

**stichting
mathematisch
centrum**



AFDELING MATHEMATISCHE BESLISKUNDE

BC 16/77

SEPTEMBER

G.J. OLSDER

DE THEORIE DER DIFFERENTIAAL SPELEN EN HAAR TOEPASSINGEN

2e boerhaavestraat 49 amsterdam

BIBLIOTHEEK MATHEMATISCH CENTRUM
AMSTERDAM

Printed at the Mathematical Centre, 49, 2e Boerhaavestraat, Amsterdam.

The Mathematical Centre, founded the 11-th of February 1946, is a non-profit institution aiming at the promotion of pure mathematics and its applications. It is sponsored by the Netherlands Government through the Netherlands Organization for the Advancement of Pure Research (Z.W.O).

De theorie der differentiaalspelen en haar toepassingen
(samenvatting van een lezing, te houden aan het MC, Amsterdam, sept. 1977)

door

G.J. Olsder^{*)}

SAMENVATTING

De inhoud van dit rapport is een samenvatting van een lezing van de auteur, gehouden tijdens de werkweek "Modelvorming met speltheorie", georganiseerd in september 1977 door de afdeling Mathematische Besliskunde van het Mathematisch Centrum. De opzet van de lezing was om de toepassingsmogelijkheden van de theorie der differentiaalspelen te belichten.

Allereerst wordt de nodige theorie in het kort behandeld. Vervolgens worden een aantal toepassingen uitgewerkt. Zowel nulsomspelen als niet-nulsomspelen komen aan de orde. Het accent ligt op singuliere oppervlakken bij de nulsomspelen en op informatiestrukturen bij de niet-nulsomspelen. Verder worden nog enkele verraderlijke eigenschappen van Nash-oplossingen bij stochastische spelen aangegeven.

TREFWOORDEN: *Differentiaalspelen, toepassingen van differentiaalspelen, nulsomspelen, niet-nulsomspelen, singuliere oppervlakken, informatiestrukturen.*

^{*)} Technische Hogeschool Twente, Enschede

1. INLEIDING

De theorie der differentiaalspelen (TDS) kan worden beschouwd als een kind van de ouders optimale besturingstheorie en speltheorie en heeft daarvan dan ook vele karaktertrekken overgenomen. Voor een studie in de TDS is kennis van deze ouders dan ook waardevol. Wat begrippen en eigenschappen betreft is de TDS echter veelzijdiger dan zowel de besturingstheorie als de speltheorie. Geprobeerd zal worden dat in dit verhaal aan te tonen.

Het verband tussen eindig dimensionale optimalisatie ("programming"), optimale besturingstheorie, speltheorie en de TDS is aangegeven in fig. 1.

	een speler	meer spelers
statisch	eindig dimensionale optimalisatie	speltheorie
dynamisch	optimale besturingstheorie	differentiaal- spelen

Fig. 1. De plaats van de theorie der differentiaalspelen

In eindig dimensionale optimalisatieproblemen bestaat de oplossing altijd uit eindig veel getallen, men denke bijv. aan Lineair Programmeren.

In de optimale besturingstheorie zal de oplossing i.h.a. een functie van de tijd zijn. Een voorbeeld van een probleem is het volgende. Een schip

moet door een gebied waar stromingen zijn varen van een gegeven beginpunt naar een gegeven eindpunt. Deze stromingen, welke zowel tijds- als positie afhankelijk kunnen zijn, worden bekend verondersteld. Hoe moet de stuurman sturen om zo snel mogelijk in het eindpunt te komen? De roerhoek wordt hier als besturingsvariable genomen en deze kan met de tijd veranderen.

Een probleem dat behoort tot de speltheorie is "boter kaas en eieren". De oplossing bestaat in dit geval uit eindig veel zetten. Het verschil met eindig dimensionale optimalisatie is dat er nu twee (i.h.a. twee of meer dan twee) spelers zijn met doelstellingen welke verschillend zijn.

Ook in de TDS is meer dan één speler aanwezig. De spelers kiezen nu functies die van de tijd en/of de toestand afhangen. Van de vier genoemde gebieden is de TDS het meest gecompliceerd en is nog niet volledig tot ontwikkeling gekomen, ofschoon reeds vele resultaten verkregen zijn. Veel inzicht wordt nog steeds verkregen door specifieke voorbeelden op te lossen.

Historisch gezien verschenen de eerste bijdragen in de optimale besturingstheorie en de TDS ongeveer gelijktijdig, 1950-1953, onafhankelijk van elkaar. De TDS is echter sterk tot ontwikkeling gekomen nadat de optimale besturingstheorie ± 1960 reeds een duidelijk gezicht had. Deze opeenvolging lijkt in eerste instantie logisch. De ontwikkeling van beide richtingen is echter duidelijk langs verschillende lijnen verlopen. Het dynamisch aspect is gemeenschappelijk, maar de TDS heeft bovendien een speltheoretisch aspect, welke de dynamica overschaduwet zodat de technieken ontwikkeld in de optimale besturingstheorie niet eenvoudig herformuleerd kunnen worden voor de TDS. Dit verschil maakt dat het beeld van de TDS nogal ongunstig uitvalt t.o.v. de optimale besturingstheorie voor de niet-expert. Mathematische strengheid en geometrische intuïtie zijn nog niet samengegaan om tot een afgeronde theorie te komen. Dit kan wetenschappelijke werkers, die best wat kennis omtrent de TDS zouden kunnen gebruiken, ontmoedigen. Een ander nadelig aspect is dat de TDS nog nauwelijks aan universiteiten en/of hogescholen onderwezen wordt.

2. PROBLEEMSTELLING

Bij een *differentiaalspel* zijn N ($N = 2, 3, \dots$) spelers genaamd P_i , $i = 1, \dots, N$ resp. betrokken. Speler i kan een functie u_i kiezen, welke de

besturing voor speler i of de beslissingsvariabele voor speler i genoemd wordt. Later zal gespecificeerd worden van welke variabelen de functies u_i zullen afhangen. De besturingen u_i , $i = 1, 2, \dots, N$ grijpen samen in op een systeem dat ten tijde t door de toestandsvektor x gekarakteriseerd wordt. De samenhang tussen x en u_i wordt gegeven door de *toestandsvergelijking*:

$$(2.1) \quad \dot{x}(t) = f(x(t), u_1, u_2, \dots, u_N),$$

waar t de rol van de tijd vervult; $\dot{}$ betekent de afgeleide naar de tijd. De vektor x heeft n componenten. $x = (x_1, x_2, \dots, x_n)'$; $'$ betekent getransponeerde. $f = (f_1, \dots, f_n)'$. Merk op dat u_i in het algemeen geen component zal zijn, maar een vektor waarvan het aantal componenten r_i genoemd zal worden. Afhankelijk van de probleemstelling kan het soms noodzakelijk zijn dat t expliciet in het rechterlid van (2.1) voorkomt. Het is dan echter altijd mogelijk weer tot de vorm (2.1) terug te komen door een extra toestandsvariabele $x_{n+1} \equiv t$ te definiëren en deze in de toestandsvergelijking op te nemen. Aangenomen wordt dat er een *beginvoorwaarde* voor (2.1) gegeven is:

$$(2.2) \quad x(t_0) = x_0.$$

Wil de uitdrukking (2.1) zin hebben, dan moeten de u_i -functies gespecificeerd worden. Dat kan op verschillende manieren. De u_i 's kunnen functies van alleen de tijd zijn: $u_i(t)$. Men spreekt dan van *open-loop* besturingen. Bij andere probleemstellingen kan het zijn dat de u_i 's afhangen van de toestandsvektor x : $u_i(x)$. Men spreekt dan van *closed-loop* of *feedback* besturingen. Uiteraard moet de oplossing van (2.1) bestaan, d.w.z. men moet bepaalde voorwaarden opleggen aan f , $u_i(t)$ of $u_i(x)$. We nemen aan dat $f \in C^1$, dus continu differentieerbaar is in alle variabelen. Ingeval we met open-loop besturingen te maken hebben, nemen we aan dat $u_i(t)$ Lebesgue-meetbaar en begrensd is. Voor closed-loop oplossingen is het wat moeilijker om algemene, voldoende zwakke, voorwaarden voor $u_i(x)$, nu *strategieën* genoemd, te geven opdat de oplossing van (2.1) bestaat. Als bijv. $u_i(x)$ niet continu is dan voldoet de differentiaalvergelijking

$$(2.3) \quad \dot{x}(t) = f(x(t), u_1(x(t)), \dots, u_N(x(t)))$$

niet aan de Lipschitz-voorwaarde en behoeft de oplossing van (2.3) niet eenduidig te zijn.

VOORBEELD 1. "De dame in het meer".

Een jongedame zwemt in een rond meer met maximum snelheid v_0 (constante). Ze kan de richting in welke ze zwemt instantaan veranderen. Een man met twijfelachtige bedoelingen, die niet kan zwemmen, rent langs de kant met maximum snelheid 1. Ook hij kan z'n richting instantaan veranderen. De jongedame wil niet eeuwig in het meer blijven; eenmaal aan land gekomen kan ze sneller rennen dan de man. Op het moment dat zij aan wal klautert wil zij de hoek tussen de man en haarzelf, gemeten vanuit het middelpunt van het meer, zo groot mogelijk hebben. De man streeft naar een zo klein mogelijke hoek.

Het is duidelijk dat open-loop besturingen hier geen zin hebben. Elke speler zal direkt reageren op hetgeen z'n opponent doet en daarom moet men closed-loop besturingen nemen om tot een zinvolle probleemstelling te komen. $\square$

VOORBEELD 2. "De prinses en het monster".

Ook dit is een achtervolgingsspel, waarbij het nu echter volkomen donker is; de spelers kunnen elkaar niet zien (of horen) en weten niet waar hun tegenstander is, tenzij ze tegen elkaar aanbotsen. De prinses en het monster bewegen zich beide langs de omtrek van dezelfde cirkel. De begintoestand is een uniforme verdeling voor beide spelers. Dit is bekend aan de spelers, maar niet hun positie in het begin of gedurende het spel, totdat de botsing heeft plaats gevonden. De maximum snelheid van het monster is 1, in beide richtingen. De snelheid van de prinses is onbegrensd. Het monster wil zo snel mogelijk botsen, de prinses wil dit moment zo lang mogelijk uitstellen.

In dit geval kunnen de spelers niet op elkaar reageren; open-loop besturingen zijn dan ook de aangewezen besturingen om dit probleem op te lossen (vooropgesteld dat beide spelers wel de tijd steeds weten). $\square$

De *eindtijd* t_f voor (2.1) kan op verschillende manieren gegeven zijn. Het kan voorgeschreven zijn of bijv. een funktie van de toestandsvektor zijn. In het laatste geval is de eindtijd t_f bijv. het eerste moment dat de toestandsvektor een voorgeschreven gebied in de toestandsruimte binnendringt. Dit gebied wordt het *doelgebied* (*target*) genoemd.

Geassocieerd aan elke speler is een *criterium* (uitkomst, *performance index*, *cost function*) J_i , welke op verschillende manieren gegeven kan zijn. De testfuncties hangen alleen van de eindtoestand af:

$$(2.4) \quad J_i = g_i(x(t_f)), \quad i = 1, \dots, N,$$

of van het hele verloop van de toestand en besturingen:

$$(2.5) \quad J_i = \int_{t_0}^{t_f} f_{0i}(x, u_1, \dots, u_N) dt, \quad i = 1, \dots, N,$$

of van een combinatie van deze twee:

$$(2.6) \quad J_i = g_i(x(t_f)) + \int_{t_0}^{t_f} f_{0i}(x, u_1, \dots, u_N) dt;$$

steeds nemen we aan $f_{0i} \in C^1$.

Door te schrijven

$$g_i(x(t_f)) = g_i(x_0) + \int_{t_0}^{t_f} (\text{grad } g_i(x(t))) f(x(t), u_1, \dots, u_N) dt,$$

gaat (2.4) over in de vorm (2.5). Omgekeerd kunnen (2.5) en (2.6) worden overgevoerd naar de vorm (2.4) door het invoeren van een extra toestandsvariabele. Alle drie vormen (2.4), (2.5) en (2.6) zijn dus even algemeen en het hangt van een concrete probleemstelling af met welke vorm men het handigst werkt.

Een *N-persoons differentiaalspel* ontstaat als elke speler z'n eigen beslissingsvariabelen zodanig kiest, dat z'n eigen criterium geminimaliseerd wordt (mits dit minimum bestaat). Toch is de probleemstelling nog niet volledig; een oplossingsconcept moet nog gegeven worden; dit gebeurt in de volgende sectie.

Er zijn een ander soort spelen, waarbij niet een kostenfunctie gedefinieerd is. Een voorbeeld is een achtervolgersprobleem. De enige vraag is: kan de achtervolger de achtervolgde te pakken krijgen? In de toestandsruimte probeert een speler de toestandsvektor in een bepaald gebied te krijgen, de ander probeert dat te voorkomen. Zulke spelen noemt men wel

kwantitatieve spelen of "*games of kind*", in tegenstelling tot de *kwantitatieve spelen* of "*games of degree*" welke in dit verhaal verder behandeld zullen worden.

Publikaties met een (gedeeltelijk) inleidend karakter zijn [1], [2] en [3].

3. OPLOSSINGSCONCEPTEN

Men maakt onderscheid in 2-persoonsspelen en meer-persoons (d.w.z. meer dan twee-persoons)spelen. Bij de 2-persoonsspelen maakt men weer onderscheid tussen *nulsomspelen* en *niet-nulsomspelen*. Voor nulsomspelen geldt dat $J_1 + J_2 = \text{constant}$, d.w.z. onafhankelijk van de gebruikte besturingen. Voor niet-nulsomspelen geldt dat niet. Bij meer-persoonsspelen kan men ook hetzelfde onderscheid maken, maar het verschil in eigenschappen is daar minder duidelijk.

Om de afhankelijkheid van J_i m.b.t. $x_0, u_1, \dots, u_N$ aan te geven zullen we soms schrijven $J_i(x_0, u_1, \dots, u_N)$. Door een constante bij een criterium op te tellen wordt de aard van het differentiaalspel niet verstoord. Voor 2-persoons nulsomspelen neemt men dan ook meestal $J_1 + J_2 = 0$; $J \triangleq J_1 = -J_2$.

DEFINITIE. Bij een 2-persoons nulsom is u_1^*, u_2^* de optimale oplossing als geldt:

$$(3.1) \quad J(x_0, u_1^*, u_2) \leq J(x_0, u_1^*, u_2^*) \leq J(x_0, u_1, u_2^*),$$

voor alle (toegelaten) u_1, u_2 . Het "punt" u_1^*, u_2^* noemt men het *zadelpunt*.

STELLING. Als zowel (x_0, u_1^*, u_2^*) als $(x_0, \tilde{u}_1, \tilde{u}_2)$ *zadelpunten* zijn, dan zijn ook $(x_0, u_1^*, \tilde{u}_2)$ en $(x_0, \tilde{u}_1, u_2^*)$ *zadelpunten* en $J(x_0, u_1^*, u_2^*) = J(x_0, u_1^*, \tilde{u}_2) = J(x_0, \tilde{u}_1, u_2^*) = J(x_0, \tilde{u}_1, \tilde{u}_2)$.

BEWIJS. Volgt direkt uit de definitie.

Hoewel de optimale beslissingen dus niet eenduidig behoeven te zijn, is de uitkomst dit wel. We definiëren:

$$V(x_0) = J(x_0, u_1^*, u_2^*),$$

waarbij V de *waarde* (Value) van het spel is.

DEFINITIE. $u_1^*, \dots, u_N^*$ is Pareto-optimaal als voor iedere (toegelaten) verzameling $u_1, \dots, u_N$ geldt dat

$$J_i(x_0, u_1, \dots, u_N) = J_i(x_0, u_1^*, \dots, u_N^*), \quad i = 1, \dots, N,$$

of voor tenminste één i :

$$J_i(x_0, u_1, \dots, u_N) > J_i(x_0, u_1^*, \dots, u_N^*).$$

Men kan dus stellen dat er geen andere oplossing is die beter of gelijk is voor alle spelers dan de Pareto-oplossing.

DEFINITIE. $u_1^*, \dots, u_N^*$ is Nash-optimaal als voor iedere i :

$$J_i(x_0, u_1^*, \dots, u_N^*) \leq J_i(x_0, u_1^*, \dots, u_{i-1}^*, u_i, u_{i+1}^*, \dots, u_N^*),$$

voor alle (toegelaten) u_i .

Het is gemakkelijk in te zien dat, als de Nash-definitie genomen wordt voor een 2-persoons nulsomspel, deze definitie overeenkomt met de zadel-puntdefinitie. Bij een tweepersoons nulsomspel is elke oplossing een Pareto-oplossing.

Meer over oplossingsconcepten kan men vinden in [1], [2] en [4].

4. INFORMATIESTRUKTUREN

In sectie 2 hebben we gezien dat u_i een functie van de tijd of toestand kan zijn. In het algemeen zal een speler alle informatie die hij heeft, gebruiken om tot een zo goed mogelijke keuze van z'n besturing te komen. Heeft hij alleen de tijd als informatie, dan spreekt men van open-loop besturing; kent hij alleen de tegenwoordige toestandsvektor (en de tijd), dan spreekt men van een closed-loop besturing. De theorie die de

invloed van verschillende informatie op de uitkomst van het spel beschrijft, noemt men de theorie der informatiestrukturen. Aangenomen wordt dat altijd aan de spelers bekend is de toestandsvergelijking (2.1), de beginvoorwaarde (2.2), de eindvoorwaarde, de criteria (d.w.z. elke speler kent alle criteria), en bovendien welke informatie iedere speler heeft om tot een beslissing te komen.

Aan de hand van het volgende voorbeeld, dat wegens het discrete karakter van de tijd een *differentiespel* zou kunnen worden genoemd, zullen we verschillende informatiestrukturen nagaan (x , u_1 en u_2 zijn scalars):

$$(4.1) \quad x(i+1) = x(i) + u_1(i) + u_2(i), \quad i = 0, 1; \quad x(0) = x_0;$$

$$(4.2) \quad J_1 = 2x^2(1) + 2x^2(2) + u_1^2(0) + u_1^2(1);$$

$$(4.3) \quad J_2 = x^2(1) + x^2(2) + u_2^2(0) + u_2^2(1).$$

We zullen alleen Nash-oplossingen beschouwen, d.w.z. $u_1^*(0)$, $u_2^*(1)$ en $u_2^*(0)$, $u_1^*(1)$ is optimaal als

$$(4.4) \quad J_1(x_0, u_1^*(0), u_1^*(1), u_2^*(0), u_2^*(1)) \leq J_1(x_0, u_1(0), u_1(1), u_2^*(0), u_2^*(1)),$$

$$J_2(x_0, u_1^*(0), u_1^*(1), u_2^*(0), u_2^*(1)) \leq J_2(x_0, u_1^*(0), u_1^*(1), u_2(0), u_2(1)),$$

voor alle $u_i(j)$, $i = 1, 2$; $j = 0, 1$. Dit noemt men soms wel de *globale Nash-oplossing*, ter onderscheiding van de *stapsgewijze Nash-oplossing*, waarbij men in de rechterleden van (4.4) slechts componentsgewijze variaties toelaat. Iedere globale Nash-oplossing is ook stapsgewijze Nash, maar het omgekeerde hoeft niet te gelden. Men berekent meestal de stapsgewijze Nash-oplossing in de hoop daarna te kunnen bewijzen dat deze oplossing ook globaal is.

Open-loop informatiestrukturen. De beslissingen $u_i(j)$ hangen alleen van de tijd j af. De vergelijkingen (4.1) worden in (4.2) en (4.3) gesubstitueerd:

$$(4.5) \quad J_1 = 2\{x_0 + u_1(0) + u_2(0)\}^2 + 2\{x_0 + u_1(0) + u_2(0) + u_1(1) + u_2(1)\}^2 + \\ + u_1^2(0) + u_1^2(1),$$

$$(4.6) \quad J_2 = \{x_0 + u_1(0) + u_2(0)\}^2 + 2\{x_0 + u_1(0) + u_2(0) + u_1(1) + u_2(1)\}^2 + \\ + u_2^2(0) + u_2^2(1).$$

Door nu te stellen $\partial J_1 / \partial u_1(j) = 0$; $\partial J_2 / \partial u_2(j) = 0$, $j = 0, 1$, krijgt men vier vergelijkingen met vier onbekenden, waarvan de oplossing is

$$(4.7) \quad \left. \begin{aligned} u_1(0) &= -\frac{10}{19} x_0; & u_1(1) &= -\frac{1}{19} x_0; \\ u_2(0) &= -\frac{5}{19} x_0; & u_2(1) &= -\frac{1}{19} x_0. \end{aligned} \right\}$$

Door de constructie is deze oplossing stapsgewijze-Nash. Het blijkt dat $\partial^2 J_j / \partial u_j(0) \partial u_j(1) > 0$ voor $j = 1, 2$ en daarom is deze oplossing tevens globaal-Nash.

Closed-loop informatiestrukturen. De beslissingen $u_i(0)$ hangen af van x_0 ; de beslissingen $u_i(1)$ hangen af van $x(1)$. De oplossingen worden verkregen door het principe van het dynamisch programmeren toe te passen. De beslissingen $u_i(1)$, $i = 1, 2$, beïnvloeden alleen het volgende deel van de criteria:

$$\begin{aligned} \tilde{J}_1 &= 2x^2(2) + u_1^2(1) = 2\{x(1) + u_1(1) + u_2(1)\}^2 + u_1^2(1), \\ \tilde{J}_2 &= x^2(2) + u_2^2(1) = \{x(1) + u_1(1) + u_2(1)\}^2 + u_2^2(1). \end{aligned}$$

De vergelijkingen $\partial \tilde{J}_j / \partial u_j(1) = 0$, $j = 1, 2$, geven nu:

$$(4.8) \quad u_1(1) = -\frac{1}{2}x(1), \quad u_2(1) = -\frac{1}{4}x(1).$$

Om nu de optimale $u_j(0)$, $j = 1, 2$, te verkrijgen wordt (4.8) gesubstitueerd in J_1 en J_2 . Ook wordt (4.1) gesubstitueerd in J_1 en J_2 , die op deze manier functies worden van x_0 en $u_j(0)$, $j = 1, 2$. Door J_1 en J_2 te minimaliseren m.b.t. resp. $u_1(0)$ en $u_2(0)$ krijgen we

$$(4.9) \quad u_1(0) = -\frac{19}{36} x_0, \quad u_2(0) = -\frac{1}{4} x_0.$$

Vergelijkingen (4.8) en (4.9) vormen samen de closed-loop Nash-oplossing.

Uiteraard zijn de closed-loop en de open-loop oplossing niet dezelfde. Een verstoring in $u_j(0)$ werkt niet door in de numerieke waarde van $u_j(1)$ ingeval de structuur open-loop is; bij closed-loop oplossingen zal een verstoring in $u_j(0)$ aanleiding geven tot andere $u_j(1)$ -waarden.

De closed-loop oplossingen hadden ook verkregen kunnen worden uit (4.5) en (4.6), door daar de volgende vergelijkingen uit te destilleren:

$$(4.9a) \quad \begin{cases} \partial J_j / \partial u_j(1) = 0, & j = 1, 2 \\ \partial J_j / \partial u_j(0) + \partial J_j / \partial u_j(1) + \partial J_j / \partial u_i(1) = 0, & i, j = 1, 2; i \neq j. \end{cases}$$

De oplossing $u_j(i)$ worden dan funkties van x_0 . Wil men deze oplossing in de vorm van (4.8), (4.9) hebben, dan moet in de uitkomsten voor $u_j(1)$ nog gesubstitueerd worden: $x_0 = x(1) - u_1(1) - u_2(1)$.

Stackelberg open-loop informatiestructuur. Hierbij is sprake van een leider en een volger. De leider kondigt zijn beslissingen aan, waarna de volger zijn beslissingen gaat nemen. Als we aannemen dat de tweede speler, P_2 , de leider is. Hij kiest $u_2(0)$ en $u_2(1)$ als funkties van x_0 en maakt deze keuzen bekend aan P_1 , welke vervolgens $u_1(0)$ en $u_1(1)$ kiest. Als P_1 eenmaal weet wat $u_2(0)$ en $u_2(1)$ zijn, worden deze gesubstitueerd in (4.1) en P_1 moet nu een standaard optimalisatieprobleem oplossen: minimaliseer J_1 met nevenvoorwaarde (4.1). De oplossing daarvan is

$$(4.10) \quad \begin{cases} u_1(0) = -\frac{1}{11} (2x_0 - 2u_2(0) - 6u_2(1)) \\ u_1(1) = -\frac{1}{11} (8x_0 - 8u_2(0) - 2u_2(1)). \end{cases}$$

Voordat P_2 z'n beslissingen bekend maakt zal hij zich realizeren dat P_1 daarna z'n beslissingen volgens (4.10) gaat kiezen. Om daarom de gunstigste $u_1(0)$, $u_1(1)$ te vinden gaat P_2 z'n kosten J_2 minimalizeren, met als nevenvoorwaarde (4.1), waarin (4.10) gesubstitueerd is. De oplossing daarvan is:

$$u_2(0) = -\frac{11}{145} x_0; \quad u_2(1) = \frac{3}{145} x_0.$$

Stackelberg closed-loop informatiestructuur. Neem weer aan P_2 is de leider, P_1 is de volger. Ook nu maakt P_2 z'n beslissingen van te voren bekend; $u_2(0)$ als een functie van x_0 ; $u_2(1)$ als een functie van $x(1)$. Het blijkt dat als we deze functies niet nader specificeren, dat uiteindelijk een stelsel gecompliceerde funktionaalvergelijkingen overblijft. Om de berekening expliciet te kunnen uitvoeren nemen we aan dat

$$(4.11) \quad u_2(0) = -\alpha_0 x(0), \quad u_2(1) = -\alpha_1 x(1),$$

waarbij α_0 en α_1 door P_2 gekozen moeten worden. Als P_2 α_0 en α_1 gekozen heeft en bekend gemaakt aan P_1 , dan moet P_1 weer een standaard optimalisatieprobleem oplossen in closed-loop vorm. De oplossing is:

$$(4.12) \quad u_1(1) = -\frac{2}{3}(x(1) - \alpha_1 x(1)); \quad u_1(0) = \frac{\mu(6+2x^2)}{9+2x^2} x_0,$$

waarbij $\mu \triangleq 1 - \alpha_0$; $x = 1 - \alpha_1$. Voordat P_2 α_0 en α_1 bekend maakt realiseert hij zich dat P_1 gaat spelen volgens (4.12). Om nu α_0 en α_1 zo gunstig mogelijk te kiezen gaat P_2 de functie J_2 in (4.3) minimalizeren, met als nevenvoorwaarde (4.1) waarin (4.12) gesubstitueerd is. Deze nevenvoorwaarden kunnen verwerkt worden in J_2 , die dan een functie wordt van x_0 , α_0 en α_1 :

$$J_2 = \left[\alpha_0^2 + \left\{ 1 + \alpha_1 + \frac{1}{9}(1-\alpha_1)^2 \right\} \cdot \left\{ \frac{3(1-\alpha_0)^2}{9+2(1-\alpha_1)^2} \right\}^2 \right] x_0.$$

Door nu te kiezen $\alpha_0 = 0$, $|\alpha_1| \rightarrow \infty$ wordt $J_2 \rightarrow 0$. De Stackelberg closed-loop oplossing bestaat hier dus niet in wiskundige zin.

Bij dit voorbeeld kan men meer informatiestructuren aangeven, bijv. P_1 open-loop, P_2 closed-loop. Dat zal hier niet verder worden uitgewerkt. Wel wordt nu aandacht besteed aan de *eenduidigheid* van de oplossing, eveneens m.b.t. bovenstaand voorbeeld. De oplossing waarbij beide spelers beiden òf open-loop òf closed-loop spelen zijn eenduidig. Men kan de klasse van closed-loop oplossingen uitbreiden door te nemen dat de huidige beslissing niet alleen afhankelijk is van de huidige tijd en de huidige toestand,

maar ook van de hele voorgeschiedenis van de toestand. In het voorbeeld zou dit betekenen dat $u_j(1)$ van zowel $x(1)$ als x_0 afhankelijk mag zijn. Om deze closed-loop besturingen te kunnen onderscheiden spreekt men wel van *closed-loop besturingen zonder, resp. met geheugen*.

Aan de hand van bovenstaand voorbeeld zullen we laten zien dat, als we als informatiestructuur voor beide spelers nemen closed-loop met geheugen, de oplossing niet meer eenduidig is. In plaats van (4.8) gaan we nu uit van

$$(4.12) \quad \begin{cases} u_1(1) = -\frac{1}{2}x(1) + p(x(1) - \bar{x}(1)), \\ u_2(1) = -\frac{1}{4}x(1) + k(x(1) - \bar{x}(1)), \end{cases}$$

waarbij p en k constanten zijn en waarbij

$$\bar{x}(1) = x_0 + \gamma_1^0(x_0) + \gamma_2^0(x_0),$$

waarbij γ_1^0 en γ_2^0 , gelijk aan $u_1(0)$ en $u_2(0)$ resp., nog bepaald moeten worden. Merk op dat voor $p, k \neq 0$ de vergelijking (4.12) functies zijn van $x(1)$ en x_0 . Merk ook op dat de numerieke waarde van $u_1(1)$ niet veranderd is door de toevoeging van de extra-term; alleen de representatie is veranderd. Een soortgelijke opmerking geldt voor $u_2(1)$. Daarom, als we eenmaal ten tijde 1 zijn aangeland, is (4.12) de optimale oplossing voor het nog resterende deel van het spel. Om nu $\gamma_1^0(x_0)$ en $\gamma_2^0(x_0)$ te vinden, wordt (4.12) in de criteria gesubstitueerd:

$$\begin{aligned} J_1 &= 2\{x_0 + u_1(0) + u_2(0)\}^2 + 2\{(\frac{1}{4} + p + k)(x_0 + u_1(0) + u_2(0)) - \\ &\quad (p + k)(x_0 + \gamma_1^0(x_0) + \gamma_2^0(x_0))\}^2 + \\ &\quad u_1^2(0) + \{(-\frac{1}{2} + p)(x_0 + u_1(0) + u_2(0)) - p(x_0 + \gamma_1^0(x_0) + \gamma_2^0(x_0))\}^2; \\ J_2 &= \{x_0 + u_1(0) + u_2(0)\}^2 + \{(\frac{1}{4} + p + k)(x_0 + u_1(0) + u_2(0)) - \\ &\quad (p + k)(x_0 + \gamma_1^0(x_0) + \gamma_2^0(x_0))\}^2 + \\ &\quad u_2^2(0) + \{(-\frac{1}{4} + p)(x_0 + u_1(0) + u_2(0)) - k(x_0 + \gamma_1^0(x_0) + \gamma_2^0(x_0))\}^2. \end{aligned}$$

J_j wordt nu gedifferentieerd naar $u_j(0)$. In deze afgeleiden, welke nul moeten zijn voor de optimale oplossing, substitueert men $\gamma_j^0(x_0) = u_j(0)$, $j = 1, 2$. Op deze manier verkrijgt men 2 vergelijkingen met de 2 onbekenden $u_1(0)$ en $u_2(0)$. De oplossing is

$$u_1(0) = -\frac{19+4k}{2\lambda} x_0; \quad u_2(0) = -\frac{9+2p}{2\lambda} x_0,$$

waarbij $\lambda \triangleq 18+p+k$. Tesamen met

$$u_1(1) = (-\frac{1}{2}+p)x(1) - \frac{4p}{\lambda} x_0;$$

$$u_2(1) = (-\frac{1}{4}+k)x(1) - \frac{4k}{\lambda} x_0,$$

is dit de Nash-oplossing. Voor verschillende keuzen van p en k krijgt men verschillende Nash-oplossingen. Voor $p = k = 0$ heeft men de closed-loop, zonder geheugenoplossing. Voor $p = \frac{1}{2}$, $k = \frac{1}{4}$ heeft men de open-loop oplossing. Voor willekeurige p en k heeft men een optimale oplossing "ergens tussen die van open- en closed-loop".

Men kan niet-lineaire Nash-oplossingen krijgen door niet van (4.12) uit te gaan, maar van

$$u_1(1) = -\frac{1}{2}x(1) + \bar{f}_1(x(1), \bar{x}(1));$$

$$u_2(1) = -\frac{1}{4}x(1) + \bar{f}_2(x(1), \bar{x}(1)),$$

met $\bar{f}_i \in C^1$ en $\bar{f}_i(a, a) = 0$ voor willekeurige a .

De niet-eenduidigheid van de Nash-oplossing is uiteraard een zwakte van de Nash-definitie. Als men echter stochastische verstoringen aan de systeemvergelijkingen toevoegt (verstoringen welke niet een één-puntsverdeling hebben), kan men de bovenstaande constructie van een twee-parameterfamilie van optimale oplossingen niet meer toepassen, omdat dan in het algemeen $x(1) \neq \bar{x}(1)$. Ook al laat men dan closed-loop informatiestrukturen met geheugen toe, de Nash-oplossing blijkt dan weer eenduidig te zijn [6].

OPMERKING. Bovenstaand voorbeeld is niet-nulsom. Als we met een nulsomspel

te maken hebben, geven de behandelde informatiestrukturen dan nog aanleiding tot verschillende uitkomsten? Antwoord: pas op, in het algemeen ja. Bij "de dame in het meer" zal de open-loop informatiestruktuur duidelijk andere antwoorden geven dan de closed-loop informatiestruktuur. Bij een nulsomspel qua "vorm" gelijk aan het niet-nulsomspel in deze sectie behandeld, zullen de uitkomsten bij closed-loop informatiestruktuur en open-loop informatiestruktuur dezelfde zijn. Een voorwaarde die dan vervuld moet zijn is dat de open-loop oplossing eenduidig is. Dit kan men zien aan de vergelijkingen (4.9a). Voor een nulsomspel zijn dan $\partial J_j / \partial u_j(1) = \partial J_j / \partial u_i(1) = 0$ en gaat (4.9a) over in

$$\partial J_j / \partial u_j(1) = 0, \quad j = 1, 2$$

$$\partial J_j / \partial u_j(0) = 0, \quad j = 1, 2,$$

hetgeen dezelfde voorwaarden zijn als die van de open-loop informatiestruktuur.

Andere informatiestrukturen m.b.t. hetzelfde voorbeeld vindt men in [5]. In [6] wordt een discussie gegeven over de niet-eenduidigheid.

5. TWEE-PERSOONS NULSOMPELEN; THEORIE

We gaan uit van (2.1), (2.2) en een kostenfunctie V ($=V_1 = -V_2$) van de vorm (2.6). De spelers worden nu P (Pursues) i.p.v. P_1 en E (Evader) i.p.v. P_2 genoemd, die nu de strategieën u resp. v tot hun beschikking hebben. De optimale strategieën u^* , v^* voldoen per definitie aan de zadel-puntvoorwaarde.

STELLING. Een noodzakelijke voorwaarde voor u^*, v^* is

$$(5.1) \quad \min_u \max_v \{f_0 + \nabla V \cdot f\} = 0,$$

$V(x) = g(x)$ in het doelgebied, overal waar ∇V bestaat; V is de waarde van het spel.

$$f_0 \triangleq f_{01} = -f_{02}; \quad g \triangleq g_1 = -g_2.$$

Aanname is dat de waarde van het spel bestaat.

BEWIJS. Dit bewijs gaat met het argument van het dynamisch programmeren. Neem aan dat ten tijde t , $t_0 \leq t < t_f$, het punt x in de toestandsruimte is bereikt. Voor dit punt x is de waarde $V(x)$. Even later, op tijdstip $t+h$, is het systeem iets verplaatst naar $x + \Delta x$. Dan geldt

$$J(x, u, v) = \int_t^{t+h} f_0(x, u, v) dt + V(x + \Delta x),$$

vanwaar

$$J(x, u, v) = f_0(x, u, v)h + V(x) + \nabla V \cdot f \cdot h + o(h^{-1}).$$

Door nu links en rechts de $\min_u \max_v$ operatie toe te passen krijgen we:

$$V(x) = V(x) + h \cdot \min_u \max_v [f_0(x, u, v) + \nabla V \cdot f + o(h^{-1})],$$

hetgeen precies (5.1) oplevert als we $h \downarrow 0$ laten gaan. $\square$

Vergelijking (5.1) wordt wel de *vergelijking van Isaacs* genoemd.

Een andere afleiding van (5.1) kan worden gegeven m.b.v. het begrip *half-doorlaatbare (semipermeable) oppervlakken*. Laat lokaal een $(n-1)$ -dimensionaal oppervlak in de buurt van een punt $x \in \mathbb{R}^n$ gegeven zijn. Dit oppervlak is halfdoorlaatbaar m.b.t. v in x als er tenminste één waarde $\bar{u}$ van u bestaat, zodanig dat, als $u = \bar{u}$, er geen vektoren v bestaan zodat de vektor $f(x, \bar{u}, v)$, opgericht in x , naar beide kanten van dit oppervlak wijst. De zijde van dit oppervlak, waarin geen enkele $f(x, \bar{u}, v)$ wijst, wordt de E-kant van dit oppervlak genoemd. Soortgelijk heet dit oppervlak halfdoorlaatbaar in x m.b.t. u als er tenminste één $\bar{v}$ bestaat, zodanig dat $f(x, u, \bar{v})$, voor alle u steeds naar één kant van het oppervlak wijst. De andere kant heet de P-kant. Het oppervlak heet halfdoorlaatbaar in x als het zowel half-doorlaatbaar m.b.t. u als v is en als E- en P-kant niet dezelfde zijn. Een oppervlak, dat in ieder punt deze eigenschap heeft, heet halfdoorlaatbaar.

Beschouw nu de kostenfunctie van de vorm (2.4). Neem aan $V(x)$ heeft tenminste twee verschillende waarden. Ieder oppervlak dat de verzamelingen

$$\{x \mid V(x) > c\} \quad \text{en} \quad \{x \mid V(x) < c\}$$

separeert, x is een constante, is halfdoorlaatbaar met de kleinere V aan de P -kant. Als bijv. bij punt x van dit oppervlak er geen $\bar{u}$ zou bestaan, dan zou P niet kunnen voorkomen dat E de toestand naar punten met grotere V -waarden zou trekken. Zo ook bestaat er een $\bar{v}$.

Dus voor x op het oppervlak moeten de spelers de strategieën $\bar{u}$ en $\bar{v}$ gebruiken en deze zijn optimaal.

Veronderstel $V \in C^1$ en niet constant in een n -dimensionale omgeving. $V(x) = \text{constant}$ is nu een halfdoorlaatbaar oppervlak met als normaal ∇V . De voorwaarde voor de spelers om op dit oppervlak te blijven is dus:

$$(5.2) \quad \min_u \max_v \nabla V \cdot f = 0,$$

hetgeen eenzelfde voorwaarde is als in (5.1). De vorm is verschillend, omdat verschillende vormen van kostfuncties beschouwd zijn.

OPMERKING 1. De halfdoorlaatbare oppervlakken zijn volledig gekarakteriseerd door de toestandsvergelijkingen; het criterium heeft er niets mee te maken. $\square$

OPMERKING 2. De noodzakelijke voorwaarden in (5.1) of (5.2) zijn lokaal, in tegenstelling tot de zadelpuntvoorwaarde (3.1) waar de "hele" strategieën met elkaar vergeleken worden. Voorwaarden (5.1) of (5.2) zijn geschikter voor de werkelijke constructie van optimale oplossingen, vooropgesteld dat we iets van $V(x)$ te weten kunnen komen. $\square$

Beschouw nu (5.1), die wel de *Hamiltoniaan* H genoemd wordt:

$$(5.3) \quad H = f_0(x, u^*, v^*) + \frac{\partial V}{\partial x} f(x, u^*, v^*) = 0.$$

We beschouwen closed-loop oplossingen $u^*(x)$, $v^*(x)$ en gaan (5.3) differentieren naar x ;

$$(5.4) \quad \frac{d}{dx} \left(\frac{dV}{dx} \right) f + \frac{dV}{dx} \frac{\partial f}{\partial u} + \frac{\partial f_0}{\partial x} + \frac{\partial H}{\partial u^*} \frac{du^*}{dx} + \frac{\partial H}{\partial v^*} \frac{dv^*}{dx} = 0.$$

Van de eerste term in (5.4) gaan we de differentiatievolgorde veranderen. De laatste termen zullen meestal nul zijn. Resultaat:

$$(5.5) \quad \frac{d}{dt} \frac{dV}{dx} = - \frac{\partial H}{\partial x};$$

dit wordt de *geadjungeerde vergelijking* genoemd. Deze vergelijking geldt alleen daar langs de optimale baan $x(t)$, waar alle tweede afgeleiden in (5.4) bestaan. De eindvoorwaarde van (5.5) is af te leiden uit $V(x) = g(x)$ langs de rand van het doelgebied.

Als we een willekeurig punt van het doelgebied nemen, dan kunnen de toestandsvergelijkingen en de geadjungeerde vergelijkingen samen worden "teruggeïntegreerd" in de tijd, waarbij (5.1) steeds de optimale u en v geeft. Door dit met elk punt van het doelgebied te doen wordt de toestandsruimte opgevuld met banen, die kandidaat staan voor optimale banen. Hier is een essentiële moeilijkheid, omdat men de vergelijkingen wel negatief terug kan integreren, maar dat in de werkelijke optimale oplossing een discontinuïteit in ∇V zit, die men nu continu voorbij integreert. Daarom geven de noodzakelijke voorwaarden alleen kandidaat-optimale oplossingen. Alleen als blijkt dat deze oplossingen achteraf aan de zadelpuntvoorwaarde voldoen, zijn ze optimaal. De genoemde discontinuïteiten kunnen van allerlei aard zijn en de theorie daarover is nog niet volledig ontwikkeld. De verzameling van alle discontinuïteitpunten in de toestandsruimte noemt men de *singuliere oppervlakken*. Voldoende voorwaarden bestaan ook. Of deze zijn tamelijk triviaal en veronderstellen daarom vele voorwaarden m.b.t. $V(x)$, die alle discontinuïteiten uitsluiten. Of deze zijn ingewikkeld, maar kunnen vele soorten discontinuïteiten aan [7].

Voor meer informatie over dit onderwerp: [3], [7] en verschillende artikelen van BREAKWELL, zoals bijv. [8]. Wiskundig gezien strenger, maar aan ideeën armer is [10].

6. TWEE-PERSOONS NULSOMSPLEN; TOEPASSINGEN

Toepassing 1. Beschouw de toestandsvergelijking:

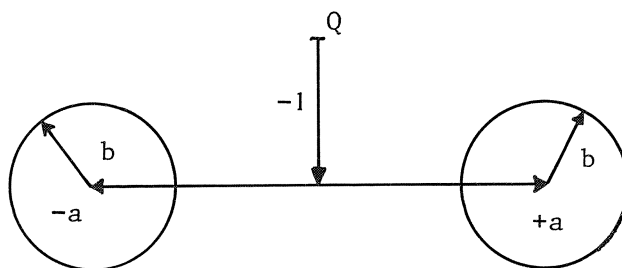
$$\dot{x}_1 = av + b \sin u,$$

$$\dot{x}_2 = -1 + b \cos u.$$

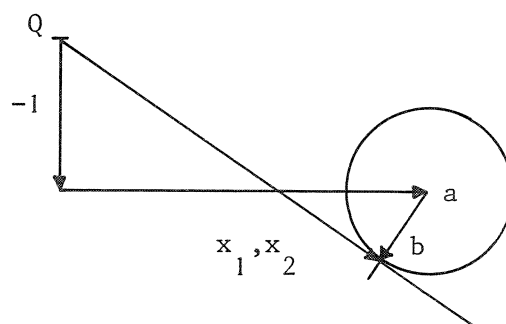
De spelruimte is het halfvlak $x_2 \geq 0$ en het doelgebied is de x -as. De parameters a en b zijn constanten met $0 < b < 1$ en $a > b$. De besturing u kan vrij gekozen worden; v moet voldoen aan $|v| \leq 1$. Het criterium is

$$(6.1) \quad J = x_f.$$

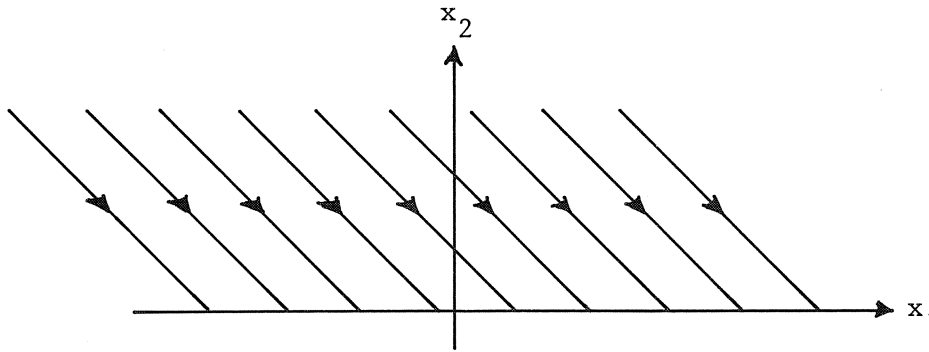
Kan elementair worden opgelost m.b.v. een *vectorgram*:



De snelheidsvektor $(\dot{x}_1, \dot{x}_2)$ zal, mits $v = \pm 1$, altijd een richting en grootte hebben van een vektor beginnend in Q en eindigend op een van de twee cirkels. E wil x_f maximaliseren en zal daarom $v = \pm 1$ kiezen. P zal de richting van b zo kiezen, dat de snelheidsvektor $(\dot{x}_1, \dot{x}_2)$ zo ver mogelijk naar links wijst, dus zoals in onderstaande figuur aangegeven:



De oplossingen u en v zijn niet toestandsafhankelijk en terug integreren van de toestandsvergelijking levert de volgende figuur:



Als men de theorie uit voorgaande sectie toepast krijgt men:

$$\min_u \max_v V_x(av + b \sin u) + V_y(-1 + b \cos u) = 0$$

$$\rightarrow v^* = \operatorname{sgn} V_x; \quad (\sin u^*, \cos u^*) \parallel (-V_x, -V_y).$$

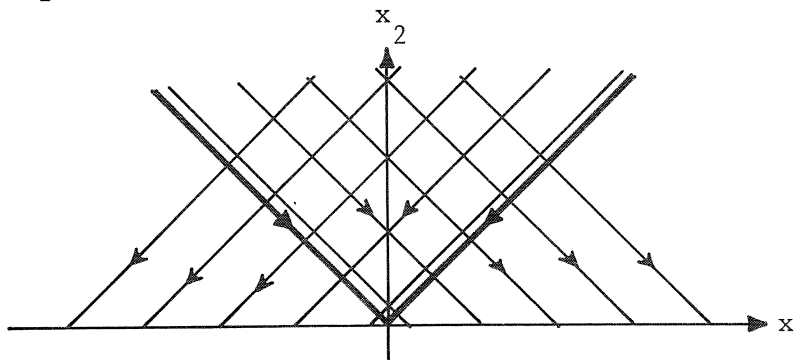
Eindvoorwaarden zijn: $V_x(x_f, 0) = 1$; wat is $V_y(x_f, 0)$?

$V_y(x_f, 0)$ krijgt men door $V_x(x_f, 0) = 1$ in te vullen in de Hamiltoniaan. Omdat deze Hamiltoniaan nul is krijgt men een waarde voor $V_y(x_f, 0)$. Nu kan men gaan terugintegreren; de banen zien er uit als in bovenstaande figuur. Men kan gemakkelijk nagaan dat deze oplossingen aan de zadelpuntvoorwaarde voldoen en daarom optimaal zijn.

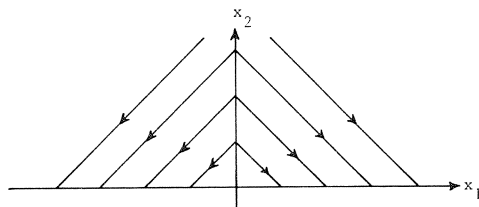
Toepassing 2. Beschouw hetzelfde probleem als boven, met alleen nu voor het criterium:

$$J = x_f^2.$$

Voor eindpunten $(x_f, 0)$ met $x_f > 0$ vindt men dezelfde banen als boven; voor eindpunten met $x_f < 0$ vindt men de gespiegelde banen, met als resultaat:



Het kegelvormig gebied met schijnbaar niet-eenduidige oplossingen laat nu het gevaar zien van "klakkeloos" terugintegreren. Binnen deze kegel en voor punten $x_1 \neq 0$ is een van de twee oplossingen duidelijk niet optimaal als we de uitkomst bekijken. Blijft het volgende plaatje over:

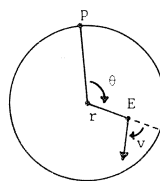


De x_2 -as is nu een "*dispersal line*", een lijn waarlangs ∇V niet continu is. Als het beginpunt op de x_2 -as ligt kan E beslissen welke kant op te gaan; daarna liggen de optimale strategieën eenduidig vast.

Toepassing 3. "De dame in het meer". P heeft de snelheid 1, en rent langs de rand van het meer dat straal R heeft. E heeft de snelheid v_0 en zwemt in het meer. De toestandsvergelijkingen worden hier in poolcoördinaten gegeven; θ is de hoek tussen P en E ($-\pi \leq \theta \leq +\pi$);

$$\dot{\theta} = \frac{v_0 \sin v}{r} - \frac{u}{R}; \quad u = \pm 1;$$

$$\dot{r} = v_0 \cos v.$$



De kostenfunctie is $g(\theta_f, r_f) = |\theta_f|$ (index f betekent final). De Isaacs vergelijking levert:

$$0 = \min_u \max_v \left\{ V_r v_0 \cos v + V_\theta \left(\frac{v_0 \sin v}{r} - \frac{u}{R} \right) \right\},$$

$$|u| = 1$$

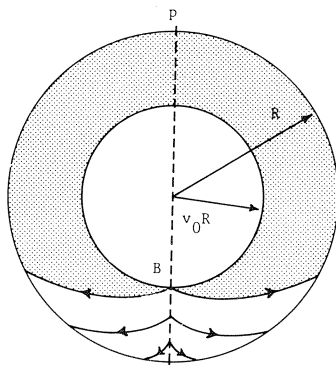
waaruit volgt:

$$(6.2) \quad \begin{cases} u^* = \operatorname{sgn} V_\theta = \text{constant} = \operatorname{sgn} \theta_f, \\ (\cos v^*, \sin v^*) \parallel (V_r, \frac{1}{r} V_\theta). \end{cases}$$

Door (6.2) in te vullen in de Isaacs vergelijking krijgen we

$$(6.3) \quad v_0 - \frac{r}{R} \sin v^* \operatorname{sgn} \theta_f = 0,$$

waaruit volgt dat E in een rechte lijn zwemt, welke raakt aan een cirkel met straal $v_0 R$ en middelpunt = middelpunt van het meer. Vergelijking (6.3) geldt niet als E binnen deze $v_0 R$ -cirkel zwemt. Binnen deze cirkel kan E een grotere hoeksnelheid ontwikkelen en daardoor kan E zwemmen naar de positie $\theta = \pi$, $r = v_0 R$ vanwaar zij langs één van beide kanten langs de raaklijn gaat zwemmen. Het punt $\theta = \pi$, $r = v_0 R$ is een zgn. *beslissingspunt*; strikt genomen kan E daar niet komen zonder de huidige besturing van P te kennen; ze kan er echter willekeurig dicht bij komen en blijven. Van sommige beginposities kan E beter direct naar de kant zwemmen i.p.v. eerst terug te zwemmen naar binnen de $v_0 R$ -cirkel. Zie onderstaande figuur. De lijn $\theta = \pi$ is weer een dispersal line (nu voor P).



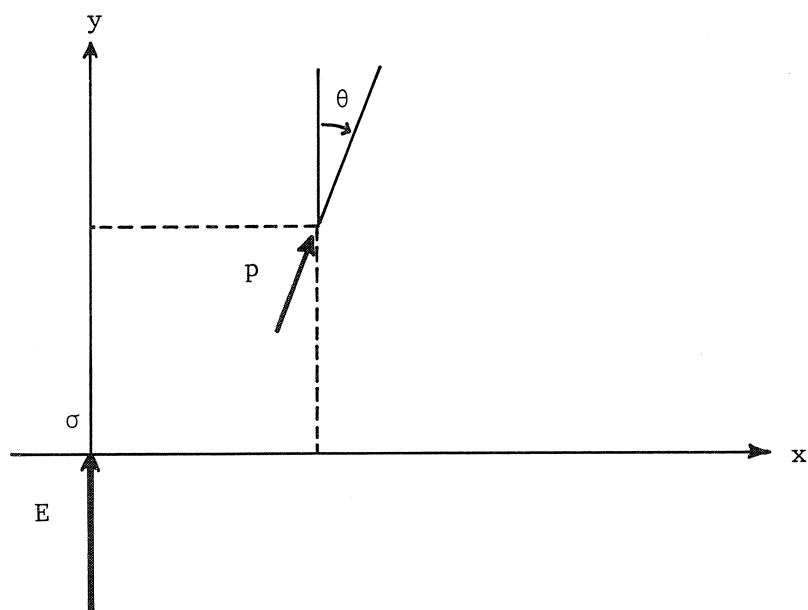
Toepassing 4. Het vermijden van botsingen bij schepen.

Gegeven zijn twee schepen in elkaars buurt in de open zee.

Kan een botsing vermeden worden?

We gaan een zgn. *worst-case design* maken; schip E probeert een botsing te vermijden; schip P probeert er één te veroorzaken. De toestandsruimte zal in twee gebieden worden verdeeld. Als de relatieve beginpositie van de schepen in het ene gebied ligt zal een botsing optreden, in het andere gebied niet. De gebieden zullen in deze toepassing worden afgebakend.

Het model is als volgt: schip E is een lijnsegment met lengte ℓ_1 ; voortwaartse snelheid in de richting van het lijnsegment v_1 ; maximale hoeksnelheid w_1 ; heeft het draaipunt voor aan de boeg. Schip P heeft een soortgelijke dynamica met parameters ℓ_2 , v_2 en w_2 . Een botsing is per definitie een overlap van de twee lijnstukjes. De hoeksnelheden zijn de respectievelijke besturingen u en v .



In het relatieve coördinatenstelsel (y langs E , $x \perp y$ en O op de boegplaats van E ; θ relatieve hoek) zijn de toestandsvergelijkingen:

$$\dot{x} = -w_1 u_1 y + v_2 \sin \theta,$$

$$\dot{y} = w_1 u_1 x - v_1 + v_2 \cos \theta,$$

$$\dot{\theta} = -w_1 u_1 + w_2 u_2,$$

waarbij $|u_i| \leq 1$; $u_i = +1$ is zo sterk mogelijk naar rechts varen (in de werkelijke situatie op zee), $u_i = 0$ is rechttuit varen, $u_i = -1$ is zo sterk mogelijk naar links varen.

Het criterium J hangt alleen van de eindtoestand af en is niet exact gedefinieerd. De volgende eigenschappen voor het criterium zijn voldoende om dit probleem op te lossen:

$$\min_{u_1} \max_{u_2} J;$$

$J > 0$: betekent een botsing,

$J < 0$: betekent geen botsing,

$J = 0$: de schepen schampen langs elkaar.

De Isaacs-vergelijking luidt:

$$\min_{u_1} \max_{u_2} \{V_x \dot{x} + V_y \dot{y} + V_\theta \dot{\theta}\} = 0$$

en daarom

$$u_1 = \operatorname{sgn}(yV_x - xV_y + V_\theta),$$

$$u_2 = \operatorname{sgn} V_\theta.$$

De geadjungeerde vergelijkingen voldoen aan

$$\dot{V}_x = -w_1 u_1 V_y,$$

$$\dot{V}_y = w_1 u_1 V_x,$$

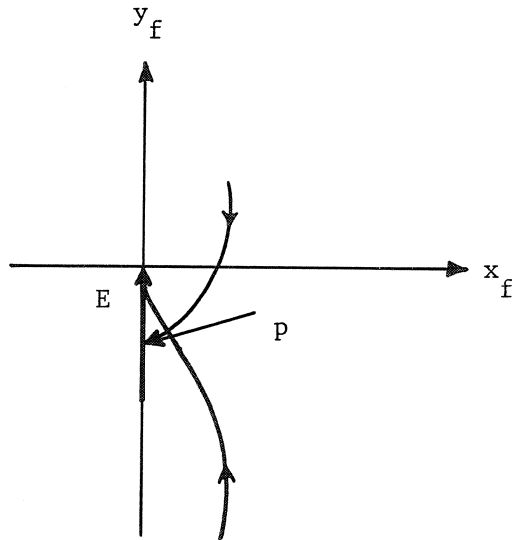
$$\dot{V}_\theta = v_2 (V_y \sin \theta - V_x \cos \theta).$$

In dit probleem kunnen singuliere banen opduiken als $V_\theta \equiv 0$ (kan alleen optreden als $u_2 \equiv 0$) en als $yV_x - xV_y + V_\theta \equiv 0$ (kan alleen optreden als $u_1 \equiv 0$).

We gaan nu verschillende deelproblemen oplossen; elk deelprobleem komt overeen met een bepaalde schampsituatie, gekarakteriseerd door de eindvoorwaarden op de toestandsvergelijkingen. Hierbij worden de eindvoorwaarden op de geadjungeerde variabelen berekend en de beide stelsels differentiaalvergelijkingen worden teruggeïntegreerd in de tijd. Elk paar van zo'n baan beschouwd als beginpunt van een werkelijk differentiaalspel zal dan aanleiding geven tot de bewuste schampsituatie en daarom zal deze baan deel uitmaken van het scheidingsoppervlak tussen het botsings- en het niet-botsingsgebied. Dit scheidingsoppervlak is halfdoorlaatbaar.

Een voorbeeld van een schampsituatie is als in het eindtijdstip t_f geldt: $x_f = 0$, $\dot{x}_f = 0$, waaruit volgt $\sin \theta_f = \frac{w_1 y_f}{v_2}$. Voor iedere y_f met grenzen tussen $\max(-\frac{v_2}{w_1}, -\ell_1) \leq y_f \leq 0$ treedt zo'n schampsituatie op. Er blijken achttien verschillende soorten schampsituaties te zijn. Voor de gedetailleerde uitwerking en oplossing worden verwezen naar [11].

De toepassingen 1, 2, 3 staan in [3]. Toepassing 4 in [11].



7. NIET-NULSOMSPELLEN; THEORIE

Gegeven de spelers $P_1, \dots, P_N$ die een interactie hebben via de toestandsvergelijking (2.1) met beginvoorwaarde (2.2), en die elk een kostenfunctie hebben van de vorm (2.5). Alleen Nash-oplossingen zullen worden behandeld. De Nash-definitie zegt dat, als alle andere dan de i -de spelers vasthouden aan hun Nash-besturing, dan de P_i zijn criterium J_i moet minimalizeren m.b.t. u_i^* . Speler P_i moet dan dus een standaard optimaal besturingsprobleem oplossen, waarin de optimale besturingen van de andere spelers als gegeven functies voorkomen. Voor iedere speler moet zo een optimaal besturingsprobleem worden opgelost om zo tot een stelsel impliciete vergelijkingen te komen, waaraan alle optimale besturingen moeten voldoen. We nemen aan dat $J_i(x, u_1^*, \dots, u_N^*)$ eenduidig is, ook al zijn de optimale besturingen dit mogelijkwijze niet, en geven deze functie aan met $V_i(x)$. Net als in sectie 5, kan nu bewezen worden dat

$$(7.1) \quad \min_{u_i} \{f_{0i} + \nabla V_i f\} = 0, \quad i = 1, \dots, N$$

overal waar ∇V_i bestaat. Omdat we alleen met een integraalcriterium werken, geldt voor de eindvoorwaarde dat $V_i(x) = 0$ op de rand van het doelgebied, en $\nabla V_i(x_f)$ staat $\perp$ op deze rand. Deze laatste voorwaarde wordt in de optimale besturingstheorie de *transversaliteitsvoorwaarde* genoemd. Als $V_i(x) \in C^2$, dan kunnen weer de geadjungeerde vergelijkingen worden afgeleid (zie (5.4)):

$$(7.2) \quad \frac{d}{dt} \frac{dV_i(x)}{dx} = - \frac{\partial H_i}{\partial x} - \sum_{j=1}^N \frac{\partial H_i}{\partial u_j^*} \frac{\partial u_j^*}{\partial x},$$

waarbij

$$(7.3) \quad H_i = f_{0i}(x, u_1^*, \dots, u_N^*) + \nabla V_i(x) f(x, u_1^*, \dots, u_N^*).$$

Formule (7.2) is afgeleid van het geval dat we closed-loop besturingen (zonder geheugen) bekijken. In het algemeen zal de term

$$\frac{\partial H_i}{\partial u_j^*} \cdot \frac{\partial u_j^*}{\partial x}$$

voor $j = i$ nul zijn, voor $j \neq i$ echter niet !

Ingeval we met open-loop oplossingen te maken hebben, krijgen we i.p.v.

(7.2) het volgende:

$$(7.4) \quad \frac{d}{dt} \frac{dV_i(x)}{dx} = - \frac{\partial H_i}{\partial x}.$$

De geadjungeerde vergelijkingen zijn dus verschillend en daarom de oplossingen i.h.a. ook naar gelang we open-loop dan wel closed-loop oplossingen bestuderen. Merk op dat (7.1) voor beide gevallen geldt. Ook nu kunnen singuliere oppervlakken bestaan waarvoor (7.2) of (7.4) geen zin heeft. Als een optimale baan z'n singulier oppervlak passeert, geldt aan beide kanten van dit oppervlak (7.2) of (7.4); tijdens de overgang kunnen er sprongen in ∇V_i optreden; hiervoor zijn enkele voorwaarden ontwikkeld, de zgn. *sprongcondities*. Zie daarvoor [2], [10].

8. NIET-NULSOMSPLEN; TOEPASSINGEN

Toepassing 1. "Onderhandelen gedurende een staking".

Het aanbod van de werkgevers is $x(t)$ en de eisen van de werknemers zijn $y(t)$ met $y(t_0) > x(t_0)$. Beide grootheden zijn scalair. Het model zal zijn:

$$\dot{x}(t) = u(t)[y(t) - x(t)],$$

$$\dot{y}(t) = -v(t)[y(t) - x(t)],$$

waarbij u en v de beslissingsvariabele van werkgevers resp. werknemers is. Het einde van het spel is bereikt als $x(t)$ en $y(t)$ voldoende dicht bij elkaar zijn:

$$y(t_f) - x(t_f) = m,$$

met $m > 0$ gegeven. De kosten zijn (P_1 = werkgevers, P_2 = werknemers):

$$J_1 = k_1 t_f + x(t_f),$$

$$J_2 = k_2 t_f - y(t_f),$$

waarbij $k_1, k_2 > 0$ gegeven. Door te definiëren $z = y - x$ wordt de toestandsvergelijking in de z -ruimte:

$$\dot{z}(t) = -[u(t) + v(t)]z(t); \quad z(t_0) = z_0 = y_0 - x_0;$$

doelgebied is $z = m$, en de kosten zijn (afgezien van constante termen):

$$\tilde{J}_1 = \int_{t_0}^{t_f} (k_1 + u(t)z(t)) dt$$

$$\tilde{J}_2 = \int_{t_0}^{t_f} (k_2 + v(t)z(t)) dt.$$

We gaan zoeken naar closed-loop oplossingen en moeten daarom beschouwen $u(z)$ i.p.v. $u(t)$ en $v(z)$ i.p.v. $v(t)$. Uit de noodzakelijke voorwaarden van de vorige sectie volgt direkt dat

$$u^*(z) = \frac{k_2}{z}; \quad v^*(z) = \frac{k_1}{z}$$

een kandidaat-optimale oplossing is. Deze blijkt ook aan voldoende voorwaarden (niet in dit verhaal geformuleerd) te voldoen en is daarom inderdaad een Nash-oplossing. Dit alles geldt onder de voorwaarde dat $V_1(x)$ eenduidig is, hetgeen niet bewezen is, en misschien niet waar is. Voor de gevonden oplossing geldt:

$$V_1(z) = V_2(z) = z - m.$$

Toepassing 2. De maatschappij P_i heeft ten tijde t : $x_i(t)$ klanten en heeft op dat moment een winst gemaakt van

$$g_i(t) = \int_0^t (c_i x_i(t) - u_i(t)) dt,$$

waarbij $c_i > 0$ constant. Adverteren (met kosten $u_i(t)$ per tijdseenheid) doet het aantal klanten toenemen volgens modelvergelijkingen. Als er twee maatschappijen zijn, zijn deze vergelijkingen:

$$\frac{dx_1}{dt} = u_1(t) - u_2(t), \quad x_1(0) = x_{10},$$

$$\frac{dx_2}{dt} = u_2(t) - u_1(t), \quad x_2(0) = x_{20}.$$

Hiet drie-maatschappijen model ziet er als volgt uit:

$$\frac{dx_1}{dt} = u_1 - \frac{x_1}{x_1+x_3} u_2 - \frac{x_1}{x_1+x_2} u_3, \quad x_1(0) = x_{10},$$

$$\frac{dx_2}{dt} = u_2 - \frac{x_2}{x_2+x_1} u_3 - \frac{x_2}{x_2+x_3} u_1, \quad x_2(0) = x_{20},$$

$$\frac{dx_3}{dt} = u_3 - \frac{x_3}{x_3+x_2} u_1 - \frac{x_3}{x_3+x_1} u_2, \quad x_3(0) = x_{30}.$$

Voor elk van de modellen gelden de volgende beperkingen:

$$u_i(t) \geq 0,$$

$$x_i(t) \geq 0,$$

$$g_i(t) \geq 0.$$

Maatschappij P_i kiest z'n $u_i(t)$ zodanig om de volgende functie te maximaliseren:

$$\max_{u_i} g_i(T),$$

waarbij $T > 0$ vast is.

Afhankelijk van de beginvoorwaarden en de c_i 's kan de open-loop oplossing er van het 2-maatschappijenmodel als volgt uit zien ($c_1 \geq c_2$):

$$u_i(t) = \begin{cases} \frac{c_i}{x_i(t)} & 0 \leq t \leq t_i, \\ 0 & t_i < t \leq T, \end{cases}$$

waarbij deze open-loop oplossing in de closed-loop vorm geschreven is, en waarbij:

$$t_1 = T - \frac{1}{c_1}, \quad t_2 = t_1 - \ln\left(\frac{c_1}{c_2} \cdot \frac{1+c_2}{1+c_1}\right).$$

Deze oplossing volgt rechtstreeks uit de vorige sectie.

Voor het 3-maatschappijenmodel wordt de oplossing ingewikkelder. Ten eerste zijn de modelvergelijkingen niet-lineair, zodat de berekeningen moeilijk uitvoerbaar zijn. Ten tweede kunnen *coalities* optreden; de twee zwakste maatschappijen (d.w.z. het minste aantal klanten) kunnen gaan samenwerken tegen de derde maatschappij. Kwalitatieve beschouwingen langs deze lijn worden beschreven in [12].

Toepassing 3. Dit niet-nulsomspel valt er een beetje buiten; het is statisch en stochastisch. Speler 1 kiest u ; speler 2 kiest v . De kosten zijn

$$\begin{aligned} J_1 &= (x+u+v)^2 c + u^2 d, \\ J_2 &= (x+u+v)^2 c + v^2 d, \end{aligned}$$

waarbij c en d scalaire grootheden zijn, > 0 en x is normaal verdeeld met verwachtingswaarde nul en variantie $\sigma > 0$; $x \sim N(0, \sigma)$. Beide spelers doen een waarneming; z_1 en z_2 resp.:

$$z_1 = x + w_1; \quad z_2 = x + w_2,$$

waarbij $w_i \sim N(0, s_i)$, $s_i > 0$; w_1 en w_2 zijn onafhankelijk. Bovenstaande gegevens zijn aan beide spelers bekend, behalve dat z_1 niet aan P_2 bekend is en z_2 niet aan P_1 . De informatiestructuur is: u is een functie van z_1 ; v is een functie van z_2 . We zoeken de Nash-oplossing. Het enige dat niet-

symmetrisch in dit spel voorkomt is dat mogelijkerwijs $s_1 \neq s_2$. We verwachten dat de speler met de kleinere s ook de kleinste waarde, J^* , zal hebben. Uitrekenen geeft echter het omgekeerde! De speler met de beste informatie is het slechtste af. Dit is duidelijk een niet-verwachte en niet-gewenste eigenschap van Nash-oplossingen. Voor de echte berekeningen, zie [13].

Toepassing 1 staat in 2, toepassing 2 staat in [12] en toepassing 3 in [13] en [14].

8. EPILOOG

Binnen de TDS bestaan verschillende stromingen, waarvan er $2\frac{1}{2}$ de aandacht hebben gekregen in dit verhaal, n.l. de singuliere oppervlakken bij nulsomspelen, de informatiestrukturen bij niet-nulsomspelen en verraderlijke eigenschappen van de Nash-oplossing bij stochastische spelen. Andere stromingen zijn het geven van existentie-stellingen voor zadelpunten, decentralisatie van kennis (elke speler kent maar een deel van het totale systeem), systeemtheoretische eigenschappen in de TDS. In de proceedings welke genoemd staan onder [5] en [9] zijn artikelen in al deze richtingen te vinden.

Getracht is in dit verhaal de sterke en zwakke punten, de mogelijkheden en de uitdagingen van de Theorie der Differentiaalspelen aan te tonen.

REFERENTIES

- [1] STARR, A.W. & Y.C. HO, *Nonzero-sum differential games and Further properties of nonzero-sum differential games*, J. of Opt. Theory and Appl., vol. 3, nos 3/4, pp. 184-219, 1969.
- [2] LEITMANN, G., *Cooperative and non-cooperative many players differential games*, Intern. Centre for Mechanical Sciences, courses and lectures nr. 190, Springer, 1974.

- [3] ISAACS, R., *Differential games*, Wiley, 2nd ed., 1975.
- [4] OWEN, G., *Game theory*, Saunders, 1968.
- [5] OLSDER, G.J., *Information structures in differential games*, in:
Differential games and control theory II, ROXIN, LIN, STERNBERG
(eds.), MARCEL DEKKER, 1977, pp. 99-135.
- [6] BAŞAR, T., *On the uniqueness of the Nash-solution in linear quadratic differential games*, Intern. J. of Game Theory, vol. 4, 1975.
- [7] BERNHARD, P., *Singular surfaces in differential games, an introduction*,
in [9].
- [8] BREAKWELL, J.V., *Zero-sum differential games with terminal payoff*,
in [9].
- [9] *Proceedings of the Workshop on differential games and applications*,
held at Twente Univ. of Technology, March 1977. Springer Lecture
Note in Control and Information Science no. 3, 1977, P. HAGEDORN,
H.W. KNOBLOCH, G.J. OLSDER (eds.).
- [10] BLAQUIÈRE, A., *An introduction to differential games*, in: *Differential games and related topics*, H.W. KUHN, G.P. SREGÖ (eds.), North
Holland, 1971.
- [11] OLSDER, G.J. & J. WALTER, *A differential approach to collision avoidance of ships*, 8th IFIP Conference on Optimisation
Techniques, Sept. 5-10, 1977, Würzburg. Proceedings will be
published by Springer.
- [12] OLSDER, G.J., *Some thoughts about simple advertising models as differential games and the structure of coalitions*, in:
Directions in large-scale systems, Y.C. HO, K. MITTER (eds.),
Plenum Press, 1976.
- [13] BAŞAR, T., *A contradictory property of the Nash solution for two stochastic nonzero-sum games*, Proceedings 10th Allerton Conference
on Circuit and System Theory, Univ. of Illinois Press, 1972,
pp. 819-827.
- [14] BAŞAR, T. & Y.C. HO, *Informational properties of the Nash solution of two stochastic nonzero-sum games*, J. of Economic Theory, vol. 7,
no. 4, 1974, pp. 370-387.