% afnam. *Het gaat daarmee als het ware*

van kwaad tot erger. De capaciteitsval is een belangrijke motivatie om verkeer zo lang mogelijk in vrije omstandigheden te laten afwikkelen, immers, zodra er congestie ontstaat neemt hierdoor de effectiviteit van een weg spontaan met 10-15% af. In het volgende hoofdstuk zullen we een aantal dynamisch verkeersmanagementmaatregelen bespreken die precies dat doen: het verkeer in vrije afwikkeling houden. We zullen ook zien dat de er niet alleen lokaal sprake is van een capaciteitsval maar dat er op netwerkniveau ook zoiets plaatsvindt.

2 VERKEERSMANAGEMENT

Verkeersmanagement is de discipline die zich bezig houdt met het managen van verkeersstromen in verkeersnetwerken. Er zijn heel veel verschillende doelen die je met verkeersmanagement kunt nastreven. De oudste en nog altijd belangrijkste doelstelling is de verkeersveiligheid. Verkeersregelinstallaties (VRI's - in de volksmond ook wel "stoplichten" genoemd) zorgen ervoor dat verkeer veilig kan afwikkelen op plekken waar conflicterende verkeersstromen elkaar kruisen. VRI's zijn een belangrijke component in stedelijke infrastructuur netwerken, we zullen daar in dit hoofdstuk de nodige aandacht aan besteden.

Maar ook het eerste volledig geautomatiseerde verkeersmanagementsysteem op de Nederlandse autosnelwegen (het zogenaamde MTM systeem) had als belangrijkste doelstelling het verbeteren van de veiligheid door voertuigen te waarschuwen voor congestie benedenstrooms. Het systeem is operationeel op de Nederlandse snelwegen sinds mid-jaren 80. Het systeem detecteert files, dwz een verlaging in de gemiddelde snelheid en waarschuwt naderende automobilisten. Het systeem was en is zeer succesvol: het aantal ongelukken nam af met circa 25%, het aantal dodelijke kopstaart botsingen daalde nog veel sterker.



FIGUUR 2.1: HET RIJSTROOKSIGNALERINGSSYSTEEM (MOTORWAY TRAFFIC MANAGEMENT: MTM)

Naast veiligheid is efficiency (het maximaliseren voor de doorstroom, voorkomen van congestie, minimaliseren van reistijd, etc.) een belangrijke doelstelling. VRI installaties worden bijna altijd zodanig ontworpen dat ze, gegeven een bepaalde verkeersvraag, zo efficiënt mogelijk de verkeersstromen kunnen laten afwikkelen. Behalve veiligheid en efficiency zijn er ook andere doelen, die met name in de afgelopen decennia steeds belangrijker zijn geworden: betrouwbaarheid (garantie aansluitingen, gegarandeerde reistijd, etc.) en milieu doelstellingen, zoals het minimaliseren van emissies (CO_2 , NO_2) en het minimaliseren van geluidsoverlast.

2.1 Stedelijk verkeersmanagement: verkeerslichten

Het eerste daadwerkelijke verkeerslicht stamt uit 1868. Het werd ontworpen door de excentrieke Engelse ingenieur John Peake Knight en was in feite een seinpaal in combinatie met een rode gaslamp, zoals die ook voor treinovergangen werd gebruikt. Het verkeerslicht werd geïnstalleerd vlakbij de Westminster bridge in London om er voor te zorgen dat voetgangers konden oversteken zonder het gevaar te lopen onder de wielen of hoeven van paardenkoetsen te belanden. Het probleem was evenwel dat dit verkeerslicht handbediend was en dat één van de dienstdoende politieagenten (die gedurende de hele dag dus niets anders deed dan de seinpaal omhoog en weer terug bewegen) al in de eerste maand zwaar gewond raakte toen de gaslamp om onbekende redenen in zijn gezicht explodeerde. Het project werd ogenblikkelijk voor onbepaalde tijd stopgezet.

Het zou dik 40 jaar duren voor er op grote schaal in London verkeerslichten werden geïnstalleerd en gebruikt.

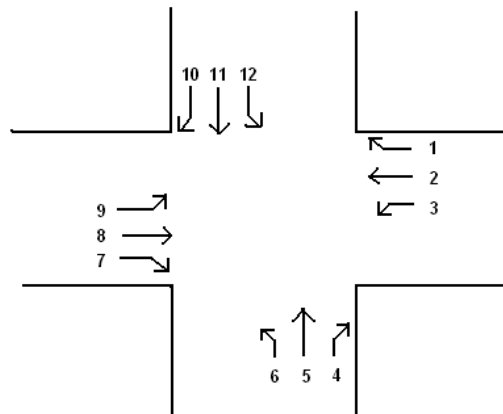
Alhoewel de meningen verdeeld zijn, wordt het eerste volledig geautomatiseerde verkeerslicht met drie kleuren (groen, oranje/geel en rood) zoals we dat nu nog steeds gebruiken, doorgaans toegeschreven aan een politieagent uit Detroit (VS), William Potts. Inmiddels zijn verkeerslichten niet meer weg te denken uit het straatbeeld. In de navolgende paragrafen zullen we kort de verschillende ontwerpaspecten van verkeersregelininstallaties (VRI's) de revue laten passeren.

2.1.1 Basisbegrippen: nummering, conflictmatrices en ontruimingstijden

Gezien de grote hoeveelheid verkeersregelininstallaties in verkeersnetwerken (een flinke stad als Rotterdam heeft al gauw honderden VRI's) is standaardisering van een aantal eigenschappen van kruispunten belangrijk. Op die manier kan ook het ontwerp van verkeersregelingen gestandaardiseerd worden. Dit is goed voor de efficiëntie en veiligheid van verkeersregelingen en vooral ook voor de herkenbaarheid van de regelingen voor weggebruikers.

Standaardnummering rijrichtingen

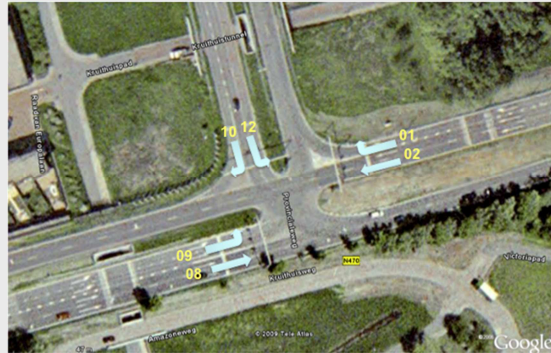
Figuur 2.2 toont de standaardnummering van rijrichtingen op een kruispunt voor autoverkeer, deze loopt van 1 t/m 12 met de klok mee. Ook voor fietsers, voetgangers, en openbaar vervoer zijn regels voor standaardnummering die we hier verder niet zullen behandelen. Door op een kruispunt nu vast te leggen welke (hoofd)rijrichting nummer 2 (en 8) krijgen, ligt de codering voor het gehele kruispunt vast. Het is overigens niet ongebruikelijk dat sommige richtingen gecombineerd worden, bijv richting 1 + 2 (die dan aangeduid worden met 2).



FIGUUR 2.2: STANDAARD NUMMERING VAN RIJRICHTINGEN OP EEN KRUISPUNT

Conflictmatrix

Nadat het kruispunt schematisch is gecodeerd (Figuur 2.2) wordt een zogenaamde conflictmatrix geconstrueerd die aangeeft welke rijrichtingen wel of niet tegelijkertijd kunnen rijden (en dus wel of niet tegelijkertijd groen kunnen krijgen). Op grond van deze conflict matrix kunnen we nu een verkeersregeling gaan ontwerpen.



FIGUUR 2.3: KRUISPUNT (KRUIHUISWEG-PROVINCIALE WEG, DELFT) MET CODERING

Naar Van	01	02	08	09	10	12
01				*		
02				*	*	*
08						*
09	*	*				*
10		*				
12		*	*	*		

TABEL 2.1: CONFLICTMATRIX WAARIN CONFLICTERENDE RICHTINGEN ZIJN AANGEGEVEN MET "*"

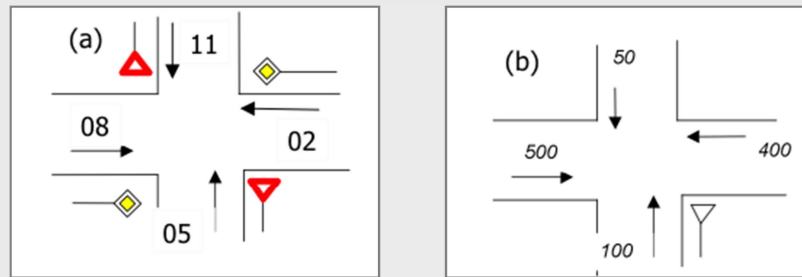
VOORBEELD 2.1: CODERING KRUISPUNT EN RESULTERENDE CONFLICTMATRIX

Wanneer regelen?

De belangrijkste vraag die eerst beantwoord moet worden is wanneer, dat wil zeggen, onder welke omstandigheden het verstandig is om het verkeer met een verkeersregeling te gaan regelen. Zo'n regeling is niet nodig als de intensiteit op elk van de richtingen zeer laag is en het kruispunt overzichtelijk genoeg is zodat bestuurders in alle gevallen kunnen beoordelen of er een voldoende groot gaatje in een conflicterende verkeersstroom is om te passeren, afhankelijk van de vigerende verkeers- en voorrangsregels. De aspecten die een rol spelen bij de beslissing wel/niet regelen zijn onder andere de volgende:

- De verkeersvraag op de hoofdrichtingen en de kansverdeling van deze aankomsten, dit geeft immers een idee over hoe groot de "gaten" zijn in die verkeersstroom
- De verkeersvraag en kansverdeling aankomsten op de overige richtingen (het verkeer dat juist naar "gaten" zoekt)
- De verkeersveiligheid (overzichtelijkheid kruispunt, aanwezigheid langzaam verkeer)
- Mogelijke prioriteit openbaar vervoer, etc.

Voorbeeld 2.2 geeft een rekenvoorbeeld "wel of niet regelen" op grond van de intensiteiten op de hoofd en zijrichtingen.



Veronderstel een kruising met alleen doorgaande richtingen, waarvan 08 en 02 de voorrangsrictingen zijn (plaatje (a)). Het zal duidelijk zijn dat de verdeling van hiaten (een ander woord voor de volgtijden h_i) tussen de voertuigen op de voorrangsrictingen bepalend is voor de hoeveelheid voertuigen die het kruispunt vanuit de andere richtingen (05 en 11) kunnen oversteken. Veronderstel dat de hiaten tussen voertuigpassages op de hoofdrichtingen i (02 en 08) geschaald exponentieel verdeeld zijn (d.w.z. een Erlang- k verdeling)⁶ met een gemiddelde $\mu = 1/q_{tot} = 1/(\sum_i q_i)$ seconden. Dat wil zeggen

$$f(h) = h q_{tot}^2 \exp(-q_{tot} h)$$

Veronderstel vervolgens dat voertuigen op richting 05 en 11 minimaal t_c seconden nodig hebben om te "durven" oversteken. De kans op een hiaat *groter dan* t_c is dan met de volgende formule uit te rekenen

$$\Pr(h \geq t_c) = (1 + q_{tot} t_c) e^{-q_{tot} t_c}$$

We kunnen nu ook uitrekenen wat de maximale hoeveelheid verkeer is dat op de richtingen 05 en 11 nog een hiaat kan vinden om over te steken. Deze wordt natuurlijk bepaald door het aantal hiaten ($> t_c$) in de hoofdrichtingen waar 1 of meer voertuigen uit de zijrichtingen doorheen durven. Veronderstellen we dat de minimale volgtijd tussen voertuigen vanuit de zijrichting 05 (of 11), wachtend op een gaatje in de hoofdstroom, 2 seconde bedraagt ($s_{zij} = 1800$ vtg/uur) dan geldt

$$q_{max} = \int_{t_c} f(h) g(h) dh$$

Met de bovengenoemde aannames voor de volgtijdverdelingen is dan de volgende benadering af te leiden

$$q_{max} = s_{zij} \exp(-q_{tot} t_c)$$

VERGELIJKING 2.1

De interpretatie is dat je de maximale intensiteit vanuit 5 en 11 (de afrijcapaciteit s_{zij}) als het ware schaalt met een factor die afhangt van hoe druk het is op de hoofdrichtingen (q_{tot}). Je kunt nu makkelijk aantonen dat een regeling in dit geval niet nodig is (vul de getallen van het plaatje linksboven in de formule 2.1 met $t_c = 4$ sec), omdat $q_{max} = 662$ voertuigen per uur en op beide richtingen de verkeersvraag kleiner is (50 respectievelijk 100 voertuigen per uur).

VOORBEELD 2.2: BEREKENING WEL/NIET REGELEN⁶

⁶ De exponentiële verdeling voor hiaten (volgtijden) heeft een kansdichtheidsfunctie

$$f(t) = \frac{1}{\mu} \exp\left(-\frac{t}{\mu}\right)$$

met gemiddelde (hiaat) μ en variantie μ^2 . De kans op een hiaat van 0 seconden is natuurlijk gelijk aan nul (tenzij voertuigen aan elkaar vastzitten) maar bedraagt volgens de exponentiële distributie (vul maar in) $1/\mu$ (>0). Dat is niet realistisch en daarom wordt voor volgtijd verdelingen vaak een zogenaamde geschaalde exponentiële distributie (ook wel een Erlang- k distributie) verondersteld:

$$f(t) = \frac{t^{k-1}}{\mu^k (k-1)!} \exp\left(-\frac{t}{\mu}\right)$$

De formule lijkt op die van de exponentiële distributie maar bevat een extra parameter k , die de vorm van de verdeling bepaalt. (zie ook http://en.wikipedia.org/wiki/Erlang_distribution)

2.1.2 Ontwerpen van verkeersregelingen

Een verkeersregeling is de set van instructies die ervoor zorgt dat binnen een bepaalde periode (de zogenaamde cyclustijd) alle richtingen op een kruispunt voldoende groentijd krijgen om veilig en efficiënt te kunnen passeren. Zo'n regeling groepeerd daarbij de richtingen die niet conflicteren in zogenaamde fasen. In Voorbeeld 2.2 kunnen bijvoorbeeld de richtingen 08, 01 en 02 tegelijkertijd groen krijgen. Maar men zou ook kunnen kiezen om de richtingen 08, 09 en 10 tegelijkertijd groen te geven. Na deze fase zouden bijvoorbeeld 01 en 02 gelijktijdig groen kunnen krijgen en daarna richting 12 eventueel vergezeld van richting 08. Maar hoe lang krijgt elke fase groen en welke volgorde van fasen is optimaal? Dit is de kerntaak een verkeersregeltechnicus.

Het ontwerpen van een efficiënte verkeersregeling is complex, zeker als het kruispunt ook langzaam verkeer (voetgangers, fietsen) verwerkt of als er bijvoorbeeld openbaar vervoer (bussen, trams) rijdt dat onder sommige (of alle) omstandigheden prioriteit heeft over "normaal" verkeer. Het wordt nog ingewikkelder als we een kruispuntregeling willen afstemmen op andere kruispuntregelingen, bijvoorbeeld om zo een "groene golf" te bewerkstelligen. Zo'n groene golf zorgt ervoor dat verkeer dat groen krijgt op een kruispunt, ook op de kruispunten benedenstrooms groen krijgt - mits het zich netjes aan de maximum snelheid houdt. Om verkeersregelingen goed op elkaar aan te passen zodat een groene golf mogelijk wordt is bijvoorbeeld schokgolftheorie nodig, en moeten geavanceerde optimalisatie technieken worden toegepast. Het is ook nodig om hierbij gebruik te maken van monitoringssystemen, zodat de regeling in real-time kan worden aangepast en gefinetuned op de daadwerkelijke verkeersvraag - daar komen we verder beneden op terug.

We zullen nu globaal de verschillende stappen bespreken om een verkeersregeling voor het bovenstaande kruispunt te ontwerpen.

Stap 1 - Bepaal/schat/meet ontruimingstijden

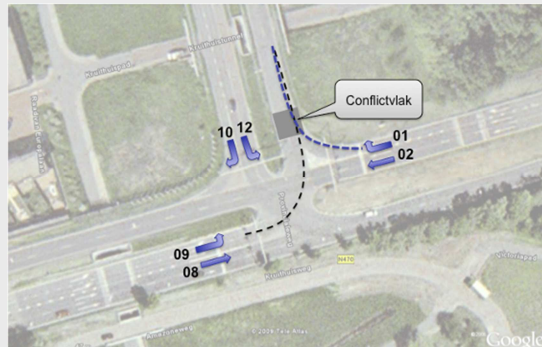
Allereerst moeten we weten welke veiligheidsmarges we moeten aanhouden in de verkeersregeling, dat wil zeggen, hoeveel extra tijd we moeten incalculeren tussen het moment dat een bepaalde richting rood krijgt en het moment dat een conflicterende richting vervolgens groen krijgt. Deze extra tijd noemen we *ontruimingstijd*. Om deze ontruimingstijden te bepalen moeten we eerst bepalen waar (op welke fysieke locaties) de conflicten uit de conflictmatrix precies plaatsvinden. Figuur 2.4 toont het zelfde kruispunt (als Figuur 2.3) met daarin het *conflictvlak* voor de richtingen 01 en 09. Te zien is dat de afstand tussen 09 en het conflictvlak groter is dan die tussen 01 en het conflictvlak. De consequentie daarvan is dat het ontruimen van het conflictvlak langer duurt als richting 09 voor richting 01 groen krijgt, dan in het geval richting 01 voor richting 09 groen krijgt.

Uit analyse of metingen zou bijvoorbeeld kunnen blijken dat de ontruimingstijd van richting 09 naar richting 01 3 seconden bedraagt, dat wil zeggen, we moeten nadat het verkeerslicht voor richting 09 op rood springt nog 3 seconden wachten voor we richting 01 groen geven, om zeker te weten dat alle voertuigen vanuit 09 het conflictvlak zijn gepasseerd. Andersom is het niet nodig om extra ontruimingstijd in te calculeren tussen richting 01 en 09. Zodra richting 01 rood heeft gekregen, kunnen de voertuigen bij richting 09 al groen krijgen. Het formele rekensommetje dat we hier maken is:

Ontruimingstijd van richting i naar j = de tijd die het langzaamste voertuig vanuit i nodig heeft om het conflictvlak te passeren minus de tijd die het snelste voertuig van uit j nodig heeft om het conflictvlak te passeren

Uiteraard is de minimum ontruimingstijd 0 seconden. Als er een negatief getal uit de bovenstaande som komt dan is er geen gevaar voor conflicten. Je kunt je voorstellen dat je deze exercitie voor alle conflictvlakken en conflicterende richtingen kunt uitvoeren. Dat resulteert dan in een conflictmatrix met ontruimingstijden in plaats van kruisjes - zie Tabel 2.2.

Naast ontruimingstijd zijn er ook andere niet door verkeer gebruikte "verliestijden" op een kruispunt. Een voorbeeld is de geeltijd (de tijd dat het verkeerslicht op geel staat), tenminste dat gedeelte van de geeltijd waarin geen voertuigen meer rijden. Een derde component van de verliestijd is het zogenaamde startverlies. Dat is de extra tijd die het eerste voertuig nodig heeft om weg te rijden, bijvoorbeeld om te detecteren dat het licht op groen is gesprongen, het voertuig in de versnelling te zetten, etc.



FIGUUR 2.4: KRUISPUNT (KRUTHUISWEG-PROVINCIALE WEG, DELFT) MET CODERING (2)

Naar Van	01	02	08	09	10	12
01				0		
02				1	3	2
08						0
09	3	2				2
10		0				
12		1	2	2		

TABEL 2.2: CONFLICTMATRIX WAARIN CONFLICTERENDE RICHTINGEN ZIJN AANGEGEVEN MET ONTRUIMINGSTIJDEN (IN SECONDEN)

VOORBEELD 2.3: CONFLICTVLAKKEN EN (HYPOTHETISCHE) CONFLICTMATRIX MET ONTRUIMINGSTIJDEN

De totale interne verliestijd voor een bepaalde richting is gelijk aan de de volgende optelsom

Interne verliestijd = Ontruimingstijd + niet gebruikte geeltijd + startverlies

Of in formele notatie

$$LT_i = LT_i^{clear} + LT_i^{amber} + LT_i^{sg} \quad \text{VERGELIJKING 2.1}$$

Waarin

LT_i^{clear} : Ontruimingstijd (Clearing time) voor richting i

LT_i^{amber} : Effectieve geeltijd (in het Engels wordt voor de tussenkleur "amber" gebruikt)

LT_i^{sg} : Opstart verliestijd (sg staat voor "start green")

Stap 2 - Bepaal de (maatgevende) conflictgroep

Een conflictgroep is een verzameling richtingen die onderling allemaal conflicteren. Het is gemakkelijk om groepen van twee conflicterende richtingen te vinden op basis van de conflictmatrix uit Tabel 2.2. Dit zijn:

- o 01-09
- o **02-09, 02-10, en 02-12**
- o 08-12
- o 09-12

Aangezien richting 02 conflicten heeft met meerdere andere richtingen, kunnen we nu nagaan of er ook groepen van drie onderling conflicterende richtingen zijn. Hiervoor checken we of de met 02 conflicterende richtingen (09, 10 en 12) ook onderling conflicteren. Het blijkt dat 09 en 12 inderdaad conflicteren (richting 09 en 10, en richting 10 en 12 conflicteren onderling niet). Dit leidt dus tot één nieuwe conflictgroep met drie conflicterende richtingen:

- o **02-09-12**

De implicatie is dat onze verkeerregeling minstens drie fasen (combinaties van niet-conflicterende richtingen) moet bevatten. Omdat er één combinatie van drie richtingen is die onderling conflicteren noemen we deze de **maatgevende** (of kritische) **conflictgroep**. Het zou ook kunnen zijn dat er meer 3-fasen conflictgroepen zijn. In dat geval wordt vaak de conflictgroep gekozen waarvoor de verzadigingsgraad het grootst is – die term zal beneden worden uitgelegd. Laten we eerst kijken in welke volgorde we de richtingen uit deze maatgevende conflictgroep groen gaan geven.

Stap 3 - Bepaal de regelstructuur / fase volgorde

Er zijn twee⁷ volgorden te bedenken voor 3-fase conflictgroepen, in dit geval zijn dat 02-09-12 of 02-12-09. Zo'n fase volgorde wordt ook wel regelstructuur genoemd. Tabel 2.4 toont de twee mogelijke regelstructuren op die we op grond van de hierboven toegelichte conflictgroepen (**vet** aangeven in Tabel 2.3) kunnen construeren. Om te bepalen welke van de twee structuren de meest geschikte is kijken we uitsluitend naar de maatgevende conflictgroep (de tweede kolom in Tabel 2.3 en eerste kolom in de regelstructuren in Tabel 2.4).

Tot nu toe hebben we bij de ontwerpstappen geen rekening gehouden met de hoeveelheid verkeer die we per richting verwachten. Dat gaan we nu doen. Daarvoor definiëren we eerst:

- q_j : Aankomstintensiteit voor richting j , d.w.z. het aantal voertuigen per tijdseenheid (vtg/u) dat gedurende de gehele cyclus (dus zowel tijdens groen als rood) arriveert bij richting j
- s_j : De zogenaamde *saturation flow* voor richting j , d.w.z. het aantal voertuigen per tijdseenheid (vtg/u) dat (tijdens een groenfase) kan weggrijpen. Saturation flow (NL: verzadigingsintensiteit) is daarmee gelijk aan de *afrijcapaciteit* dat wil zeggen, de maximale intensiteit waarmee voertuigen uit een wachtrij (een file) kunnen weggrijpen.
- $x_j = \frac{q_j}{s_j}$: De zogenaamde *verzadigingsgraad* van een richting. Als deze grootheid >1 dan komt er meer verkeer aan dan er weggrijdt (wat leidt tot lange wachtrijen en dus vertragingen)

Voorbeeld 2.5 laat op grond van gemiddelde waarden voor q_j en s_j zien hoe je handmatig kunt uitrekenen wat de minimaal benodigde cyclustijd is voor elk van de twee mogelijke regelstructuren. Het blijkt dat structuur I efficiënter is. De structuur levert minder interne verliestijd op (16 sec t.o.v. 21 voor structuur II) en dus een kleinere minimaal benodigde cyclustijd (minder dan een minuut bij de verkeersbelasting uit dit voorbeeld).

⁷ Er zijn bijvoorbeeld 6 verschillende manieren om een 4-fase conflictgroep af te wikkelen.

2 fase conflictgroep	3 fase conflictgroep	4 fase conflictgroep
01-09	-	-
02-09 02-10 02-12	02-09-12	-
08-12	-	-
09-12	-	-

TABEL 2.3: CONFLICTGROEPEN, DE (ENIGE) 3-FASE CONFLICTGROEP IS IN DIT GEVAL MAATGEVEND

	Regelstructuur I				Regelstructuur II			
↓ volgorde	..	..	..	..	..	..	..	..
	02	01	02	08	02	01	02	08
	09	09	10	08	12	01	10	12
	12	01	10	12	09	09	10	08
	..	..	..	..	..	..	..	..

TABEL 2.4: TWEE MOGELIJKE REGELSTRUCTUREN OP BASIS VAN DE CONFLICTGROEPEN UIT BOVENSTAANDE TABEL

VOORBEELD 2.4: CONFLICTGROEPEN (TABEL 2.3) EN MOGELIJKE REGELSTRUCTUREN (TABEL 2.4)

richting	q_i (vtg/u)	s_i (vtg/u)
02	300	1800
09	600	1800
12	400	1800

TABEL 2.5: INTENSITEITEN EN SATURATION FLOWS VOOR DE DRIE RICHTINGEN

De tabel hierboven geeft de te verwachten intensiteiten en saturation flows (afrijcapaciteiten) weer voor de drie richtingen uit de maatgevende conflictgroep. Op grond van de saturation flow (de afrijcapaciteit) weten we dat een voertuigpassage gemiddeld 2 seconden duurt (1800 vtg/u delen door 3600 seconden - denk aan formule 1.1!). Dit impliceert dat we om de verkeersvraag op richting 02, 09 en 12 af te wikkelen per uur gemiddeld $(300+600+400) \times 2 = 2600$ seconden nodig hebben. Dan blijven er in een uur nog $3600-2600 = 1000$ seconden over waarin geen verkeer hoeft worden afgewikkeld. Deze 1000 seconden kunnen gebruikt worden voor de al eerder genoemde *interne verliestijden*.

N.B. Merk op dat voor de richtingen 02, 09 en 12 steeds 1 rijstrook beschikbaar is. Indien er 2 rijstroken beschikbaar zouden zijn, wordt de saturation flow gelijk aan 3600 vtg/u.

Met behulp van vergelijking 2.1 kunnen we de verliestijd voor alle richtingen in de maatgevende conflictgroep uitrekenen via

$$LT = \sum_i LT_i$$

Veronderstellen we dat voor alle richtingen de gemiddelde geeltijd 3 seconden is en de startgroen verliezen 1 seconden, dan levert dit op:

$$LT_{(I)} = 3 \times (1+3) + 1 + 2 + 1 = 16 \text{ sec (Fasevolgorde 02-09-12-...)}$$

$$LT_{(II)} = 3 \times (1+3) + 2 + 2 + 2 = 18 \text{ sec (Fasevolgorde 02-12-09-...)}$$

Regelstructuur I is dus duidelijk efficiënter dan regelstructuur II. Per uur kun je met structuur I in de "overgebleven" 1000 seconden (de tijd dat er geen verkeer hoeft worden afgewikkeld) in totaal $1000/16 = 62,5$ cycli afwerken, en slechts $1000/18 = 55,6$ cycli met structuur II. De minimale cyclustijd voor beide regelstructuren bedraagt dus

$$t_{C,(I)}^{\min} = 3600/62,5 = 57,6 \text{ sec}$$

$$t_{C,(II)}^{\min} = 3600/55,6 = 64,8 \text{ sec}$$

Deze minimale cyclustijden zijn ook direct te berekenen met formule 2.3 ¹

$$t_c^{\min} = \frac{\sum_j LT_j}{1 - \sum_j x_j}$$

VERGELIJKING 2.3

VOORBEELD 2.5: BEREKENING MINIMALE CYCLUSTIJD

2.1.3 Toepassingen van schokgolftheorie bij geregelde kruispunten

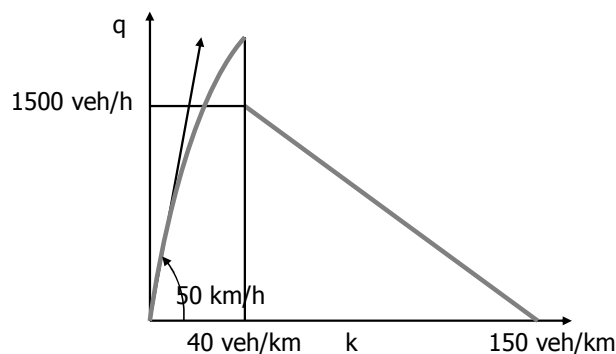
In deze paragraaf laten we zien hoe je schokgolftheorie (besproken in paragraaf 1.5) kunt toepassen voor het analyseren en voorspellen van de verkeerscondities bij kruispunten. Beschouw een geregeld kruispunt met voor de hoofdrichting groentijd t_g en roodtijd t_r . Veronderstel dat voor deze richting het fundamenteel diagram geldt uit Figuur 2.5. In dit fundamenteel diagram zien we een zogenaamde *capaciteitsval*, een sprong in de intensiteit voorbij de kritische dichtheid. Links van die "sprong" zien we de vrije capaciteit, rechts de al eerder genoemde *afrijcapaciteit*. De capaciteitsval wordt onder andere veroorzaakt omdat voertuigen accelererend vanuit stilstand grotere volgtijden zullen aanhouden dan als ze in volle vaart achter elkaar rijden. In de laatste paragraaf van het vorige hoofdstuk zagen we dat zo'n capaciteitsval ook op snelwegen plaatsvindt.

Welke toestanden verwachten we?

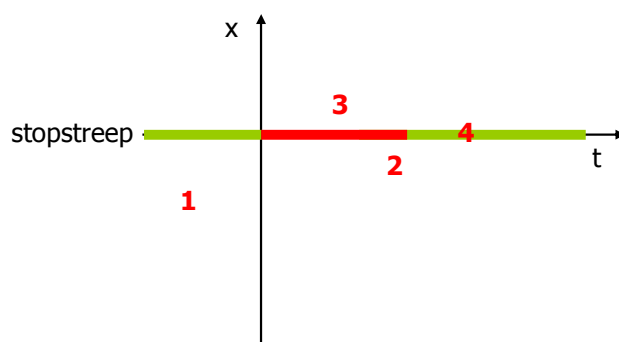
De eerste stap bij het toepassen van schokgolftheorie is weer het bepalen van de verschillende toestanden die we verwachten. In dit geval verwachten we er vier:

1. We veronderstellen dat de ongestoorde verkeerscondities (toestand 1) worden gegeven door $q_1 = 1000$ vtg/h en $k_1 = 18$ vtg/km.
2. Toestand 2: toestand die ontstaat stroomopwaarts van de stoplijn op het moment dat het verkeerslicht op rood staat.
3. Toestand 3: toestand die ontstaat stroomafwaarts van de stoplijn op het moment dat het verkeerslicht op rood staat.
4. Toestand 4: toestand die ontstaat wanneer het verkeer uit de wachtrij die is ontstaan stroomopwaarts voor de toerit wegrijdt.

In Figuur 2.6 is weergegeven waar we deze gebieden verwachten.



FIGUUR 2.5: FUNDAMENTEEL DIAGRAM VOOR STEDELIJKE WEG.



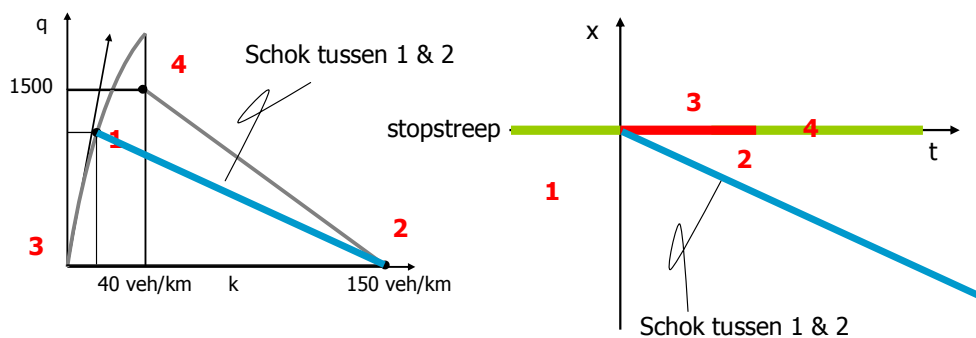
FIGUUR 2.6: WAAR VERWACHTEN WE WELKE VERKEERSTOESTANDEN?

Schokgolven tussen verkeerstoestanden

Nu we weten waar de verschillende toestanden optreden, kunnen we schokgolftheorie toepassen. Dit kan vrij eenvoudig door de schokgolfvergelijking 1.12 toe te passen om de grenzen tussen de geïdentificeerde gebieden te beschrijven. Op deze manier vinden we voor de snelheid van de schokgolf tussen gebied 1 (onverstoorde verkeerscondities) en gebied 2 (wachtrij voor het verkeerslicht):

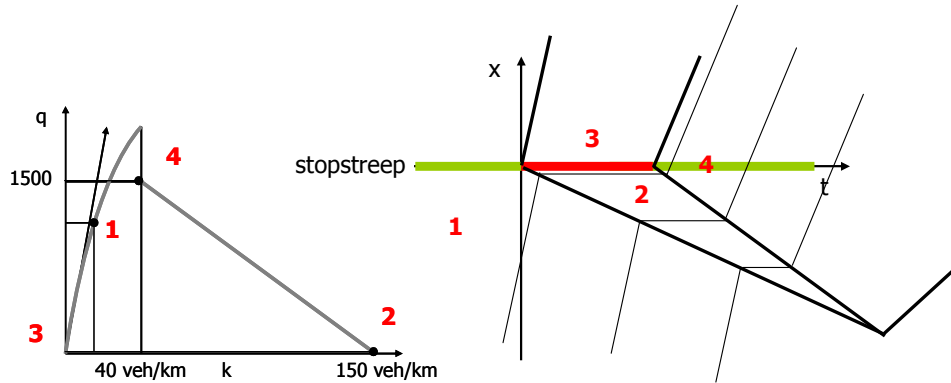
$$\omega_{12} = \frac{0 - 1000}{150 - 18} = -7.6 \text{ km/u}$$

Figuur 2.7 laat de schokgolf grafisch zien.



FIGUUR 2.7: SCHOKGOLF S_{12} MET EEN SNELHEID VAN -7.6 km/h SCHEIDT GEBIEDEN 1 EN 2.

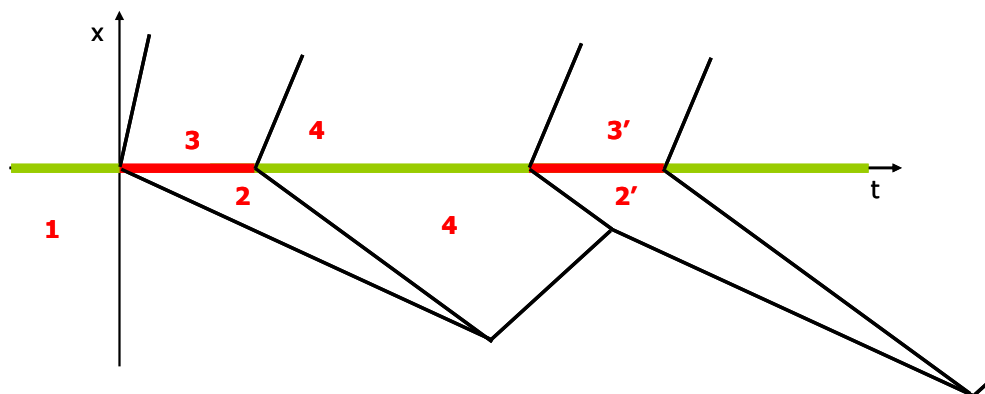
We kunnen dit proces herhalen voor alle grenzen tussen de geïdentificeerde gebieden om zo een volledige beschrijving te geven van het zich voltrekkende proces. In Figuur 2.8 is dit weergegeven. Daarnaast zijn enkele trajectorieën geschetst van voertuigen die zich door de verschillende verkeerstoestanden heen bewegen.



FIGUUR 2.8: DOOR SCHOKGOLFTHEORIE VOORSPELDE CONDITIES BIJ GEREGLD KRUISPUNT.

Voorbeeld toepassing: maximale roodtijd zonder "queue overflow"

Met behulp van dit resultaat kunnen we nu interessante vragen beantwoorden. Zo kunnen we bepalen op welk moment een nieuwe roodfase kan beginnen, zodat gedurende elke cyclus de wachtrij volledig kan oplossen. Is dit niet het geval (Figuur 2.9 illustreert dit) dan ontstaat er een zogenaamde "overflow queue", d.w.z. de wachtrij lost niet meer op en zal per cyclus verder in lengte toenemen. Dit is een van de oorzaken voor netwerk "gridlock", we komen daar verderop uitgebreid op terug. Dit is feitelijk de basis van een verkeersafhankelijke regelaar, waar de groenfase wordt verlengd zolang de gemeten intensiteit vlak voor de stopstreep voldoende hoog is (om en nabij de capaciteit van de weg). We komen ook hier in de volgende paragraaf op terug.



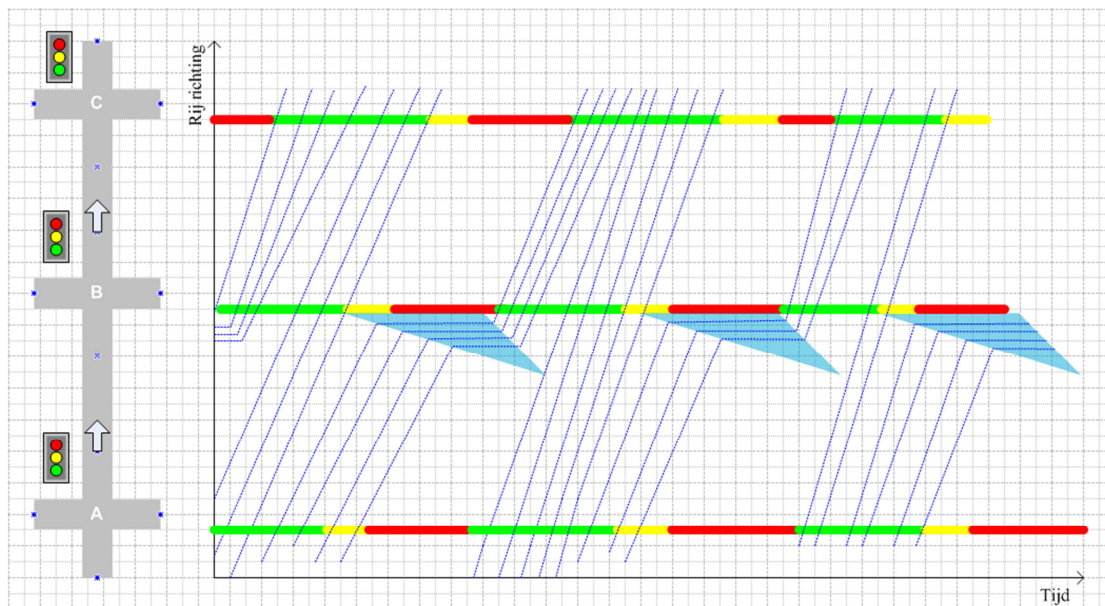
FIGUUR 2.9: RESULTERENDE VERKEERSCONDITIES INDIEN TWEEDE ROODFASE TE VROEG START.

Je kunt met behulp van schokgolftheorie dus grofweg uitzoeken hoe je achtereenvolgende verkeerslichten op elkaar zou moeten afstemmen om ervoor te zorgen dat (a) de wachtrijen bij verkeerslichten binnen een cyclus kunnen oplossen en (b) te voorkomen dat deze per cyclus groeien.

Voorbeeld toepassing: de groene golf

Je kunt nog een stap verder gaan en proberen om de verkeersregelingen zodanig op elkaar af te stemmen dat automobilisten helemaal niet meer in een wachtrij terecht komen. Dit principe wordt wel groene golf genoemd. Figuur 2.10 illustreert dit principe. Verkeersregeling A is niet afgestemd op regeling B, wat resulteert in het feit dat in ongeveer de helft van de gevallen voertuigen vertrekkend vanaf A vertraging ondervinden bij B. De verkeersregelingen bij B en C zijn op wel

elkaar afgestemd en voertuigen vertrekkend bij B ondervinden geen vertraging bij C. Dit laatste noemen we een groene golf.



FIGUUR 2.10: SCHEMATISCHE WEERGAVE GROENE GOLF MET VOERTUIGTRAJECTORIEËN EN SCHOKGOLFTHEORIE.

Een groene golf voor één rijrichting is nog wel met behulp van pen en papier te ontwerpen. Voor het optimaliseren van de verkeersregelingen voor *alle* rijrichtingen voor elk kruispunt langs een bepaalde route of corridor, en het bewerkstelligen van een groene golf voor de hoofdrichtingen, is geavanceerder rekengereedschap nodig. We komen hier verderop in dit hoofdstuk nog op terug.

2.1.4 Voertuigafhankelijke (verkeers)regelingen: gebruik maken van detectiemiddelen

Het aantal voertuigen op een bepaalde richting zal per regelcyclus sterk fluctueren. In het geval van de groene golf heeft dat tot directe consequentie dat de noodzakelijke "groentijd vensters"⁸ groot genoeg moeten worden gekozen zodat het maximale (of gemiddelde, of ...) aantal voertuigen dat we verwachten op een richting, ook daadwerkelijk kan worden bediend. Dat is niet erg effectief. Meer in het algemeen geldt dat het niet erg effectief is om voor elke richting een vaste hoeveelheid groentijd te reserveren. Het is veel logischer om de groentijd te laten afhangen van de daadwerkelijke hoeveelheid verkeer die gedurende een periode bij een kruispunt arriveert.

Om deze reden zijn er in de afgelopen decennia zogenaamde voertuig-afhankelijke verkeersregelingen ontwikkeld waarbij gebruikt wordt gemaakt van voertuigdetectie, d.w.z. van allerlei sensoren in de weg die meten hoeveel verkeer er per richting arriveert. Met behulp van voertuigdetectie kunnen verkeersregelingen op grofweg drie manieren worden uitgebreid (verbeterd):

1. De groentijd voor een bepaalde richting kan worden verlengd zodat de wachtrij volledig kan oplossen. Dit wordt gedaan door de hiaattijden te meten van voertuigen die vanaf de stopstreep weggrijden. Zolang er een wachtrij staat bedragen deze ongeveer 2 seconden (de afrijcapaciteit!). Worden de hiaten veel groter dan 2 seconden dan is de rij waarschijnlijk opgelost en kan de groenfase worden afgekapt. Overigens is er natuurlijk wel een maximum *verlenggroen* per richting, anders zou een cyclus te lang gaan duren. Het bepalen van de hoeveelheid verlenggroen per richting kan nog verder worden verfijnd door niet alleen naar hiaten te kijken maar ook te meten (schatten) of er daadwerkelijk nog een wachtrij aanwezig is. Dat kan bijvoorbeeld met behulp van zgn "lange inductielussen". Tenslotte kunnen we zelfs de verkeersvraag verder stroomopwaarts mee wegen in de mate waarin we de groenfase voor een bepaalde richting willen verlengen of niet.

⁸ Een groentijd venster is de groentijd die bij locatie C nodig is op alle vertrekkende voertuigen bij B te kunnen faciliteren (zie Figuur 2.10)

2. Een tweede toepassing van detectiemiddelen bij verkeersregelininstallaties is om bepaalde richtingen pas groen te geven als er daadwerkelijk verkeersvraag is. In een dergelijk geval krijgen de hoofdrichtingen altijd groen (zgn. *wachtgroen*) en de niet-hoofdrichtingen staan dan standaard op rood. Er wordt een extra fase ingelast zodra er verkeersvraag is (gemeten!) op de niet-hoofdrichtingen. We kennen dit principe ook voor voetgangers en fietsers (waarbij er wordt gedetecteerd of er verkeersvraag is d.m.v. het knopje waar je op moet drukken).
3. Ten slotte biedt sensortechnologie de mogelijkheid om (mogelijk conditioneel en situatie afhankelijk) de verkeersregeling dynamisch te prioriteren voor verschillende groepen weggebruikers zoals bussen en trams, fietsers, voetgangers of speciaal vervoer zoals ambulances en hulpdiensten. In Delft is er bijvoorbeeld sprake van conditionele prioriteit voor het openbaar vervoer, d.w.z., ligt een bus achter op schema, dan krijgt deze direct prioriteit en worden alle andere richtingen afgekapt. Ligt de bus keurig op schema, dan wordt een normale verkeersregelcyclus gedraaid. De sensortechnologie bestaat hier uit transponders in de bus zelf, die communiceren met de VRI. Het is niet moeilijk te bedenken dat ook andere detectiemiddelen zouden kunnen worden ingezet (in de weg, maar ook in de bus, bijv GPS tracking).

Het is tenslotte belangrijk om op te merken dat een voertuig geactueerde regeling bij zware congestie in feite reduceert tot een "fixed" regeling, omdat er een maximum zit aan de hoeveelheid extra groen die een richting per cyclus kan krijgen. In zwaar belaste verkeersnetwerken is de lokale winst (in termen van vertragingstijden) van voertuig-geactueerde VRI's vaak niet zo groot. Maar in zo'n geval kan de voertuig detectie voor heel andere doeleinden juist wel zeer belangrijk en in feite onmisbaar zijn.

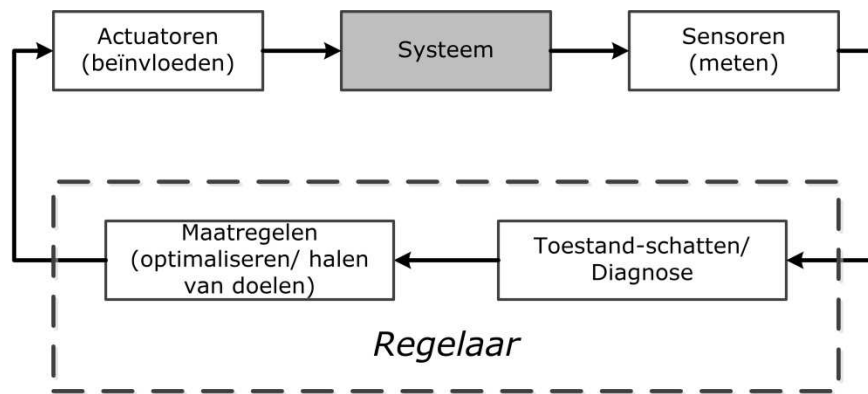
2.2 Voorbeelden van dynamisch verkeersmanagement op snelwegen

Er zijn een aantal fundamentele verschillen tussen stedelijk verkeer en verkeer op snelwegen. In de eerste plaats zijn er in snelwegennetwerken geen conflicterende verkeerstromen en diensgevolge ook geen verkeersregelingen⁹. Daarnaast is de ontwerpsnelheid op snelwegennetwerken beduidend hoger dan in stedelijke netwerken. Ten derde is er de scheiding tussen langzaam en snel (gemotoriseerd) verkeer, op snelwegen is de eerste categorie simpelweg niet toegestaan. Een vierde verschil is de netwerkstructuur, en – gerelateerd hieraan – het aantal verschillende routes tussen een bepaalde herkomst en bestemming. Er zijn tenslotte nog andere verschillen, bijvoorbeeld wat betreft de typen detectiemiddelen (bijv inductielussen, camera's) en zogenaamde actuatoren (zoals dynamische route-informatiepanelen) die je op een snelwegennetwerk kunt gebruiken.

2.2.1 Regelsystemen en regelcirkels

Om dynamisch verkeersmanagement goed te begrijpen is basiskennis van meet- en regeltechniek en systeemdynamica onontbeerlijk. De meeste regelsystemen (ook de voertuig-geactueerde verkeersregelininstallaties uit de vorige paragraaf) kunnen worden geschematiseerd in een zogenaamde regelcirkel (Figuur 2.11). Grofweg bevat een regelcirkel 4 componenten, te weten (1) het "te regelen" systeem; (2) de sensoren waarmee dit systeem (gedeeltelijk) wordt bemeten; (3) een regelaar / regelininstallatie, die de sensor informatie interpreteert (de systeemtoestand bepaalt) en – op grond van vooraf gestelde doelen – deze vertaalt naar te nemen maatregelen; en (4) zogenaamde actuatoren, die de maatregelen daadwerkelijk implementeren in het te regelen systeem.

⁹ alhoewel dit afhankelijk is van de manier waarop "conflicterend" is gedefinieerd; een samenvoeging van twee snelwegen (weefvak) of een oprit kan best als een conflictvlak worden gezien, en een toeritdosering is in feite niets anders dan een verkeerslicht met een speciaal soort regeling (één of twee voertuigen per keer).



FIGUUR 2.11: DE REGELCIKKEL

Voorbeeld – centrale verwarming

Neem bijvoorbeeld een centrale verwarmingsinstallatie (CVI). Het te regelen systeem is hier de temperatuur in een huis. Meestal is er maar één sensor, een temperatuurmeter op een strategische plek in de huiskamer, maar in sommige gevallen zijn er meerdere. De regelaar bestaat uit de verwarmingsketel, die op grond van de gemeten temperatuur en de gewenste (instel)temperatuur (het doel) bepaalt of er moet worden bijgestookt of niet. Moderne HR verwarmingsketels doen meer dan alleen aanslaan als de gemeten temperatuur te laag is, ze maken een inschatting van de huidige en toekomstige situatie en “beoordelen” op grond daarvan welke verwarmingsstrategie kan worden gevolgd. Het doel is dan niet alleen de gewenste temperatuur handhaven, maar ook zo min mogelijk gas verbruiken. Soms anticiperen deze ketels zelfs; als de temperatuur over een half uur terug mag naar 16 graden, dan zullen dit soort ketels niet meer vol gas geven om de huidige temperatuur op 20 graden te houden. De ketel transporteert vervolgens door middel van een pomp en allerlei leidingen heet water naar de beschikbare radiatoren – dit zijn de actuators in het CVI voorbeeld. Hierdoor zal het systeem veranderen – de temperatuur stijgt of daalt, zullen de sensoren deze verandering gedeeltelijk meten en zal de regelaar op grond daarvan weer een nieuwe inschatting maken voor de te nemen acties.

Een regelcirkel is dus een dynamisch systeem. In het ideale geval meten we precies de toestand in een systeem (afhankelijk van hoe we die toestand definiëren natuurlijk) en kunnen we op grond daarvan precies voorspellen hoe het systeem zich zonder ingrijpen zou gedragen en wat er gebeurt als we wel ingrijpen – sterker, we kunnen in het ideale geval precies uitrekenen welke maatregelen nodig zijn om het vooraf gestelde doel te bereiken en we hebben ook de perfecte actuators om deze maatregelen te implementeren. Dit soort ideale systemen komen alleen voor in laboratoria. Zelfs de CVI is verre van ideaal. De sensoren meten bijvoorbeeld niet de (gevoels)temperatuur in een kamer maar de temperatuur op één of enkele plekken. Daarnaast zijn de meeste ketels reactief en niet proactief en er zit bijvoorbeeld ook vaak een forse vertraging in het regelen en de daadwerkelijke effectuering (denk maar aan de tijd die het kost om handmatig de douche op de juiste temperatuur te krijgen). Niettemin, zelfs hele simpele reactieve regelsystemen zijn in staat om complexe systemen op een stabiele en betrouwbare manier te regelen. We zullen nu kort de verschillende componenten uit de regelcirkel behandelen in de context van verkeersmanagement.

Sensoren en Actuators

In de vorige paragraaf hebben we al kennigemaakt met het gebruik van sensoren/detectiemiddelen bij verkeerslichtregelingen. Uit het vorige hoofdstuk is duidelijk geworden dat het in het verkeer heel erg uitmaakt wat je precies meet en waar je dat doet. Op een locatie/dwarsdoorsnede kun je per definitie alleen lokale grootheden meten, zoals volgtijden, intensiteit, lokale snelheden of bijvoorbeeld bezettingsgraad¹⁰, zie paragraaf 0. Dit is een goed voorbeeld van sensoren die iets anders meten (bv. lokale bezettingsgraad) dan datgene wat we – bijvoorbeeld in het geval van een VRI – graag zouden willen weten (de lengte van een wachtrij achter een verkeerslicht). Er zijn ook detectie middelen die ruimtelijke grootheden meten,

¹⁰ De bezettingsgraad is het % van de tijd dat een sensor bedekt wordt door voertuigen en kan als benadering voor de dichtheid worden gebruikt

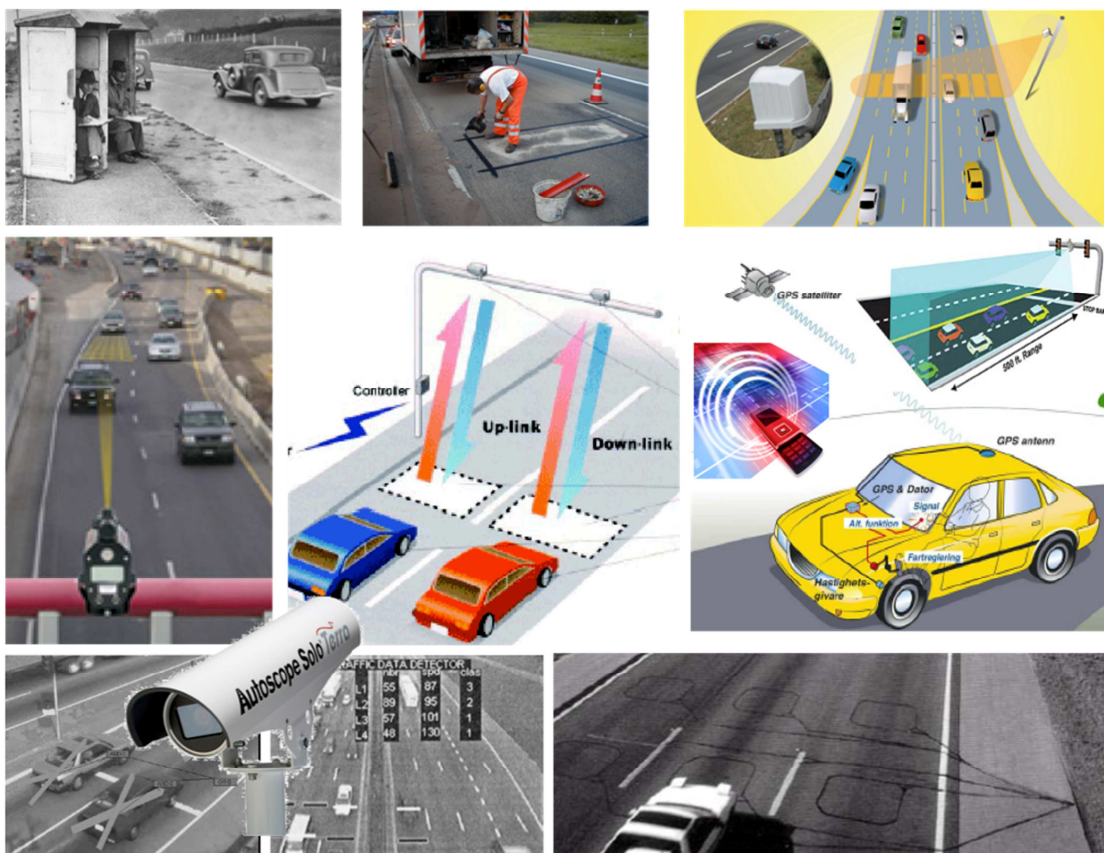
bijvoorbeeld geavanceerde video detectie systemen of de gegevens die uit voertuigen zelf komen (zogenoemde floating car data - FCD) d.m.v. GPS en/of GSM plaatsbepaling.

Figuur 2.12 laat een collage zien van allerlei real-time detectie middelen en technieken die tegenwoordig worden ingezet voor het monitoren van (auto)verkeer, variërend van handmatig tellen (linksboven), inductielusdetectie (midden-boven, rechtsonder), radar detectie (rechtsboven), laser detectie (linksmidden en rechtsmidden), tot (IR) camera technologie (linksonder).

Dit is natuurlijk maar een kleine subset van de sensoren die mogelijk voor verkeersmanagement belangrijk zijn. Het is goed denkbaar dat voor sommige verkeersmanagementtoepassingen ook andere sensoren een rol spelen:

- Detectiemiddelen die gegevens in andere transportmodi (openbaar vervoer, railvervoer, in de lucht- en scheepvaart, etc.) verzamelen.
- Detectiemiddelen (bijvoorbeeld smartphones, Tom-Tom's, enquêtes) die gegevens over tactisch en strategisch reisgedrag meten (herkomsten, bestemmingen, route- en vertrektijdstipkeuze).
- Detectiemiddelen die gegevens over gerelateerde systemen verzamelen (weer, gesteldheid van infrastructuur, informatie over evenementen, etc).

De bekendste actuatoren binnen verkeer en vervoer zijn de verkeerslichten (onderdeel van VRI's) die we al eerder hebben gezien. Andere voorbeelden van actuatoren binnen dynamisch verkeersmanagement zijn verkeerslichten bij toeritdoseerinstallaties (TDI's), dynamische snelheidslimieten of rijstrooktoewijzing d.m.v. de signaleringsportalen boven de snelwegen, dynamische route-informatiepanelen (DRIPs) en de vele mogelijke varianten daarop (grafische panelen, berm DRIPs), dynamische wegmarkeringen en allerlei soorten van borden, panelen en signaalgevers. In de volgende paragrafen zullen we kort een aantal voorbeelden van regelsystemen bespreken die we dagelijks in het verkeer tegenkomen.



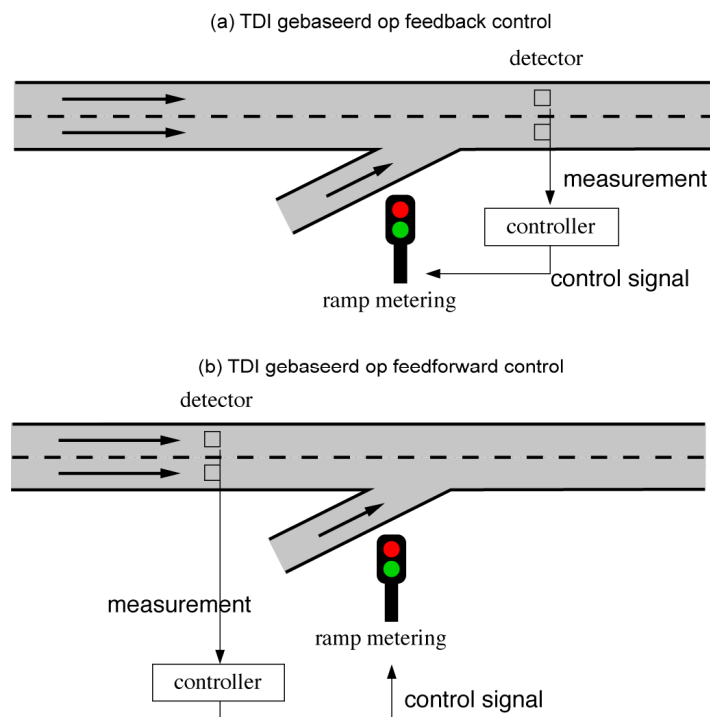
FIGUUR 2.12: COLLAGE VAN VERSCHILLENDE REAL-TIME DETECTIE SYSTEMEN

2.2.2 Voorbeeld 1: toeritdosering

Een toeritdoseerinstallatie (TDI) is niets anders dan een VRI met maar één richting/fase en een zeer korte cyclustijd zodat per groenfase niet meer dan één of twee voertuigen kunnen weggrijden. De twee belangrijkste redenen om TDI's toe te passen zijn

1. Voorkomen dat snelweg verkeer in congestie raakt, immers, dan is er sprake van een capaciteitsval en verliezen we een fors percentage (soms wel tot 30%) van de capaciteit.
2. Voorkomen dat een file te lang wordt en daarmee voorkomen dat deze kan terugslaan naar verder bovenstrooms gelegen op- of afritten. We komen verderop nog terug op het fenomeen fileterugslag en waarom dit zo'n groot probleem is.

Een TDI is een evident voorbeeld van een regelcirkel. Figuur 2.13 geeft schematische voorbeelden van twee veel voorkomende type TDI-regelsystemen.



FIGUUR 2.13: TWEE VOORBEELDEN VAN TDI REGELCIRKELS, FEEDBACK CONTROL EN FEED FORWARD CONTROL (NB: RAMP METERING = TOERIT DOSEREN)

In beide gevallen uit Figuur 2.13 wordt geprobeerd om de volgtijden van voertuigen die door de TDI worden doorgelaten zodanig af te stemmen dat de verkeertoestand op de snelweg zo lang mogelijk in de vrije tak van het fundamenteel diagram blijft (zie bijv. Figuur 1.3, Figuur 1.4, en Figuur 2.5).

Feedback control (Figuur 2.13a): In dit geval wordt benedenstrooms van de TDI gemeten wat de verkeersstoestand is, bijvoorbeeld door de bezettingsgraad te meten (en / of de dichtheid te schatten). Het bekendste voorbeeld van een feedbackcontroller is het zogenaamde ALINEA algoritme¹¹ dat wereldwijd op veel plekken wordt toegepast. Het algoritme wordt *feedback* genoemd omdat het kijkt naar het netto resultaat van de TDI benedenstrooms. De TDI instellingen worden op grond van elke meting zodanig aangepast dat de bezettingsgraad (dichtheid) benedenstrooms vlak onder een kritische bezettingsgraad (dichtheid) blijft.

Feed forward control (Figuur 2.13b): In dit geval wordt niet gekeken naar het resultaat van eerdere toeritdosering, maar wordt uitsluitend gekeken naar de verkeersvraag bovenstrooms en wordt een veronderstelling gedaan over hoeveel verkeer er door de bottleneck kan (de capaciteit!). De TDI zal dan maximaal die capaciteit min de bovenstroomse verkeersvraag doorlaten. In Nederland wordt een dergelijk algoritme op veel plekken toegepast (het zogenaamde RWS-algoritme).

¹¹ Er is zeer veel gepubliceerd over ALINEA en de vele varianten op het algoritme, dat oorspronkelijk is ontwikkeld door Prof Marcos Papageorgiou (University of Crete, Greece).

Er kleven voor- en nadelen aan beide regelprincipes. In beide gevallen moeten er veronderstellingen worden gedaan over een aantal van de parameters (kritische bezettingsgraad in het feedback geval, restcapaciteit in het feed forward geval). De effectiviteit van de TDI hangt nauw samen met de correcte instelling van die parameters. Vaak zijn feedbackcontrollers iets robuuster en stabielere dan feedforwardsystemen, de precieze redenen hier voor zullen we hier niet behandelen. In het algemeen geldt dat TDI's mits goed ontworpen en getuned gunstige effecten hebben op de lokale verkeersafwikkeling op snelwegen. De keerzijde daarvan is dat dit ten koste gaat van de verkeersafwikkeling van het onderliggend wegennet, immers, er ontstaat door een TDI een wachtrij bij de oprit. In de meeste gevallen mag deze wachtrij niet te groot worden, want anders ontstaat fileterugslag op het onderliggend wegennet en is het netto resultaat qua reistijdwinst klein of zelfs negatief. Een tweede aspect waar men bij TDI's rekening moet houden is de relatie met andere TDI's. Want als op locatie bij oprit A een TDI staat en een kilometer verder benedenstrooms bij locatie B niet, dan zou het zomaar kunnen gebeuren dat automobilisten de wachtrij bij A ontvluchten en verder benedenstrooms invoegen. In dat geval hebben we de problemen alleen maar verplaatst. Om TDI's effectief in te zetten is enige vorm van coördinatie dus wenselijk, zowel met andere TDI's als met VRI's in het aansluitende onderliggende wegennet.

2.2.3 Voorbeeld 2: de 80 km/u zones

De 80km/u-maatregel had als primaire doel om de jaargemiddelde concentraties NO₂ nabij autosnelwegen te verminderen. De idee om dit te doen door een verlaging van de maximumsnelheid naar 80 km/u en deze te handhaven d.m.v. trajectcontrole is tussen 2003 en 2005 getest op de A13 tussen vliegveld Zestienhoven en knooppunt Kleinpolderplein. Mits de maatregel aantoonbaar geen negatieve effecten zou opleveren voor de doorstroming zou de maatregel ook worden ingevoerd op andere trajecten. Op 1 november 2005 is, na positieve evaluatie op de pilot op de A13, en positieve ex ante evaluaties¹² voor de andere locaties, besloten de maatregel ook in te voeren op vier andere trajecten in de Randstad (Figuur 2.14).



- A10 bij Amsterdam op de ring West
- A12 bij Den Haag op de Utrechtsebaan
- A12 bij Utrecht op de parallelrijbanen ring Zuid
- A20 bij Rotterdam op de ring Noord



FIGUUR 2.14: OVERZICHT 80 KM/U ZONES

Ondanks dat de 80 km/u maatregel in beginsel een heel ander doel dient dan bijvoorbeeld een TDI of VRI, moet ook de 80 km/u zone worden gezien als een dynamisch regelsysteem. Het te regelen systeem is hier rijgedrag (specifiek de maximumsnelheid). De sensoren zijn in dit geval geavanceerde kentekenplaatherkenningscamera's aan het begin en einde van de trajecten, die het passage moment van elk passerend voertuig opslaan. De regelaar is de gehele keten van systemen op grond van de ruwe kentekens en tijdstempels de gemiddelde snelheid (=lengte traject / reistijd) bepaalt, vergelijkt met het toegestane maximum en al dan niet hier een consequentie aan verbindt. Behalve de signaalgevers die de maximum snelheid boven de weg aangeven (en ook gebruikt worden voor reguliere rijstrooksignaling) zou je ook de bekeuring die enige weken later op de deurmat valt als de actuator in dit gecombineerde regelsysteem kunnen beschouwen.

¹² *Ex ante* betekent letterlijk *vooraf*. Ex ante onderzoek in ons vakgebied geschiedt vaak met verkeerssimulatie modellen, variërend van gedetailleerde microscopische modellen (die het gedrag van elke individuele verkeersdeelnemer simuleren) tot macroscopische benaderingen zoals schokgolftheorie of wachtrijmodellen. Of zo'n model valide voor het onderzoek is hangt af van welke onderzoeksvraag moet worden beantwoord en welk gedrag hiervoor moet worden voorspeld.

A priori overwegingen

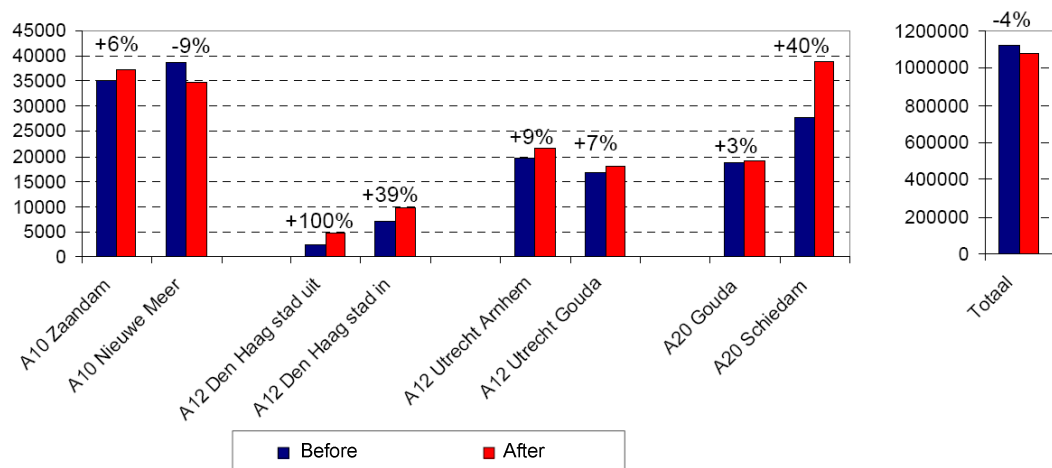
We hebben gezien dat de maximumsnelheid waarmee voertuigen rijden nauw samenhangt met de gemiddeld minimale volgtijd die ze daarbij nog durven aanhouden op hun voorgangers. Een verandering van maximum snelheid kan dus potentieel de capaciteit van een weg veranderen. We weten ook dat de snelheid waarbij de capaciteit gemiddeld het hoogst is, in Nederland zo rond de 85-90 km/u ligt. Tenslotte weten we dat de capaciteit een zeer stochastische grootte is, en dat het dus niet onmogelijk is dat bij een verlaging van de maximum snelheid tot 80 km/u de capaciteit slechts marginaal zal afnemen. Met andere woorden, a priori is er geen reden om aan te nemen dat er grote problemen ontstaan bij de invoering van de 80 km/u zone. De pilot studie op de A13 bevestigde dit beeld, er was een (zij het geringe) afname in uitstoot te zien en nagenoeg geen effect in termen van extra files en vertragingen. De al eerder gememoreerde ex ante studies met simulatie modellen wezen uit dat deze effecten ook naar andere locaties waren te extrapoleren.

Het ministerie van Verkeer en Waterstaat heeft direct na de invoering van de 80 km/u zones op de andere locaties een evaluatieonderzoek uitgevoerd naar de effecten van de 80 km/u maatregel. Het onderzoek concludeert dat de maatregel niet op alle locaties heeft geleid tot (de verwachte) positieve effecten voor de afwikkeling van het verkeer, op trajecten met complexe weefvakken en invoegingen is sprake van een minder effectief gebruik van de wegcapaciteit na invoering van de snelheidsverlaging.



FIGUUR 2.15: COLLAGE PERSBERICHTEN NAAR AANLEIDING VAN DE 80-KM ZONES.

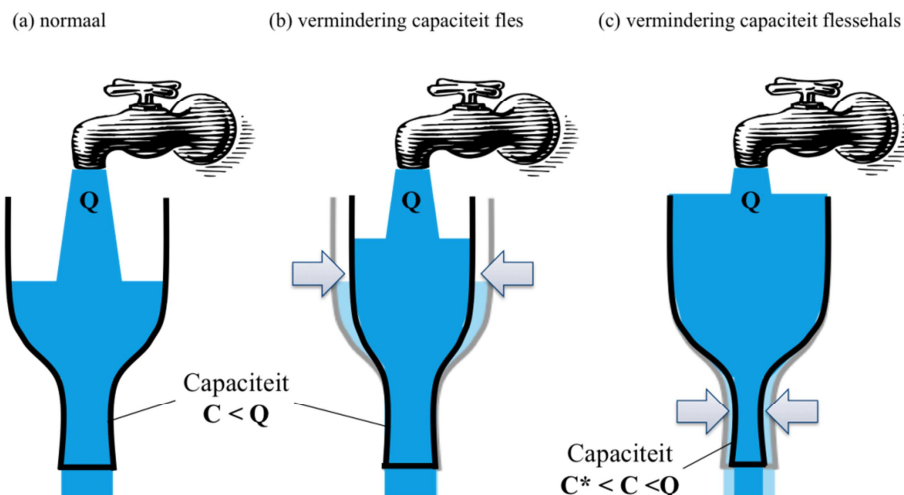
Dit is nogal eufemistisch uitgedrukt. De daadwerkelijke effecten op de verkeersafwikkeling waren in sommige gevallen ronduit dramatisch. Binnen een aantal weken na de invoering van de 80 km/u zones rapporteerden de ANWB (www.anwb.nl) en de verkeersinformatiedienst (de VID - www.vid.nl) een sterke toename van de filedruk (aantal files maal de respectievelijke filelengte), hetgeen in de evaluatiestudie van het ministerie later ook wordt bevestigd (Figuur 2.16). Er blijkt dat de filedruk op de A20 en A12 met 40% tot zelfs 100% is toegenomen. De kernvraag is natuurlijk waarom dit het geval was en – misschien nog wel belangrijker – waarom de ex ante evaluatiestudies er kennelijk zo ver naast zaten. Hiervoor zijn grofweg twee redenen aan te geven, (1) er is niet goed gekeken naar netwerk effecten, en (2) de modellen gebruikt voor de ex ante evaluatie waren niet valide om de effecten van een (gehandhaafde) 80 km/u snelheidslimiet te evalueren.



FIGUUR 2.16: RESULTATEN EVALUATIESTUDIE 80 KM/U ZONES.

Oorzaken: netwerk effecten

A priori zagen we dat de 80 km/u maatregel een lichte daling van de capaciteit tot gevolg kon hebben, omdat bij 80 km/u voertuigen gemiddeld een net iets grotere minimale volgtijd zouden aanhouden dan bij bijvoorbeeld 90 km/u. Maar of een daling in capaciteit ook leidt tot meer files en vertragingen hangt natuurlijk samen met het feit of de capaciteitsverlaging plaatsvindt op een traject met een actieve bottleneck (flessenhals) of niet. Dit is het eenvoudigst te begrijpen door weer het voorbeeld van de omgekeerde flessenhals te beschouwen (Figuur 2.17). Stel dat we het brede gedeelte van de fles versmallen, dan zal er bij gelijke vraag Q de resulterende file evenveel water bevatten, maar (proportioneel aan de versmalling) wel iets langer worden. Versmallen we echter de flessenhals, dan zal de file bij een gelijke vraag Q sneller groeien omdat in dit geval de capaciteit van de fles wordt verlaagd naar $C^* < C$ en zal de file dus niet alleen langer zijn maar ook meer water bevatten.



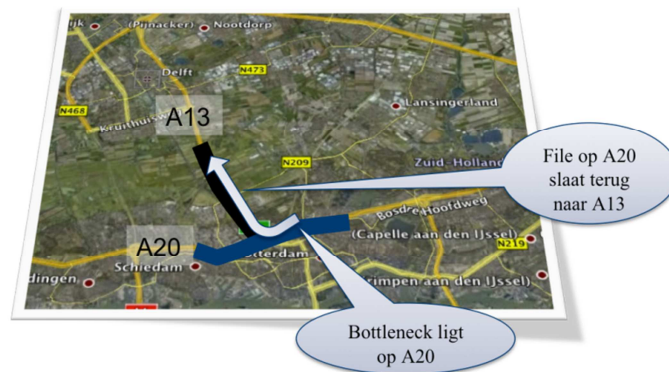
FIGUUR 2.17: "FILE" IN DE OMGEKEERDE FLES (A) EN HET EFFECT VAN HET VERKLEINEN VAN DE CAPACITEIT VAN DE FLES (B) EN HET VERKLEINEN VAN DE CAPACITEIT VAN DE FLESSENHALS (C).

Dit verklaart waarom de 80 km/u zone op de A13 nauwelijks invloed had op de filedruk, want de belangrijkste oorzaak van de dagelijkse files op de A13 ligt niet op de A13 maar stroomafwaarts op de A20 richting Utrecht bij de afrit Rotterdam centrum.

In de meeste gevallen ontstaat daar zo rond 16:00 's middags file, welke vervolgens over het knooppunt Kleinpolderplein terugslaat naar de A13. Figuur 2.18 illustreert dit schematisch. Als men

bij de ex ante evaluatie beter had gekeken naar de precieze locaties van bottlenecks dan had deze inschattingfout niet gemaakt hoeven worden.

Maar dit is maar een gedeelte van het verhaal. Zoals we in het vorige hoofdstuk hebben geconstateerd, is capaciteit niet een eigenschap van een stuk weg maar de directe consequentie van menselijk rijgedrag.



FIGUUR 2.18: REGULIERE FILE OP DE A13 WORDT VEROORZAAKT DOOR FILETERUGSLAG VANAF DE A20.

In het geval dat de modellen waarmee dat rijgedrag wordt gesimuleerd niet in staat zijn het rijgedrag onder de nieuwe veranderde situatie (één met een streng gehandhaafde 80 km/u snelheidslimiet) te voorspellen, dan is het logisch dat de evaluatieresultaten met de nodige reserves moeten worden geïnterpreteerd. Om die tools geschikt te maken om iets te zeggen over een streng gehandhaafde 80 km/u snelheidslimiet zijn empirische gegevens nodig. Dit lijkt op een "kip-en-ei" situatie, omdat we nu juist die rekenmodellen willen gebruiken om *voor invoering* van zo'n maatregel te voorspellen wat de response van het systeem zal zijn. Er zijn echter ook alternatieve manieren om – zonder een maatregel echt uit te voeren – toch gegevens te verzamelen. Je kunt bijvoorbeeld ook door middel van enquêtes en rij- / gedragssimulaties a priori data verzamelen, waarmee de modellen kunnen worden verfijnd zodat ze in elk geval bij benadering de response op bijvoorbeeld de 80 km/u maatregel kunnen nabootsen.

Oorzaken: invalide ex ante onderzoek c.q. fundamentele veranderingen in rijgedrag

In een studie aan de TU Delft (Harms, 2006) is in detail onderzocht welke fundamentele veranderingen in rijgedrag optraden ten gevolge van de 80 km/u zones. Grofweg betreft dat vier aspecten.

1. **Snelheid.** Als gevolg van de strikte handhaving van de maximumsnelheid m.b.v. trajectcontrole is de gemiddelde snelheid op de rijbaan niet 80 km/u maar gedaald tot ruim onder de 80 km/u. Het effect op capaciteit is daarmee groter dan voorspeld werd.
2. **Rijstrookgebruik.** Door afname in het snelheidsverschil is de wens om van rijstrook te wisselen afgenomen. Omdat in Nederland de wet voorschrijft om zoveel mogelijk rechts te rijden heeft dit tot consequentie dat de rijstrookverdeling anders wordt: meer (personen)verkeer maakt gebruik van de rechterrijstrook, terwijl de linker (en middelste) rijstrook onderbenut wordt. Het vrachtverkeer daarentegen is meer gebruik gaan maken van de middelste rijstrook.
3. **Rijstrookwisselgedrag.** Door de kleinere snelheidsverschillen en het toegenomen gebruik van de rechter rijstrook zijn de mogelijkheden om van rijstrook te wisselen sterk afgenomen. Voertuigen voeren de verplichte rijstrookwisselingen dan ook zo veel mogelijk aan het begin van het weefvak uit. Indien het niet mogelijk is om direct deze rijstrookwisseling uit te voeren, verminderen de voertuigen hun snelheid om een geschikt hiaat af te wachten. Ook de doorgaande voertuigen anticiperen door hun snelheid te verlagen om zodoende een geschikt hiaat voor de invoegende voertuigen te creëren. Het gevolg: een zeer inefficiënt gebruik van de beschikbare ruimte.
4. **Volggedrag.** Met name vlak voor het ontstaan van de file, zien we op de 80 km/u zones hogere dichtheden dan op andere snelweg segmenten. Tijdens congestie zijn de volgfstanden op de 80 km/u zones juist hoger (en de dichtheid dus lager), waardoor het langer duurt voordat de file oplost en de lengte van de file, bij een gelijk aantal voertuigen,

is toegenomen. Dit hangt samen met het veranderde rijstrookgebruik: de linkerrijstrook wordt minder goed benut bij het wegvrijen uit een file.

De belangrijkste les die we hieruit kunnen leren is dat we voor het ontwerpen en ex ante evalueren van regelsystemen kennis nodig hebben van de response van dat systeem op de regelaar. Aan de andere kant, de belangrijkste bron voor die kennis ligt in het vakgebied transport en planning bij het bestuderen van data uit daadwerkelijke experimenten zoals de invoering van de 80 km/u zones, en het generaliseren van de daaruit gedestilleerde informatie.

2.2.4 Voorbeeld 3: real-time routeadvies

"We kunnen onze klanten niet beloven dat zij nooit meer in een file terecht zullen komen. Het nieuwe navigatiesysteem stelt bestuurders echter wel in staat gemakkelijk de slimste route naar de eindbestemming te plannen en beter op de hoogte te zijn van de werkelijke vertragingen. Op deze manier kunnen zij de juiste route kiezen op het juiste moment, wetende dat wanneer zij een alternatieve route kiezen deze ook daadwerkelijk beter is."

De bovenstaande claim klinkt logisch: door het verschaffen van verkeersinformatie (of door route adviezen te geven op grond van die verkeersinformatie) zullen reizigers betere routekeuzes maken, d.w.z. voor routes kiezen met een kortere reistijd. Ook hier zijn we duidelijk een regelcirkel terug. Het systeem is in dit geval het routekeuze proces van alle automobilisten. Als sensoren maakt TomTom gebruik van alle detectielussen op het hoofdwegennet en van GSM en GPS locatie tracking op het onderliggend wegennet¹³, en van zeer gedetailleerde actuele wegenkaarten. De actuatoren zijn uiteraard de routenavigatiemodules in het voertuig (of portable). De regelaar (de software) tenslotte berekent de kortste / snelste route dan ongeveer als volgt:

1. Bepaal de *weerstand* (kosten + vertragingstijd + mogelijk andere criteria) voor elk wegsegment in een netwerk.
2. Bepaal nu met een routezoek algoritme¹⁴ de *route met de minste weerstand* tussen begin- en eindlocatie.

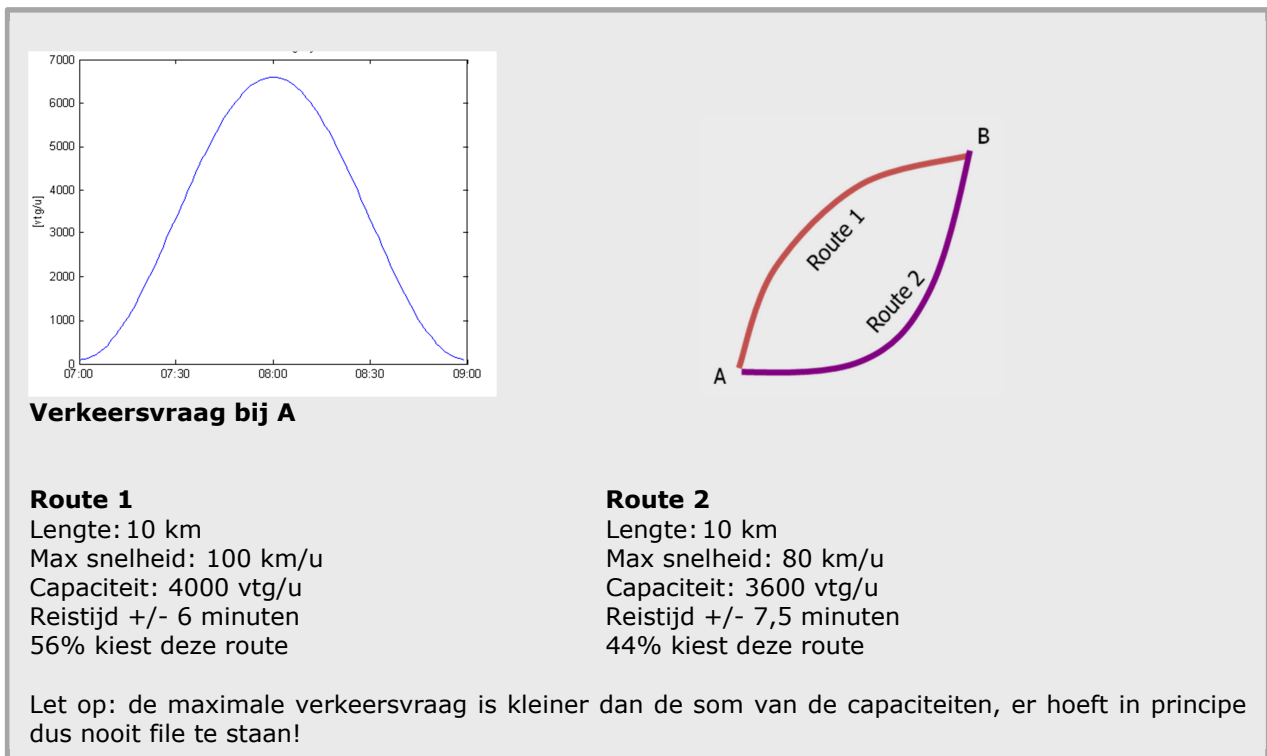
Bij klassieke routenavigatiesystemen (en ook de vele routeplanners op internet) wordt voor stap 1 gebruik gemaakt van de toegestane maximumsnelheid en in sommige gevallen (bijvoorbeeld www.tripcast.nl) van een historisch gemiddelde (of percentiel) reistijd. In het geval van real-time routenavigatiesystemen wordt (op dit moment!) gebruik gemaakt van de meest recente gegevens verzameld met sensoren. Men noemt de reistijd (vertraging) die men op deze manier berekent ook wel de **instantane reistijd** (of instantane vertraging). De instantane reistijd veronderstelt dus dat de huidige verkeerssituatie constant zal blijven gedurende de rit van de gebruiker (en daarna ook).

Illustratief voorbeeld

Stel je nu de situatie voor zoals weergegeven in Voorbeeld 2.6, waarin er twee alternatieve routes tussen de plaatsen A en B bestaan. Op grond van ervaring en persoonlijke voorkeuren blijkt 56% van de reizigers de iets snellere route 1 te kiezen. De verkeersvraag in een typische ochtendspits is ook weergegeven, merk op dat de maximale verkeersvraag kleiner is dan de som van de capaciteiten van beide routes. Er hoeft dus nooit file te ontstaan.

¹³ De claim dat TomTom veel betere verkeersinformatie verzamelt dan Rijkswaterstaat is alleen waar in termen van geografisch bereik (aantal kilometers weg waarover gegevens beschikbaar zijn), en zeker niet wat betreft de kwaliteit (nauwkeurigheid en betrouwbaarheid) van de data

¹⁴ Het bekendste algoritme is dat van Dijkstra, zie bijvoorbeeld <http://nl.wikipedia.org/wiki/Kortstepad-algoritme>.



VOORBEELD 2.6: ROUTENAVIGATIE – TWEE ALTERNATIEVE ROUTES VAN A NAAR B.

Stel nu dat er een (toenemend) percentage van p (%) reizigers hun routekeuze gaat baseren op basis van een real-time navigatie systeem¹⁵. Dat systeem adviseert de reiziger die route te nemen met de kleinste (verwachte) reistijd, die berekend wordt door te kijken naar de huidige filelengte vlak voor locatie B op de route j , d.w.z.

$$T_j(t) = T_j^{\text{free}} + \frac{N_j(t)}{C_j} \quad \text{VERGELIJKING 2.2}$$

Waarin

T_j^{free} : De vrije reistijd op route j

$N_j(t)$: Het aantal voertuigen in de file voor B op route j op tijdstip t

C_j : De capaciteit van route j

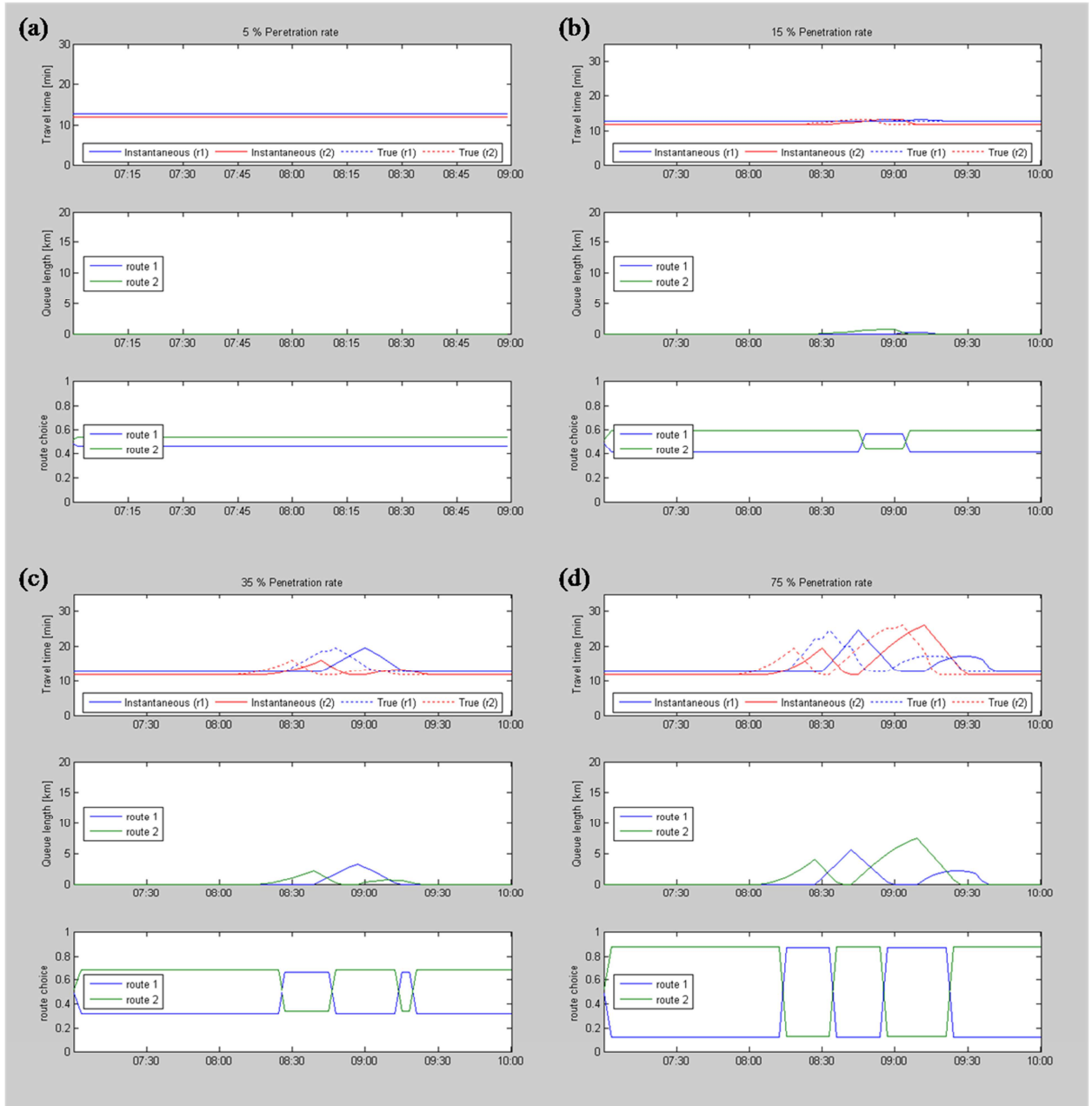
De “verwachte” reistijd $T_j(t)$ is dus – net als die op de DRIPs op ons hoofdwegennet en die in TomTom wordt gebruikt¹⁶ – een *instantane* reistijd, want we veronderstellen dat het aantal voertuigen in de file voor B constant blijft gedurende de trip. Die instantane reistijd is niets anders dan een (naïeve) voorspelling van de werkelijke, zogenaamde **actuele reistijd**, die reizigers gaan ondervinden als ze op het huidige moment t vertrekken. Het is belangrijk om je te realiseren dat je de actuele reistijd niet kunt meten, maar moet voorspellen, immers, voertuigen moeten de trip nog gaan maken!

Figuur 2.19 laat voor vier verschillende percentages gebruikers, (a) 5%, (b) 15%, (c) 35%, en (d) 75%, zien wat de consequenties zijn van routekeuzegedrag gebaseerd op instantane reistijd voor de verkeersafwikkeling op beide routes en de daaruit resulterende (daadwerkelijke) reistijden en files. Voor elk percentage gebruikers toont Figuur 2.19 de instantane en daadwerkelijke reistijden

¹⁵ Dat systeem kan in het voertuig zitten maar ook een dynamisch route informatie paneel (DRIP) bij het keuzepunt A zijn – dat maakt voor dit voorbeeld niet uit.

¹⁶ Dit is in elk geval t/m 2009 / 2010 het geval

op beide routes over de ochtendspits (bovenste grafiekje); de filelengte bij B (middelste plaatje); en de fractie reizigers die voor route 1 dan wel 2 kiest (samen uiteraard altijd gelijk aan 1 of 100% - onderste plaatje).



FIGUUR 2.19: INSTANTANE EN DAADWERKELIJKE REISTIJDEN; FILELENGTE EN ROUTEKEUZE FRACTIE VOOR VIER OPOPENDE PENETRATIEGRADEN (A T/M D) VAN REAL-TIME ROUTENAVIGATIE GEBRUIKERS OP BASIS VAN INSTANTANE REISTIID

De eerste observatie die we kunnen maken (dat is overigens lastig in de grafieken te zien) is dat bij een kleine penetratie graad (Figuur 2.19a) de gemiddelde reistijd over *alle* reizigers verbeterd. Reizigers die voorheen route 2 namen worden nu geadviseerd om de (snellere) route 1 te nemen. Die reizigers zijn dus sowieso beter af. Daardoor neemt de reistijd op route 1 iets toe en die op route 2 iets af en het netto resultaat is een verlaging van het gemiddelde.

Is het percentage gebruikers echter toegenomen tot 15% (Figuur 2.19b) dan neemt het gemiddelde voordeel voor iedereen gestaag af en zien we dat net na 8:30 de actuele reistijd (stippeltjes lijn) op route 1 groter wordt dan die op route 2. Even later zien we dat ook de instantane reistijd op route 2 kleiner is dan op route 1, hetgeen leidt tot een gewijzigd route advies

voor geëquiperde reizigers (te zien in de onderste grafiek in Figuur 2.19b). Bij nog verdere groei van het percentage gebruikers (Figuur 2.19c en d) zien we dat het systeem instabiel wordt en dat de routekeuze sterk gaat oscilleren van route 1 naar route 2 en weer terug. Het gevolg is file en forse vertragingen op beide routes en uiteindelijk reistijd verlies voor iedereen, zowel gebruikers als niet-gebruikers van de real-time routenavigatie!

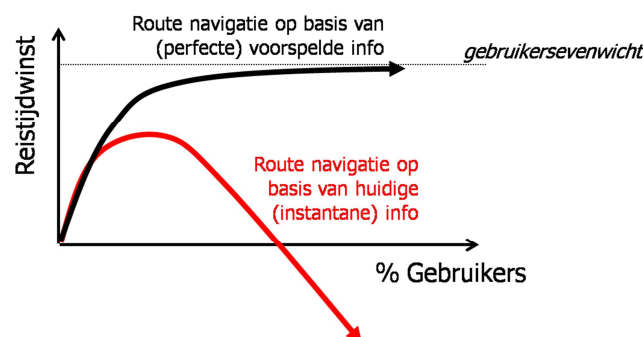
Achterliggende mechanismen en waarom voorspellen zo belangrijk is

Het voorbeeld hierboven laat zien wat er gebeurt als we op grond van de zogenaamde instantane reistijd proberen het routekeuzegedrag te regelen. Dat dit leidt tot een instabiel en oscillerend verkeerssysteem is het gemakkelijkst als volgt te begrijpen. Stel dat er op tijdstip t een kleine wachtrij staat op route 1 en geen wachtrij op route 2. Het is heel goed mogelijk dat de wachtrij allang is opgelost tegen de tijd dat voertuigen bij de bottleneck arriveren en dat het verschil in de daadwerkelijke reistijd heel klein is. Maar omdat de instantane reistijd uit vergelijking 2.2 nu voor route 1 groter is dan voor route 2, wordt iedereen (met een navigatie systeem) over route 2 gestuurd. Doordat iedere uitgeruste gebruiker nu over route 2 rijdt, zal de wachtrij op route 1 inderdaad snel oplossen, echter, door de sterke toename in de vraag op route 2 zal daar nu na verloop van tijd een wachtrij ontstaan. Die wachtrij zorgt ervoor dat even later alle uitgeruste gebruikers weer en masse over route 1 worden gestuurd, et cetera. Omdat de feedback (terugkoppeling van sensor informatie) in dit regelsysteem vertraagd en onjuist is (instantane reistijd ipv de werkelijke reistijd, slaat de regelaar als het ware op hol. Ondanks het feit dat er voldoende capaciteit op beide routes aanwezig is, ontstaan er hierdoor toch lange files en forse vertragingen.

In termen van regeltechniek is er hier sprake van een instabiele regelaar die duidelijk niet voldoet – zeker niet als we de quote over TomTom HD Traffic aan het begin van deze paragraaf er weer bij pakken. Een regelsysteem dat op grond van reistijd routekeuzegedrag aanpast zou dat moeten doen op basis van de reistijd die voertuigen ook daadwerkelijk gaan ondervinden en dat betekent dus dat deze per definitie moet worden voorspeld. Met andere woorden, i.p.v. de instantane reistijd uit vergelijking 2.2 moeten we de reistijd berekenen met

$$T_j(t) = T_j^{free} + \frac{N_j(t + T_j(t))}{C_j} \quad \text{VERGELIJKING 2.3}$$

Vergelijking 2.3 laat zien dat reistijd voorspellen lastig is – we moeten immers de verkeerstoestand (het aantal voertuigen N_j) kennen op het moment waarop een reiziger zijn trip beëindigd (na $T_j(t)$ tijdseenheden). Om dat te doen moeten we dus ook voorspellen hoeveel verkeer er tot die tijd elk van de twee routes zal kiezen. Dat leidt dan tot een kip en ei probleem, want die routekeuze wordt nu juist gedeeltelijk weer bepaald door de reistijd! We zullen niet ingaan op de mogelijk aanpakken om zo'n voorspelling te maken maar kort laten zien wat er gebeurt als we *zouden* voorspellen.



FIGUUR 2.20: SCHEMATISCHE WEERGAVE VAN DE RELATIE TUSSEN DE REISTIJDWINST DIE ROUTENAVIGATIE OPLEVERT (VOOR GEBRUIKERS) EN HET PERCENTAGE GEBRUIKERS.

Figuur 2.20 geeft een schematische weergave van de relatie tussen de reistijdwinst die routenavigatie oplevert (voor gebruikers) en het percentage gebruikers van routenavigatie. Als de routeadviezen gebaseerd zijn op instantane reistijd hebben we al gezien dat er winst zit bij kleine

percentages, maar dat de zaak daarna van kwaad tot erger gaat. Bij (perfecte!) voorspelde informatie blijft de reistijdwinst wel stijgen zij het dat de er sprake is van verminderde meeropbrengst. In het hypothetische geval dat iedereen perfecte voorspelde route adviezen krijgt is niemand meer in staat om zijn reistijd te verbeteren door een andere route kiezen. Deze toestand wordt wel het gebruikersevenwicht (user equilibrium - UE) genoemd¹⁷. De belangrijkste conclusies uit dit voorbeeld zijn dus:

- Routenavigatie op grond van instantane reistijden leidt tot zowel individuele als collectieve reistijdwinst mits het percentage gebruikers klein is (ordegrootte van een paar procent)
- Zodra de penetratiegraad groter wordt neemt de winst voor zowel gebruikers als niet gebruikers af - vanaf een bepaald percentage wordt het systeem zelfs instabiel en is de reistijd voor iedereen slechter dan zonder routenavigatie.
- De enige remedie is om de reistijd (en/of de onderliggende verkeersafwikkeling) te voorspellen.

Voor aanbieders van routenavigatiesystemen en services (zoals TomTom) lijkt het dus verstandig om niet te snel naar een groot marktaandeel voor real-time routenavigatie te willen groeien, want dan bijt het systeem zichzelf lelijk in de staart. Voor verkeersmanagers (Rijkswaterstaat, gemeenten en provincies) is er een nog groter belang: bij toenemende penetratiegraden gaan routenavigatie aanbieders – of ze dit willen of niet – in de stoel van de verkeersmanager zitten. Als de routeadviezen niet stroken met andere vormen van verkeersmanagement, dan zullen geen van beide succesvol zijn, in het ernstigste geval met gevolgen zoals beschreven in deze paragraaf.

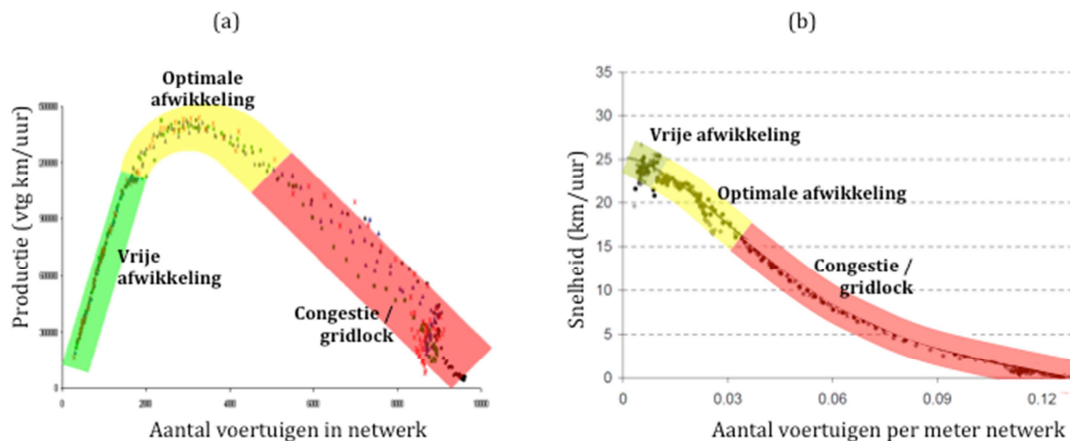
2.3 Integraal verkeersmanagement in netwerken

Uit het voorbeeld van de 80 km/uur-zones blijkt dat de verkeersafwikkeling niet alleen het resultaat is van (vaak slecht voorspelbaar) collectief menselijk gedrag, maar dat daarnaast ook netwerkeffecten (file terugslag, locatie bottlenecks etc.) een cruciale rol spelen. Het routenavigatie voorbeeld geeft bovendien aan dat verkeersmanagement strenge eisen stelt aan de kwaliteit van de middelen (sensoren, actuatoren en regelaars) en de modellen en algoritmes waarmee we managen. Een verkeerd algoritme of twee slecht op elkaar afgestemde regelsystemen kan er voor zorgen dat een verkeerssysteem volledig in de soep draait. Voor we – kwalitatief – zullen ingaan op integraal netwerkmanagement, zullen we eerst kennismaken met het zogenaamde netwerk fundamenteel diagram, ook wel NFD (Figuur 2.21).

2.3.1 Het netwerk fundamentele diagram

Het is nog niet zo lang bekend dat de 'gemiddelde verkeersafwikkeling' in een heel netwerk met een betrekkelijk eenvoudige relatie kan worden beschreven: het netwerk fundamenteel diagram (NFD). Dit NFD beschrijft de relatie tussen de aantallen trips die binnen een uur kunnen worden voltooid maal de lengte van die trips (de zogenaamde verkeersproductie) en de belasting van het netwerk – uitgedrukt in het aantal voertuigen dat zich op een gegeven moment in het netwerk bevindt (Figuur 2.21a). Op de verticale as staat de productie (aantal voltooide trips per uur maal de gemiddelde lengte van die trips), op de horizontale as het aantal voertuigen aanwezig in het netwerk.

¹⁷ Dit klinkt als een erg zware en onrealistische veronderstelling (en dat is het ook), maar deze evenwichtssituatie (geformuleerd door J.G. Wardrop, http://en.wikipedia.org/wiki/John_Glen_Wardrop) is de basis voor bijna alle transport-planningsmodellen die in de praktijk worden gebruikt, bijvoorbeeld om het nationale verkeer- en vervoersbeleid voor de komende kabinetsperioden door te rekenen.



FIGUUR 2.21: MACROSCOPISCH (OOK WEL NETWORK) FUNDAMENTEEL DIAGRAM VAN EEN DEEL VAN HET SAN FRANCISCO NETWORK.

Simpel gezegd, hoe groter de productie, hoe meer trips er worden voltooid. Een andere manier om op netwerkschaal de afwikkelingskwaliteit te beschrijven is om de gemiddelde snelheid en het gemiddelde aantal voertuigen per eenheid lengte in een netwerk tegen elkaar uit te zetten (Figuur 2.21b). Ook hier zien we een verrassende sterke correlatie. Beide figuren laten deze NFD relaties zien voor (een deel van) het netwerk in San Francisco en in beide figuren is schematisch aangegeven in welke tak van het NFD een bepaalde toestand in een netwerk zich bevindt, dat wil zeggen, in vrije afwikkeling (niets aan de hand), in een toestand van optimale afwikkeling (netwerk draait op "netwerkkapaciteit") of in toestand van congestie / gridlock (het netwerk slibt dicht en de prestatie neemt snel af).

Wat we hieruit kunnen zien is dat we, als we maar weten hoeveel voertuigen zich in een netwerk bevinden, heel aardig kunnen schatten hoe de afwikkeling in grote verkeersnetwerken er voor staat. Misschien nog wel het meest opmerkelijke van dit alles, is dat dit verband vrijwel onafhankelijk is van de herkomsten en de bestemmingen van de reizigers die zich in het netwerk begeven. Als je dus eenmaal in een netwerk zit, maakt het voor het netwerk als geheel niet meer uit waar je vandaan komt en waar je naartoe gaat. Dit bijzondere resultaat is onderbouwd met diverse simulatiestudies en empirische metingen in netwerken. Behalve dat het opmerkelijk is dat we voor een heel netwerk de condities kunnen beschrijven met een eenvoudige relatie, zien we in de vorm van het netwerk fundamenteel diagram ook een van de belangrijkste kenmerken van verkeersnetwerken terug: wanneer er teveel voertuigen in het netwerk aanwezig zijn, dan neemt de prestatie in het netwerk af. Met ander woorden: gedurende een uur bereiken gemiddeld minder voertuigen hun bestemming, wat het probleem natuurlijk alleen nog maar groter maakt. Ook op de schaal van hele verkeersnetwerken zien we dus dat files zichzelf verergeren, zelfs tot op het punt dat er helemaal geen beweging meer in het netwerk mogelijk is. Het is belangrijk om je te realiseren dat deze afname in prestatie (outflow) bij toenemende verkeersdruk (inflow) een unieke eigenschap van verkeersnetwerken. Als men in andere "civiele" netwerken, zoals in waternetwerken de druk opvoert dan neemt de outflow juist toe¹⁸. Bij gebrek aan betere terminologie noemen we dit fenomeen in dit dictaat het *meer is minder* effect.

2.3.2 Het "meer is minder" effect: waarom netwerkverkeersmanagement belangrijk is.

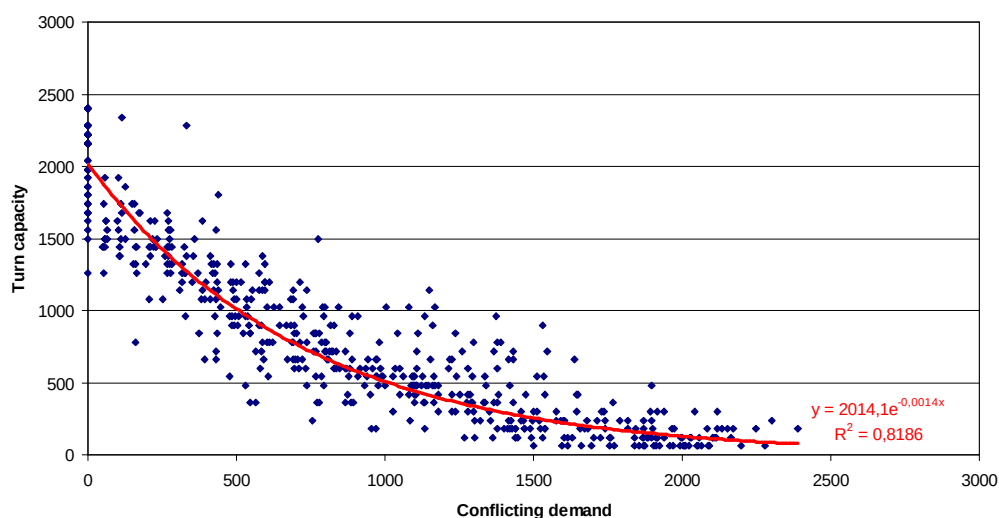
In dit en het vorige hoofdstuk zagen we dat in verkeersnetwerken zowel lokaal (op het niveau van een enkel wegsegment), als globaal (op corridor- en netwerkniveau) dat de verkeersafwikkeling bij toenemende belasting steeds slechter wordt. Er is, met andere woorden, zowel lokaal als globaal sprake van prestatie afname bij toenemende belasting: het bovengenoemde "meer is minder" effect.

¹⁸ Gridlock komt overigens ook in datanetwerken voor (bijv in gecombineerde internet en peer-peer netwerken)

Het lokale "meer is minder" effect: nogmaals de capaciteitsval

We zagen al dat door veranderend rijgedrag er sprake is van een capaciteitsval, dat wil zeggen, in congestie kunnen er minder voertuigen per eenheid tijd een bepaalde dwarsdoorsnede passeren dan onder vrije omstandigheden. Hoe groter het snelheidsverschil tussen congestie en vrije omstandigheden, hoe groter de capaciteitsval is. Een andere voorbeeld van lokale prestatie-erosie (Figuur 2.22) is dat het bij toenemende drukte steeds moeilijker wordt (het per voertuig langer duurt) om in- en uit te voegen, c.q. (bij kruispunten) om een voldoende groot hiaat te vinden om over te steken. Wat de figuur in bredere zin illustreert is dat conflicterende verkeerstromen bij toenemende verkeersdruk elkaar in toenemende mate beïnvloeden.

Je zou dit lokale MiM effect kunnen vergelijken met een caissière in een supermarkt die ten gevolge van een (groeïende) rij klanten voor haar kassa zodanig in de stress schiet dat ze minder hard gaat werken (en dus per klant meer tijd nodig heeft). Op die manier zal bij een gelijkblijvende toestroom van klanten de rij voor de kassa steeds groter worden.



FIGUUR 2.22: IN DEZE GRAFIEK ZIE JE HOE DE CAPACITEIT VAN EEN LINKSAF-RICHTING OP EEN KRUISPUNT OMLAAG GAAT BIJ TOENEMENDE VERKEERSVRAAG OP EEN HOOFDRICHTING (GESCHAT OP BASIS VAN MICROSIMULATIE).

Het globale "meer is minder" effect: fileterugslag en gridlock

Zodra er lokale prestatie afname plaatsvindt, neemt de kans toe dat er ook op corridor- of netwerkniveau problemen ontstaan. De belangrijkste onderliggende mechanismen voor dit globale "meer is minder" effect zijn fileterugslag en (ten gevolge daarvan) conflicterende (elkaar blokkerende) richtingen (de zgn. turncapaciteit op conflicterende richtingen gaat naar nul - Figuur 2.22). We noemen een netwerk waarin alle richtingen elkaar blokkeren ook wel een netwerk in *gridlock* (Figuur 2.23).

Bij fileterugslag wordt een file die ontstaat bij een knelpunt (een flessenhals) zo lang dat bijvoorbeeld afritten, weefvakken of zelfs complete knooppunten verder stroomopwaarts worden geblokkeerd. De consequentie is dat daarmee ook reizigers worden benadeeld die al ruim voor het knelpunt een andere richting op willen gaan en in feite dus geen "schuld" hebben aan de file voor het knelpunt. Zij zullen dus onnodige vertragingen ondervinden. Het effect versterkt zichzelf bovendien (waar hebben we dat al eerder gezien?), omdat door die extra "onschuldige" reizigers de file ook nog eens veel sneller groeit. Stel bijvoorbeeld dat er file ontstaat bij de Coentunnel (Westelijke ringweg A10) in Amsterdam, die helemaal terugslaat tot aan Schiphol (de A4 van Den Haag naar Amsterdam). In dat geval staan in deze file ook reizigers van Den Haag of Leiden naar Utrecht, Amstelveen, de RAI of Diemen, reizigers die dus in de verste verte niet van plan zijn door de Coentunnel te gaan rijden.



FIGUUR 2.23: GRIDLOCK IN EEN STEDELIJK VERKEERSNETWERK

Gridlock (Figuur 2.23) is natuurlijk een vergelijkbaar fenomeen, zij het dat in dit geval verschillende richtingen elkaar blokkeren waardoor het verkeer letterlijk niet meer van haar plaats komt. Gridlock is natuurlijk nauw gerelateerd aan file terugslag, want een wachtrij voor een verkeerslicht kan terugslaan naar het vorige kruispunt stroomopwaarts (en verder) om daar vervolgens voertuigen van andere richtingen te blokkeren.

2.3.3 Geïntegreerd en gecoördineerd netwerkmanagement

Wat het netwerk fundamenteel diagram en de bovenstaande analyse van de onderliggende mechanismen (het lokale en globale "meer is minder" effect) ons leren, is dat voor een vlotte afwikkeling in het netwerk we ervoor moeten zorgen dat het netwerk niet te zwaar wordt belast. In zekere zin geldt hier: voorkomen is beter (in elk geval effectiever) dan genezen, waarbij voorkomen in dit geval equivalent is aan het managen van de verkeersvraag (met een duur woord: mobiliteitsmanagement) b.v. met behulp van woon-werkregelingen bij bedrijven.

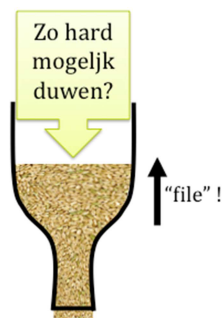
Bijvoorbeeld door mensen aan de randen van het netwerk tegen te houden, of alternatieve routes om het netwerk aan te bieden – dit wordt wel *perimeter control* genoemd. Een tweede optie is om de verkeersvraag beter te spreiden over de dag. We zien hier dus dat in stedelijke netwerken behalve rijgedrag ook (on-trip) routekeuzegedrag een grote rol speelt bij het wel of niet ontstaan van problemen.

Vaak is het een kwestie van prioritering en afstemming tussen verschillende netwerkbeheerders (Rijkswaterstaat versus de provincies of stedelijke wegbeheerders), en sinds een paar jaar ook met aanbieders van routenavigatiesystemen zoals TomTom of Garmin. Hoewel deze commerciële partijen zich in de media sterk maken dat ze met verkeersmanagement niets te maken (willen) hebben en uitsluitend een goed product voor (auto)mobilisten willen leveren (de snelste route, de beste informatie), worden deze aanbieders van autonavigatie systemen en services *de facto* – of dit nu de bedoeling is of niet – straks de grootste en belangrijkste verkeersmanagers van Nederland (en daarbuiten). Het collectieve routekeuzegedrag bepaalt immers de verdeling van verkeer over een netwerk. In het beste geval leidt dit tot een optimale doorstroming, in het slechtste geval, tot gridlock. Een duidelijker voorbeeld van verkeersmanagement door de inzet van dynamische maatregelen (routeadviezen) is nauwelijks denkbaar. Het mes snijdt hier natuurlijk

ook aan twee kanten. Door routekeuzen slim op elkaar af te stemmen kan een voor iedereen gunstigere reistijd worden bereikt en daarmee ook betere verkeersinformatie en routeadviezen.

2.4 Tenslotte – nogmaals de omgekeerde fles

In het vorige hoofdstuk zagen we dat files het resultaat zijn van zowel individueel als collectief rij- en reisgedrag. Hoewel dat gedrag vaak erg variabel en moeilijk voorspelbaar is, kun je toch met verrassend eenvoudige theorieën en modellen heel goed begrijpen en verklaren hoe files ontstaan en weer oplossen. Bijvoorbeeld met behulp van vloeistof dynamica: plaats een fles zonder bodem ondersteboven onder een kraan. Zodra daar meer water uitkomt dan er per tijdseenheid door de flessenhals kan ontstaan er file in de fles. Maar we zagen ook dat die vergelijking op heel veel punten spaak loopt. Automobilisten gedragen zich heel anders dan moleculen, verkeer heeft in veel opzichten een hele andere dynamica dan vloeistoffen. In dit hoofdstuk zagen we hoe je zowel lokaal als in een heel netwerk aan allerlei knoppen kunt draaien om te zorgen dat het verkeer zich veilig, efficiënt en schoon afwikkelt. Zolang de verkeersvraag niet te hoog is lukt dat goed, en er zijn veel succesvolle voorbeelden van verkeersregeltechniek en verkeersmanagement – van gecoördineerde verkeersregelingen, toeritdosering tot het verschaffen van individuele routegeleiding. Echter, zodra de verkeersvraag in de buurt van de capaciteit komt, wordt de verkeerssituatie instabiel. Op zowel route- als netwerkniveau zien we bijvoorbeeld dat files zichzelf verergeren, zo erg zelfs dat er sprake kan zijn van *gridlock*, waarbij een netwerk volledig volstaat met verkeer dat zichzelf volledig vast heeft gezet.



In dat opzicht gedraagt verkeer zich veel meer als een granulair medium zoals zand of rijstkorrels dan een vloeistof zoals water (maar ook deze vergelijking gaat maar heel beperkt op!). Probeer maar eens uit wat de snelste manier is om een kilo rijst door een trechter heen te gieten: (a) gelijkmatig gieten; (b) alles in de trechter gieten en dan hard duwen. Het antwoord is natuurlijk (a). Het is niet moeilijk om op YouTube voorbeelden te vinden van de verkeersequivalent van die tweede “oplossing”. Zoek maar eens met de termen “crazy traffic jam” en “Cairo”, “Shanghai” of “Mumbai” naar filmpjes op het internet. Dat verkeersnetwerken kwetsbaar en onbetrouwbaar worden bij hoge belasting en dat oplossingen gezocht moeten worden in de sfeer van (met beleid) doseren en verdelen, geeft in elk geval duidelijk aan dat er (veiligheids- efficiency en misschien zelfs milieu) winst zit in het slim inzetten van mobiliteits- en verkeersmanagement, en dat deze in feite onmisbaar zijn om de verkeersafwikkeling in een netwerk betrouwbaar en efficiënt te houden.

