BA

stichting
mathematisch
centrum

MC

AFDELING MATHEMATISCHE BESLISKUNDE BC 12/74 DECEMBER

H.C. TIJMS

SEMI-MARKOV BESLISSINGSTHEORIE MET TOEPASSINGEN OP WACHTTIJDPROBLEMEN MET BESTURING

2e boerhaavestraat 49 amsterdam

INHOUD

1. HET SEMI-MARKOV BESLISSINGSMODEL MET EEN AFTELBARE TOESTANDSRUIMTE EN ONBEGRENSDE KOSTEN

1.1. *Inleiding*

In dit hoofdstuk beschouwen wij het semi-Markov beslissingsmodel met een aftelbare toestandsruimte, een eindige aktieruimte en kosten die niet uniform begrensd behoeven te zijn. Het optimaliteitskriterium waarin wij primair geïnteresseerd zijn is het gemiddelde kosten kriterium. Voorwaarden zullen worden gegeven waaronder niet alleen een optimale politiek bestaat, maar tevens een optimale politiek bestaat die tot de klasse van de stationaire politieken behoort. Ten overvloede wellicht merken wij op dat zelfs voor het Markov beslissingsmodel met uniform begrensde kosten het kan zijn dat een gemiddeld kosten optimale politiek niet bestaat of dat een gemiddeld kosten optimale politiek bestaat, maar geen enkele gemiddeld kosten optimale politiek tot de klasse van de stationaire politieken behoort. Voor voorbeelden van deze verschijnselen verwijzen wij naar ROSS (1970).

De behandeling in dit hoofdstuk is gebaseerd op werk van ROSS (1970) waar de klassieke theorie met uniform begrensde kosten wordt gegeven en werk van LIPPMAN (1973a,1973b) waar onbegrensde kosten worden toegelaten. Voor het semi-Markov beslissingsmodel met onbegrensde kosten zijn ook in HARRISON (1970,1972a) en in HORDIJK (1973) voorwaarden gegeven waaronder een optimale politiek bestaat. De voorwaarden en de aanpak van LIPPMAN (1973a,1973b) lijken het meest toepasbaar om het bestaan van een optimale politiek aan te tonen en om tevens de structuur van een dergelijke politiek te karakteriseren. Bovendien zijn deze voorwaarden uitermate geschikt voor toepassingen in wachttijdtheorie. De voorwaarden die wij zullen geven zijn overigens minder algemeen dan de voorwaarden in LIPPMAN (1973a,1973b). De reden hiervan is dat het bewijs van stelling 2 in LIPPMAN (1973a) incorrect is en deze stelling speelt een basisrol in het werk van LIPPMAN (1973a,1973b).

De te volgen aanpak van het gemiddeld kosten kriterium is standaard. Dit kriterium wordt geanalyseerd door eerst resultaten af te leiden voor het aanzienlijk eenvoudiger te analyseren kriterium van de totale verwachte verdisconteerde kosten. Hieruit worden vervolgens door toepassing van Abel stellingen resultaten voor het gemiddeld kosten kriterium afgeleid.

1.2. *Het model en notatie.*

Beschouw een systeem dat op stochastisch gelegen tijdstippen wordt waargenomen in één van de toestanden van een aftelbare toestandsruimte S. Wij beschouwen het geval dat de toestand s geschreven kan worden als $s = (i,\xi)$ waarbij de eerste component i een niet-negatief geheel getal is en de component ξ een willekeurig object mag zijn dat behoort tot een (mogelijk lege) verzameling S_i. De component i moet worden opgevat als de hoofdvariabele terwijl de component ξ gezien moet worden als een hulpvariabele die op zichzelf tamelijk complex mag zijn. Bijvoorbeeld, in een wachttijdprobleem kan i het aantal aanwezige klanten zijn en kan ξ aangeven of de bediende bezig is of niet. Als de toestand van het systeem wordt waargenomen, dan moet een aktie gekozen worden. Voor elke toestand $s \in S$ is een eindige verzameling A_s van toegelaten akties gegeven. De toestand van het systeem wordt voor het eerst waargenomen op tijdstip 0. Als het systeem wordt waargenomen in toestand s en de aktie $a \in A_s$ wordt genomen, dan gebeurt, onafhankelijk van het verleden van het systeem, het volgende:

(1) de volgende toestand van het systeem is toestand s' met kans $q(s'|s,a)$,

(2) gegeven dat de volgende toestand s' is, de hoeveelheid tijd tot de overgang van toestand s naar toestand s' plaatsvindt is een stochastische variabele met verdelingsfunktie $F(t|s,a,s')$.

Na de overgang naar de volgende toestand moet opnieuw een aktie gekozen worden waarna het bovenstaande zich herhaalt, etc.

De overgangskansen $q(s'|s,a)$ en de verdelingsfunkties $F(t|s,a,s')$ worden bekend verondersteld. Om ervoor te zorgen dat met kans 1 het aantal toestandsovergangen in een eindig tijdsinterval eindig is, maken wij de volgende aanname:

Aanname 1. *Er bestaan positieve getallen* ε *en* δ *zo dat*

$$\sum_{s' \in S} q(s'|s,a)\{1-F(\delta|s,a,s')\} \geq \varepsilon \quad \textit{voor alle } s \in S \textit{ en } a \in A_s.$$

D.w.z., voor elke toestand s en aktie $a \in A_s$ geldt dat de tijd tot de eerstvolgende toestandsovergang groter dan δ is met een kans die tenminste ε is. Voorts nemen wij aan dat voor elke $s \in S$ en $a \in A_s$ de grootheid

$$(1.1) \qquad \tau(s,a) = \sum_{s' \in S} q(s'|s,a)\, \tau(s,a,s')$$

eindig is, waarbij $\tau(s,a,s') = \int_0^\infty t dF(t|s,a,s')$. Merk op dat $\tau(s,a)$ de verwachte tijd is tot de eerstvolgende toestandsovergang als in toestand s de aktie a genomen wordt.

Wij beschouwen de volgende kostenstructuur. Als in toestand s de aktie a wordt genomen, dan worden direkte kosten $d_1(s,a)$ gemaakt en worden voor elke tijdseenheid tot de eerstvolgende toestandsovergang kosten $d_2(s,a)$ gemaakt. Wij nemen aan dat

$$(1.2) \qquad d_1(s,a) \geq 0 \quad \text{en} \quad d_2(s,a) \geq 0 \qquad \text{voor alle } s \in S \text{ en } a \in A_s.$$

Definieer voor alle $s \in S$ en $a \in A_s$,

$$(1.3) \qquad c(s,a) = d_1(s,a) + \tau(s,a) d_2(s,a),$$

dan is $c(s,a)$ de verwachte kosten tot de eerstvolgende toestandsovergang als in toestand s de aktie a genomen wordt.

Een politiek R voor de besturing van het systeem is een regel die op elk beslissingstijdstip voorschrijft hoe de te nemen aktie gekozen moet worden. Deze keuze mag afhangen van de tot het beslissingstijdstip waargenomen toestanden en gekozen akties. Voorts mag het zo zijn dat tussen de mogelijke akties geloot wordt welke aktie gekozen wordt. De klasse van alle mogelijke politieken geven wij aan door C. Een belangrijk deelklasse van de klasse C is de klasse F van de stationaire politieken. Een stationaire politiek f is een voorschrift dat aan elke toestand $s \in S$ één en slechts één aktie $f(s) \in A_s$ toevoegt, zodat altijd aktie f(s) wordt genomen wanneer het systeem wordt waargenomen in toestand s.

Voor gegeven begintoestand s en politiek $R \in C$, definieer $\underline{x}_n$ en $\underline{T}_n$ als de toestand na de n^{de} toestandsovergang en als het tijdstip van de n^{de} toestandsovergang voor $n=0,1,\ldots$, waarbij wij afspreken dat de 0^{de} toestandsovergang plaatsvindt op tijdstip 0 (dus $\underline{x}_0 = s$ en $\underline{T}_0 = 0$). Voorts definieer $\underline{a}_n$ als de aktie die genomen zal worden op tijdstip $\underline{T}_n$, $n \geq 0$. Gegeven dat de begintoestand $s \in S$ is en politiek $R \in C$ wordt gevolgd, definieer verder

(1.4) $V(s,R,t)$ = verwachting van de totale kosten in $[0,t]$, $t \geq 0$

en

(1.5) $$g(s,R) = \limsup_{t\to\infty} t^{-1}V(s,R,t).$$

Wij merken op dat $V(s,R,t)$ en $g(s,R)$ welgedefinieerd zijn op grond van het feit dat de kosten niet-negatief zijn. De grootheid $g(s,R)$ kan geïnterpreteerd worden als de gemiddelde verwachte kosten per tijdseenheid op de lange duur als de begintoestand s is en politiek R wordt gevolgd. Wij zullen een politiek $R^* \in C$ *gemiddeld kosten optimaal* noemen als $g(s,R^*) \leq g(s,R)$ voor alle $s \in S$ en $R \in C$.

Zoals reeds gesteld zullen wij het gemiddeld kosten model analyseren door eerst het verdisconteerde kostenmodel te bestuderen. Laat α een willekeurig positief getal zijn. Voor elke $s \in S$ en $R \in C$, definiëren wij

(1.6) $$V_\alpha(s,R) = \int_0^\infty e^{-\alpha t} dV(s,R,t).$$

De grootheid $V_\alpha(s,R)$ kan geïnterpreteerd worden als de totale verwachte verdisconteerde kosten over een oneindig lange tijdsperiode als de kosten verdisconteerd worden door een faktor α, de begintoestand s is en politiek R wordt toegepast. Het verdisconteren van de kosten met een faktor α houdt in dat 1 gulden ontvangen op tijdstip t de waarde van $e^{-\alpha t}$ gulden op tijdstip 0 heeft.

Het is vrij eenvoudig na te gaan dat voor alle $s \in S$ en $R \in C$ geldt

(1.7) $$V_\alpha(s,R) = E_{s,R}\Big[\sum_{n=0}^\infty e^{-\alpha \underline{T}_n}\{d_1(\underline{x}_n,\underline{a}_n)+d_2(\underline{x}_n,\underline{a}_n)\int_{\underline{T}_n}^{\underline{T}_{n+1}} e^{-\alpha t}dt\}\Big],$$

waarbij $E_{s,R}$ de verwachting aanduidt voor begintoestand s en politiek R. Deze relatie zullen wij gebruiken bij de analyse van het verdisconteerde kosten model. Definieer $V_\alpha(s) = \inf_{R\in C} V_\alpha(s,R)$ voor $s \in S$. Voor vaste α heet een politiek R^* *α-optimaal* als $V_\alpha(s,R^*) = V_\alpha(s)$ voor alle $s \in S$.

Vervolgens definiëren wij voor alle $s \in S$ en $a \in A_s$

$$c_\alpha(s,a) = d_1(s,a) + d_2(s,a) \sum_{s'} q(s'|s,a) \int_0^\infty dF(t|s,a,s') \int_0^t e^{-\alpha y}dy,$$

m.a.w. $c_\alpha(s,a)$ is de verwachte verdisconteerde kosten tot de eerstvolgende toestandsovergang bij verdisconteringsfaktor α als in toestand s de aktie a genomen wordt. Wij merken op dat voor alle s en a geldt

(1.8) $$c_\alpha(s,a) \le c(s,a) \text{ voor } \alpha > 0, \quad \text{en } \lim_{\alpha\to 0} c_\alpha(s,a) = c(s,a).$$

Tenslotte definiëren wij voor alle s, a en s',

(1.9) $$\beta_\alpha(s,a,s') = \int_0^\infty e^{-\alpha t} dF(t|s,a,s').$$

Met behulp van aanname 1 leidt men eenvoudig af dat voor alle $\alpha > 0$,

(1.10) $$\beta_\alpha(s,a,s') \le \beta_\alpha < 1 \qquad \text{voor alle } s,s' \in S \text{ en } a \in A_s,$$

waarbij $\beta_\alpha = 1-\varepsilon(1-e^{-\alpha\delta})$. Merk op dat $\beta_\alpha(s,a,s')$ geïnterpreteerd kan worden als de waarde op tijdstip 0 van een bedrag van 1 gulden, ontvangen op tijdstip $\underline{T}_1$ als op tijdstip 0 de aktie a in toestand s genomen wordt en op tijdstip $\underline{T}_1$ een overgang naar toestand s' plaatsvindt.

1.3. *Het verdisconteerde kosten model.*

In deze paragraaf zullen wij aantonen dat onder zekere voorwaarden voor elke $\alpha > 0$ een α-optimale politiek bestaat. Wij maken de volgende aannames.

Aanname 2. *Er is een eindig getal* c *en een niet-negatief getal* K *zo dat*

$$c(s,a) \le c(i \vee 1)^K \qquad \textit{voor alle } s=(i,\xi) \in S \textit{ en } a \in A_s,$$

waarbij

$$x \vee y = \max(x,y).$$

Aanname 3. *Er is een eindig getal* $d \ge 1$ *zo dat voor alle* $k=0,\ldots,K$ *geldt*

$$\sum_{s'=(i',\xi')} (i' \vee 1)^k \, q(s'|s,a) \le (i+d)^k$$

voor alle $s=(i,\xi) \in S$ en $a \in A_s$.

Merk op dat aanname 1 eist dat de kosten $c((i,\xi),a)$ begrensd zijn door een polynoom van de orde K en dat aanname 2 stelt dat voor $k=0,\ldots,K$ het k^{de} moment van de eerste component van de eerstvolgende toestand na toestand (i,ξ) begrensd is door $(i+d)^k$ voor een eindige d.

Kies nu $\alpha > 0$ vast. Wij zullen nu aantonen dat een α-optimale politiek bestaat die tot de klasse van de stationaire politieken behoort en dat een α-optimale politiek bepaald kan worden met een behulp van een funktionaal-vergelijking waaraan $V_\alpha(s)$ voldoet. Het bewijs is gebaseerd op de theorie van de contractie afbeeldingen. Daartoe moeten wij een verzameling van funkties definiëren waartoe de funktie $V_\alpha(s,R)$, $s \in S$ behoort voor elke R en op deze verzameling moeten wij een metriek definiëren waardoor deze verzameling een compleet metrische ruimte wordt.

Laat B de volgende verzameling zijn van reëlwaardige funkties u gedefinieerd op S,

$$(1.11) \qquad B = \{u \mid \sup_{s=(i,\xi)\in S} |u(s)| \,/\, (i \vee 1)^K < \infty\}.$$

LEMMA 1.1. *Veronderstel dat de aannames* 1 – 3 *vervuld zijn. Voor elke politiek* $R \in C$ *geldt dat de funktie* $V_\alpha(s,R) \in B$. *Bovendien, de funktie* $V_\alpha(s) \in B$.

Bewijs. Definieer de stochastische variabelen $\tilde{\underline{\tau}}_1, \tilde{\underline{\tau}}_2, \ldots$ op dezelfde wijze als in het bewijs van Lemma 7.4 in ROSS (1970). Beschouw vervolgens de uitdrukking die gelijk is aan het rechterlid van (1.7) met $e^{-\alpha \underline{T}_n}$ vervangen door $e^{-\alpha(\tilde{\underline{\tau}}_1+\ldots\tilde{\underline{\tau}}_n)}$. Op deze wijze verkrijgt men een bovenschatting voor $V_\alpha(s,R)$ en van de zo verkregen uitdrukking gaat men eenvoudig na dat deze tot B behoort. Uit $0 \le V_\alpha(s) \le V_\alpha(s,R)$ voor alle s en R volgt tenslotte dat de funktie $V_\alpha(s) \in B$. □

LEMMA 1.2. *Veronderstel dat aanname* 1 *vervuld is. De funktie* $V_\alpha(s)$ *voldoet aan*

$$(1.12) \qquad V_\alpha(s) = \min_{a\in A_s} \{c_\alpha(s,a) + \textstyle\sum_{s'} q(s'|s,a)\beta_\alpha(s,a,s')V_\alpha(s')\}, \qquad s \in S.$$

Bewijs. Het bewijs is vrijwel identiek aan het bewijs van stelling 6.1 in ROSS (1970). Kies $s \in S$ en $R \in C$. Laat $\pi(a|s)$ de kans zijn dat op tijdstip 0 politiek R de aktie a kiest in toestand s. Gegeven dat op tijdstip 0 de aktie a genomen wordt in toestand s en dat de toestand op het eerstvolgende beslissingstijdstip $\underline{T}_1$ gelijk aan s' is, laat $W_\alpha(s,a,s',R)$ de totale ver-verwachte verdisconteerde kosten vanaf tijdstip $\underline{T}_1$ zijn waarbij deze kosten worden verdisconteerd naar tijdstip $\underline{T}_1$. Gebruikmakend van (1.7) gaat men eenvoudig na dat

$$V_\alpha(s,R) = \sum_a \pi(a|s)\{c_\alpha(a,s) + \sum_{s'} q(s'|s,a)\beta_\alpha(s,a,s')W_\alpha(s,a,s',R\} \geq$$

$$\geq \sum_a \pi(a|s)\{c_\alpha(a,s) + \sum_{s'} q(s'|s,a)\beta_\alpha(s,a,s')V_\alpha(s')\} \geq$$

$$\geq \min_{a \in A_s} \{c_\alpha(a,s) + \sum_{s'} q(s'|s,a)\beta_\alpha(s,a,s')V_\alpha(s')\}.$$

Aangezien R willekeurig gekozen was, volgt nu dat $V_\alpha(s)$ groter dan of gelijk aan het rechterlid van (1.12) is. Om de omgekeerde ongelijkheid te bewijzen, kies s vast en laat a_0 een aktie zijn waarvoor het rechterlid van (1.12) minimaal is. Kies $\varepsilon > 0$, en voor elke $s' \in S$ laat politiek $R_{s'} \in C$ zō zijn dat $V_\alpha(s',R_{s'}) \leq V_\alpha(s')+\varepsilon$. Definieer politiek $\pi \in C$ als volgt. Op tijdstip 0 neemt politiek π de beslissing a_0 in toestand s, en, als op het volgende beslissingstijdstip $\underline{T}_1$ de toestand s' is, dan doet politiek π alsof op tijdstip $\underline{T}_1$ het systeem opnieuw begint met s' als begintoestand en vanaf dat moment volgt π de politiek $R_{s'}$. Derhalve vinden wij met behulp van (1.10)

$$V_\alpha(s) \leq V_\alpha(s,\pi) = c_\alpha(s,a_0) + \sum_{s'} q(s'|s,a_0)\beta_\alpha(s,a_0,s')V_\alpha(s',R_{s'}) \leq$$

$$\leq c_\alpha(s,a_0) + \sum_{s'} q(s'|s,a_0)\beta_\alpha(s,a_0,s')\{V_\alpha(s')+\varepsilon\} \leq$$

$$\leq c_\alpha(s,a_0) + \sum_{s'} q(s'|s,a)\beta_\alpha(s,a_0,s')V_\alpha(s')+\varepsilon\beta_\alpha.$$

Aangezien ε willekeurig klein gekozen mag worden, volgt nu dat $V_\alpha(s)$ kleiner dan of gelijk aan het rechterlid van (1.12) is. Hiermee is het bewijs compleet. □

Definieer $V_{0,\alpha}(s) = 0$ voor alle $s \in S$, en voor $n=0,1,\ldots$, definieer

$$(1.13)\qquad V_{n+1,\alpha}(s) = \min_{a\in A_s} \{c_\alpha(s,a) + \textstyle\sum_{s'} q(s'|s,a)\beta_\alpha(s,a,s')V_{n,\alpha}(s')\},\ s \in S.$$

Opmerking 1.1. Voor elke vaste $N \geq 1$ geldt dat $V_{N,\alpha}(s_0)$ gelijk is aan de minimale totale verwachte verdisconteerde kosten tot het $(N+1)^{\text{ste}}$ beslissingstijdstip als de begintoestand s_0 is. Deze minimale kosten worden bereikt door voor $n=0,\ldots,N-1$ op het $(n+1)^{\text{ste}}$ beslissingstijdstip $\underline{T}_n$ de aktie $f_{n,\alpha}(s)$ te nemen als op dit tijdstip het systeem in toestand s is, waarbij $f_{n,\alpha}(s)$ een aktie is waarvoor het rechterlid van (1.13) minimaal is. De bewijzen van deze "vanzelfsprekende" beweringen zijn standaard en geheel analoog aan de bewijzen in hoofdstuk 2 in DERMAN (1970).

STELLING 1.1. *Veronderstel dat de aannames* 1 - 3 *vervuld zijn.*
(a) *De funktie* $V_\alpha(s)$, $s \in S$, *is de enige funktie uit de klasse* B *die voldoet aan de "optimaliteitsvergelijking",*

$$(1.14)\qquad V_\alpha(s) = \min_{a\in A_s} \{c_\alpha(s,a) + \textstyle\sum_{s'} q(s'|s,a)\beta_\alpha(s,a,s')V_\alpha(s')\},\qquad s \in S.$$

Verder geldt

$$(1.15)\qquad \lim_{n\to\infty} V_{n,\alpha}(s) = V_\alpha(s)\qquad \text{voor alle } s \in S.$$

(b) *Laat* $f_\alpha \in F$ *zo zijn dat* $f_\alpha(s)$ *het rechterlid van* (1.14) *minimaliseert voor alle* $s \in S$. *Dan is de stationaire politiek* f_α *α-optimaal.*
(c) *Voor* $n=0,1,\ldots$, *definieer* $f_{n,\alpha} \in F$ *als een politiek zodat* $f_{n,\alpha}(s)$ *het rechterlid van* (1.13) *minimaliseert voor alle* $s \in S$. *Veronderstel dat een politiek* $g_\alpha \in F$ *bestaat zodat voor elke* $s \in S$ *een deelrij* $\{n_k\}$ *met* $n_k \to \infty$ *als* $k \to \infty$ *bestaat zodat* $f_{n_k,\alpha}(s) = g_\alpha(s)$. *Dan is politiek* g_α *α-optimaal.*

Bewijs. (a) Voeg aan elk tweetal funkties $u,v \in B$ toe het getal

$$(1.16)\qquad \rho(u,v) = \sup_{s=(i,\xi)\in S} \{|u(s)-v(s)| \,/\, (i\vee 1)^K\}.$$

Gebruikmakend van het Cauchy kriterium voor een rij van uniform convergente funkties, gaat men eenvoudig na dat (B,ρ) een compleet metrische ruimte is, vgl. de appendix. Voor elke funktie $u \in B$ definieer de funktie Tu door

$$(1.17) \qquad Tu(s) = \min_{a \in A_s} \{c_\alpha(a,s) + \textstyle\sum_{s'} q(s'|s,a)\beta_\alpha(s,a,s')u(s')\}, \qquad s \in S.$$

Om aan te tonen dat de afbeelding T een uniek vast punt bezit, moeten wij eerst nagaan dat $Tu \in B$ voor alle $u \in B$. Kies $u \in B$ vast, en voor alle $s \in S$ laat $f(s)$ een aktie zijn waarvoor het rechterlid van (1.17) minimaal is. Laat M een bovengrens zijn voor $|u(s)| \,/\, (i \vee 1)^K$ voor $s=(i,\xi) \in S$, vgl. (1.11). Op grond van relatie (1.10) en de aannames 2 en 3, vinden wij voor alle $s=(i,\xi) \in S$,

$$|Tu(s)| = |c_\alpha(s,f(s)) + \textstyle\sum_{s'} q(s'|s,f(s))\beta_\alpha(s,f(s),s')u(s')| \le$$

$$\le c(i \vee 1)^K + \beta_\alpha M \textstyle\sum_{s'=(i',\xi')} q(s'|s,f(s))(i' \vee 1)^K \le$$

$$\le c(i \vee 1)^K + \beta_\alpha M(i+d)^K \le \{c+\beta_\alpha M(1+d)^K\}(i \vee 1)^K.$$

Dus $Tu \in B$. Het is eenvoudig in te zien dat de afbeelding T continu is. Definieer de funktie $T^n u$ door $T^n u = T(T^{n-1}u)$ voor $n \ge 2$ waarbij $T^1 u = Tu$. Dan geldt $V_{n,\alpha} = T^n V_{0,\alpha}$. Vanwege de lemma's 1.1 en 1.2 volgt de bewering (a) uit stelling 1 in de appendix wanneer wij kunnen aantonen dat een getal β met $0 \le \beta < 1$ en een natuurlijk getal N bestaan zodat

$$(1.18) \qquad \rho(T^N u, T^N v) \le \beta\rho(u,v) \qquad \text{voor alle } u,v \in B.$$

Om (1.18) te bewijzen, zullen wij aantonen dat voor alle $n=1,2,\ldots$ geldt

$$(1.19) \qquad |T^n u(s) - T^n v(s)| \le \beta_\alpha^n \rho(u,v)(i+nd)^K \text{ voor alle } u,v \in B \text{ en } s=(i,\xi) \in S.$$

Uit (1.16) en (1.19) volgt $\rho(T^n u, T^n v) \le \beta_\alpha^n \rho(u,v)(1+nd)^K$ voor alle $n \ge 1$, waaruit (1.18) volgt aangezien $0 \le \beta_\alpha < 1$. Voor het bewijs van (1.19) merken wij eerst op dat uit (1.17) direkt volgt

$$(1.20) \qquad Tv_1(s) \le Tv_2(s) \text{ voor alle } s \in S \text{ als } v_1(s) \le v_2(s) \text{ voor alle } s \in S.$$

Uit (1.16) volgt $u(s) \le v(s) + (i\vee 1)^K \rho(u,v)$ voor alle $s=(i,\xi) \in S$ en alle u,v. Derhalve volgt uit (1.20), (1.17), (1.10) en aanname 3 dat voor alle $s=(i,\xi) \in S$,

$$Tu(s) \le \min_\alpha\{c_\alpha(s,a) + \sum_{s'=(i',\xi')} q(s'|s,a)\beta_\alpha(s,a,s')[v(s') + $$
$$+ (i'\vee 1)^K \rho(u,v)]\} \le$$
$$\le \min_a\{c_\alpha(s,a) + \sum_{s'} q(s'|s,a)\beta_\alpha(s,a,s')v(s')\} +$$
$$+ \beta_\alpha \rho(u,v)(i+d)^K.$$

Dus $Tu(s) \le Tv(s) + \beta_\alpha \rho(u,v)(i+d)^K$. Om redenen van symmetrie geldt dan ook $Tv(s) \le Tu(s) + \beta_\alpha \rho(v,u)(i+d)^K$. Hieruit volgt dan (1.19) voor $n = 1$. Stel nu dat (1.19) bewezen is voor $n=1,\dots,m$. Dan geldt

$$(1.21) \qquad T^m u(s) \le T^m v(s) + \beta_\alpha^m \rho(u,v)(i+md)^K \qquad \text{voor alle } s=(i,\xi) \in S.$$

Uit (1.21), (1.20), (1.17), (1.10) en aanname 3 volgt nu dat voor alle $s=(i,\xi)$,

$$T^{m+1}u(s) \le \min_a\{c_\alpha(s,a) + \sum_{s'} q(s'|s,a)\beta_\alpha(s,a,s')[T^m v(s') +$$
$$+ \beta_\alpha^m \rho(u,v)(i'+md)^K]\} \le$$
$$\le \min_a\{c_\alpha(s,a) + \sum_{s'} q(s'|s,a)\beta_\alpha(s,a,s')T^m v(s') +$$
$$+ \beta_\alpha^{m+1} \rho(u,v) \sum_{j=0}^{K} \binom{K}{j} (md)^{K-j} \sum_{s'} q(s'|s,a)(i')^j\} \le$$
$$\le \min_a\{c_\alpha(s,a) + \sum_{s'} q(s'|s,a)\beta_\alpha(s,a,s')T^m v(s')\} +$$
$$+ \beta_\alpha^{m+1} \rho(u,v) \sum_{j=0}^{K} \binom{K}{j} (md)^{K-j} (i+d)^j.$$

Dus $T^{m+1}u(s) \le T^{m+1}v(s) + \beta_\alpha^{m+1}\rho(u,v)\{i+(m+1)d\}^K$. Door de rollen van u en v te verwisselen, volgt nu (1.19) voor $n=m+1$. Hiermee is het bewijs van bewering (a) voltooid.

(b) Kies $f \in F$ willekeurig. Uit (1.7) volgt direkt dat $V_\alpha(s,f)$ voldoet aan

$$(1.22)\qquad V_\alpha(s,f) = c_\alpha(s,f(s)) + \sum_{s'} q(s'|s,f(s))\beta_\alpha(s,f(s),s')V_\alpha(s',f),\ s \in S.$$

Door het bewijs van stelling 1.1(a) letterlijk te herhalen met A_s vervangen door $A'_s = \{f(s)\}$ voor alle $s \in S$, vinden wij dat de funktie $V_\alpha(s,f)$ de enige funktie uit de klasse B is die aan (1.22) voldoet. Aangezien

$$V_\alpha(s) = c_\alpha(s,f_\alpha(s)) + \sum_{s'} q(s'|s,f_\alpha(s))\beta_\alpha(s,f_\alpha(s),s')V_\alpha(s'),\quad s \in S$$

volgt dat $V_\alpha(s,f_\alpha) = V_\alpha(s)$ voor alle $s \in S$, m.a.w. f_α is α-optimaal.

(c) Met induktie volgt uit (1.13) dat voor elke $n \geq 0$ geldt $V_{n+1,\alpha}(s) \geq$ $\geq V_{n,\alpha}(s) \geq 0$ voor alle $s \in S$. Derhalve geldt voor alle s en a,

$$\lim_{n\to\infty} \sum_{s'} q(s'|s,a)\beta_\alpha(s,a,s')V_{n,\alpha}(s') = \sum_{s'} q(s'|s,a)\beta_\alpha(s,a,s')V_\alpha(s').$$

Kies s vast. Laat $\{n_k\}$ een rij zijn met $n_k \to \infty$ als $k \to \infty$ zodat $f_{n_k,\alpha}(s) =$ $= g_\alpha(s)$ voor alle k. Door k naar ∞ te laten gaan in (1.13) met $n = n_k$, leidt men nu eenvoudig af dat $g_\alpha(s)$ het rechterlid van (1.14) minimaliseert. Bewering b) volgt nu uit bewering c). □

Opmerking 1.2. Veronderstel dat de aannames 1 - 4 vervuld zijn. Dan geldt voor elke α-optimale stationaire politiek f dat f(s) het rechterlid van (1.14) minimaliseert voor alle $s \in S$. Het bewijs van deze bewering is eenvoudig. Stel dat er een toestand s_0 is zodat $f(s_0)$ het rechterlid van (1.14) niet minimaliseert. Dan geldt $V_\alpha(s,f) \leq c_\alpha(s,f(s)) +$ $+ \sum q(s'|s,f(s))\beta_\alpha(s,f(s),s')V_\alpha(s',f)$ voor alle $s \in S$, waarbij het < teken geldt voor $s = s_0$. Door deze ongelijkheid voor $s = s_0$ te itereren, vindt men $V_\alpha(s_0)) < V_\alpha(s_0,f)$. Deze tegenspraak bewijst bovenstaande bewering.

1.4. *Het bestaan van een gemiddeld kosten optimale politiek.*

In deze paragraaf zullen wij aantonen dat onder zekere voorwaarden een gemiddeld kosten optimale politiek bestaat. Tevens zullen wij de "optimaliteitsvergelijking" voor het gemiddelde kosten kriterium afleiden.

Aanname 4. (a) *Er is een politiek* $f^* \in F$ *en een rij* $\{\alpha_n, n \geq 1\}$ *met* $\alpha_n \to 0$ *als* $n \to \infty$ *zodat voor elke* $n \geq 1$ *de stationaire politiek* f^* α_n*-optimaal is.*

(b) *Er is een toestand* $s^* \in S$ *zodat*

$$E_{s,f^*}(\underline{\tau}) < \infty \quad \textit{en} \quad E_{s,f^*}(\underline{z}(\underline{\tau})) < \infty \qquad \textit{voor alle } s \in S,$$

waarbij $\underline{\tau}$ *het eerste tijdstip vallend na tijdstip* 0 *is waarop een overgang naar toestand* s^* *plaatsvindt en* $\underline{z}(t)$ *de totale kosten in* $[0,t)$ *is,* $t \geq 0$.*)

Op grond van een standaard resultaat uit de theorie van de regeneratieve processen geldt dat aanname 4(b) impliceert (zie bijvoorbeeld stelling 3.16 in ROSS (1970)),

$$\lim_{t\to\infty} t^{-1} E_{s,f^*}(\underline{z}(t)) = E_{s^*,f^*}(\underline{z}(\underline{\tau}))/E_{s^*,f^*}(\underline{\tau}) \qquad \text{voor alle } s \in S.$$

Derhalve geldt (zie (1.4))

$$(1.23) \qquad g(s,f^*) = \lim_{t\to\infty} t^{-1} V(s,f^*,t) \text{ voor alle } s, \text{ waarbij } g(s,f^*) \text{ eindig is en onafhankelijk van } s \in S.$$

Aanname 5. *Voor alle* $s \in S$ *en* $a \in A_s$ *geldt*

$$\textstyle\sum_{s'} q(s'|s,a) E_{s',f^*}(\underline{\tau}) < \infty \quad \text{en} \quad \sum_{s'} q(s'|s,a) E_{s',f^*}(\underline{z}(\underline{\tau})) < \infty.$$

In vele toepassingen leidt de analyse van het verdisconteerde model tot het resultaat dat voor alle α voldoende klein een α-optimale politiek bestaat die behoort tot een eindige klasse van stationaire politieken met een bepaalde struktuur. De volgende stelling leert dan dat ook een gemiddeld kosten optimale stationaire politiek met deze struktuur bestaat.

STELLING 1.2. *Veronderstel dat de aannames* 1 - 4 *vervuld zijn.*

---

*) De resultaten uit deze paragraaf blijven geldig als aanname 4(a) verzwakt wordt tot: er is een politiek $f^* \in F$, een eindig getal M en een rij $\{\alpha_n\}$ met $\alpha_n \to 0$ als $n \to \infty$ zodat $V_{\alpha_n}(s,f^*) \leq V_{\alpha_n}(s)+M$ voor alle $s \in S$ en $n \geq 1$. De bewijzen behoeven slechts ondergeschikte veranderingen.

(a) *De politiek* f^* *uit aanname* 4 *is gemiddeld kosten optimaal. De minimale gemiddelde kosten* $g^*(s) = \min_{R\in C} g(s,R)$ *zijn onafhankelijk van de begintoestand* $s \in S$, *m.a.w.* $g^*(s) \equiv g^*$ *voor een constante* g^*.

(b) $\lim_{n\to\infty} \alpha_n V_{\alpha_n}(s) = g^*$ *voor alle* $s \in S$.

(c) *Er is een eindige funktie* $h(s)$, $s \in S$ *en een deelrij* $\{\alpha_{n_k}\}$ *met* $\alpha_{n_k} \to 0$ *als* $k \to \infty$ *zodat*

$$(1.24) \qquad \lim_{k\to\infty}\{V_{\alpha_{n_k}}(s) - V_{\alpha_{n_k}}(s^*)\} = h(s) \qquad \text{voor alle } s \in S.$$

Veronderstel dat bovendien aanname 5 *vervuld is. Dan geldt tevens*

(d) *De constante* g^* *en de funktie* $h(s)$ *voldoen aan de "optimaliteitsvergelijking" van het gemiddelde kosten kriterium,*

$$(1.25) \qquad h(s) = \min_{a\in A_s} \{c(s,a) - g^*\tau(s,a) + \sum_{s'\in S} h(s')q(s'|s,a)\}, \qquad s \in S.$$

Bewijs. (a) Kies $R \in C$ willekeurig. Uit (1.4)-(1.6), (1.23) en stelling 2 in de appendix volgt dat voor alle $s \in S$ geldt

$$(1.26) \qquad g(s,R) \geq \limsup_{\alpha\to 0} \alpha V_\alpha(s,R) \quad \text{en} \quad g(s,f^*) = \lim_{\alpha\to 0} \alpha V_\alpha(s,f^*).$$

Uit aanname 4 volgt nu dat voor alle $s \in S$ geldt

$$g(s,R) \geq \limsup_{\alpha\to 0} \alpha V_\alpha(s,R) \geq \limsup_{\alpha\to 0} \alpha V_\alpha(s) \geq \limsup_{n\to\infty} \alpha_n V_{\alpha_n}(s) =$$
$$= \limsup_{n\to\infty} \alpha_n V_{\alpha_n}(s,f^*) = \lim_{n\to\infty} \alpha_n V_{\alpha_n}(s,f^*) = g(s,f^*).$$

Derhalve is politiek f^* gemiddeld kosten optimaal aangezien R willekeurig gekozen was. De tweede bewering uit (a) volgt uit het feit dat $g(s,f^*)$ onafhankelijk van s is, zie (1.23).

(b) Deze bewering volgt direkt uit de tweede relatie in (1.26), aanname 4(a) en het feit dat $g(s,f^*) = g^*$ voor alle s.

(c) Wij zullen eerst aantonen dat voor elke vaste s de rij $\{V_{\alpha_n}(s) - V_{\alpha_n}(s^*), n\geq 1\}$ begrensd is. Uit aanname 4 en (1.7) volgt dat voor alle $s \in S$ en $n \geq 1$ geldt

$$(1.27)\qquad V_{\alpha_n}(s)-V_{\alpha_n}(s^*) = V_{\alpha_n}(s,f^*)-V_{\alpha_n}(s^*) \le$$

$$\le E_{s,f^*}(\underline{z}(\underline{\tau}))+V_{\alpha_n}(s^*,f^*)-V_{\alpha_n}(s^*) = E_{s,f^*}(\underline{z}(\underline{\tau})).$$

omgekeerd, met behulp van de ongelijkheid van Jensen (zie bijvoorbeeld blz. 148 in ROSS (1970)), de ongelijkheid $1-e^{-x} \le x$ voor $x \ge 0$ en bewering b) volgt uit (1.7) dat voor alle $s \in S$ geldt

$$(1.28)\qquad V_{\alpha_n}(s)-V_{\alpha_n}(s^*) =$$

$$= V_{\alpha_n}(s,f^*)-V_{\alpha_n}(s^*) \ge E_{s,f^*}(e^{-\alpha_n\underline{\tau}})V_{\alpha_n}(s^*,f^*)-V_{\alpha_n}(s^*) \ge$$

$$\ge e^{-\alpha_n E_{s,f^*}(\underline{\tau})} V_{\alpha_n}(s^*)-V_{\alpha_n}(s^*) \ge -E_{s,f^*}(\underline{\tau})\alpha_n V_{\alpha_n}(s^*) \ge$$

$$\ge \gamma E_{s,f^*}(\underline{\tau})$$

voor een eindige constante γ. Uit (1.27) en (1.28) volgt nu dat

$$\gamma E_{s,f^*}(\underline{\tau}) \le V_{\alpha_n}(s)-V_{\alpha_n}(s^*) \le E_{s,f^*}(\underline{z}(\underline{\tau})) \quad \text{voor alle } s \in S \text{ en } n \ge 1.$$

Toepassing van stelling 3 uit de appendix geeft nu dat (1.24) vervuld is voor een deelrij $\{\alpha_{n_k}\}$ en een eindige funktie $h(s)$ met

$$(1.29)\qquad \gamma E_{s,f^*}(\underline{\tau}) \le h(s) \le E_{s,f^*}(\underline{z}(\underline{\tau})) \qquad \text{voor alle } s \in S.$$

(d) Wij merken eerst op dat uit aanname 5 en (1.29) volgt

$$(1.30)\qquad \textstyle\sum_{s'} |h(s')|q(s'|s,a) < \infty \qquad \text{voor alle } s \in S \text{ en } a \in A_s.$$

Kies nu s vast. Aangezien A_s eindig is, kunnen wij een aktie $a_0 \in A_s$ en een deelrij $\{\alpha_{m_k}\}$ van $\{\alpha_{n_k}\}$ met $\alpha_{m_k} \to 0$ als $k \to \infty$ vinden zodat voor alle k geldt dat a_0 het rechterlid van (1.14) met $\alpha = \alpha_{m_k}$ minimaliseert. Derhalve geldt voor alle k en $a \in A_s$

(1.31) $$V_{\alpha_{m_k}}(s)-V_{\alpha_{m_k}}(s^*) \geq c_{\alpha_{m_k}}(s,a)+\textstyle\sum_{s'}q(s'|s,a)\beta_{\alpha_{m_k}}(s,a,s')\{$$
$$V_{\alpha_{m_k}}(s')-V_{\alpha_{m_k}}(s^*)\}+\textstyle\sum_{s'}q(s'|s,a)\{1-\beta_{\alpha_{m_k}}(s,a,s'\}V_{\alpha_{m_k}}(s^*),$$

waarbij het gelijkheidsteken geldt voor $a = a_0$. Uit (1.9) volgt dat voor alle s, a en s',

(1.32) $$\lim_{\alpha\to 0}\beta_\alpha(s,a,s') = 1 \text{ en } \lim_{\alpha\to 0}\alpha^{-1}\{1-\beta_\alpha(s,a,s'\} = -\int_0^\infty t dF(t|s,a,s').$$

Laat nu k naar ∞ gaan in (1.31). Dan volgt met behulp van de relaties (1.32), (1.30), (1.24), (1.18) en (1.1), bewering b) en de begrensde convergentiestelling dat

$$h(s) \geq c(s,a)+\textstyle\sum_{s'}q(s'|s,a)h(s')-g^*\tau(s,a) \qquad \text{voor alle } a \in A_s,$$

waarbij het gelijkheidsteken geldt voor $a = a_0$. Hiermee is bewering d) bewezen.*) □

Tenslotte zullen wij laten zien dat onder zekere voorwaarden uit de funktionaalvergelijking (1.25) een gemiddeld optimale politiek bepaald kan worden.

STELLING 1.3. *Veronderstel dat de aannames* 1 - 5 *vervuld zijn. Laat de stationaire politiek* f_0 *zo zijn dat* $f_0(s)$ *het rechterlid van* (1.25) *minimaliseert voor alle* $s \in S$. *Stel dat er een toestand* $s_0 \in S$ *bestaat zodat*

(1.33) $$E_{s,f_0}(\underline{\tau}_0) < \infty \quad \text{en} \quad E_{s,f_0}(\underline{z}(\underline{\tau}_0)) < \infty \qquad \textit{voor alle } s \in S,$$

waarbij $\underline{\tau}_0$ *het eerste tijdstip vallend na tijdstip* 0 *is waarop een overgang naar toestand* s_0 *plaatsvindt en* $\underline{z}(t)$ *de totale kosten in* $[0,t)$ *is. Neem tevens aan*

---

*) Uit het bewijs van bewering d) en opmerking 1.2 volgt dat $f^*(s)$ het rechterlid van (1.25) minimaliseert voor alle $s \in S$. Wij merken echter op dat in het algemeen niet geldt dat elke gemiddeld kosten optimale politiek $f \in F$ de eigenschap heeft dat $f(s)$ het rechterlid van (1.25) minimaliseert voor alle $s \in S$.

(1.34) $\qquad E_{s_0,f_0}(h(\underline{x}_k)) < \infty$ *voor alle* $k \geq 1$, *en* $\quad k^{-1}E_{s_0,f_0}(h(\underline{x}_k)) \to 0$ *als* $k \to \infty$. *)

Dan is politiek f_0 *gemiddeld kosten optimaal.*

Bewijs. Uit het bewijs van stelling 7.5 in ROSS (1970) volgt (in dit bewijs wordt aangenomen dat de beslissingkosten begrensd zijn; deze aanname kan echter zonder meer door de tweede aanname in (1.33) vervangen worden),

$$\lim_{t\to\infty} t^{-1}E_{s_0,f_0}(\underline{z}(t)) = \frac{\lim_{n\to\infty} n^{-1}\sum_{k=0}^{n} E_{s_0,f_0}[c(\underline{x}_k,f_0(\underline{x}_k))]}{\lim_{n\to\infty} n^{-1}\sum_{k=0}^{n} E_{s_0,f_0}[\tau(\underline{x}_k,f_0(\underline{x}_k))]}, \tag{1.35}$$

waarbij tevens geldt dat zowel de teller als de noemer van het rechterlid van (1.35) eindig zijn. Dit laatste impliceert overigens dat voor alle $k \geq 1$,

$$E_{s_0,f_0}[c(\underline{x}_k,f_0(\underline{x}_k))] < \infty \quad \text{en} \quad E_{s_0,f_0}[\tau(\underline{x}_k,f_0(\underline{x}_k))] < \infty. \tag{1.36}$$

Uit het feit dat onder politiek f_0 vanuit elke begintoestand het systeem met kans 1 na een eindige tijd een overgang maakt naar toestand s_0 waarbij de verwachting van de kosten te maken gedurende deze tijd eindig is, volgt eenvoudig dat

$$\lim_{t\to\infty} t^{-1}E_{s,f_0}(\underline{z}(t)) = \lim_{t\to\infty} t^{-1}E_{s_0,f_0}(\underline{z}(t)) \qquad \text{voor alle } s \in S, \tag{1.37}$$

m.a.w. $g(s,f_0)$ hangt niet van de begintoestand s af.

Gebruikmakend van het feit dat $f_0(s)$ het rechterlid van (1.25) minimaliseert voor alle s, volgt nu dat voor alle k=1,2,... geldt

$$E_{s_0,f_0}[h(\underline{x}_k)\,|\,\underline{x}_{k-1}=s] = \textstyle\sum_{s'} q(s'|s,f_0(s))h(s') =$$

$$= c(s,f_0(s))-g^*\tau(s,f_0(s)) + \textstyle\sum_{s'} q(s'|s,f_0(s))h(s')-c(s,f_0(s))+g^*\tau(s,f_0(s)) =$$

$$= h(s)-c(s,f_0(s))+g^*\tau(s,f_0(s)) \qquad \text{voor alle } s.$$

*) Uit (1.29) volgen direkt voldoende voorwaarden voor (1.34).

Derhalve vinden wij voor $k=1,2,\dots$

$$E_{s_0,f_0}[h(\underline{x}_k)] =$$
$$= E_{s_0,f_0}[h(\underline{x}_{k-1})-c(\underline{x}_{k-1},f_0(\underline{x}_{k-1}))+g^*\tau(\underline{x}_{k-1},f_0(\underline{x}_{k-1}))].$$

Gebruikmakend van (1.34) en (1.36) vinden wij vervolgens door beide leden van deze gelijkheid te sommeren over $k=1,\dots,n$,

$$E_{s_0,f_0}[h(\underline{x}_n)] - h(s_0) =$$
$$= -\sum_{k=1}^{n} E_{s_0,f_0}[c(\underline{x}_{k-1},f_0(\underline{x}_{k-1}))] + g^*\sum_{k=1}^{n} E_{s_0,f_0}[\tau(\underline{x}_{k-1},f_0(\underline{x}_{k-1}))]$$

voor alle $n \geq 1$. Delen wij beide leden van deze gelijkheid door n en laten wij n naar ∞ gaan, dan vinden wij met behulp van (1.34) en (1.35) dat $g(s_0,f_0) = g^*$. Vervolgens volgt uit (1.37) dat $g(s,f_0) = g^*$ voor alle $s \in S$, waarmee dan bewezen is dat politiek f_0 gemiddeld kosten optimaal is. □

Opmerking 1.3. In stelling 1.3 mogen wij de aanname (1.34) door een andere aanname vervangen. Daartoe merken wij eerst het volgende op. Op grond van (1.33) geldt dat onder politiek f_0 de Markov keten $\{\underline{x}_n\}$ een unieke stationaire kansverdeling $\{\pi_{f_0}(s), s\in S\}$ bezit (aanname 1 en (1.33) impliceren dat voor $\underline{n} = \inf\{n\geq 1 \mid \underline{x}_n = s_0\}$ geldt dat $P_{s,f_0}(\underline{n} < \infty) = 1$ en $E_{s_0,f_0}(\underline{n}) < \infty$ (zie Lemma 7.4 in ROSS (1970)), waaruit volgt dat onder politiek f_0 de Markov keten $\{\underline{x}_n\}$ een unieke stationaire kansverdeling bezit, zie bijv. blz. 32-34 in CHUNG (1960)). Stelling 1.3 blijft nu geldig als wij (1.34) vervangen door

$$(1.38) \qquad \textstyle\sum_s |h(s)|\pi_{f_0}(s) < \infty.$$

Het bewijs is eenvoudig en volgt door beide leden van de gelijkheid

$$h(s) = c(s,f_0(s)) - g^*\tau(s,f_0(s)) + \textstyle\sum_{s'} q(s'|s,f_0(s))h(s'), \quad s \in S$$

te vermenigvuldigen met $\pi_{f_0}(s)$, te sommeren over $s \in S$ en gebruik te maken van de relaties (vgl. bijv. paragraaf 2.3 in TIJMS (1974))

$$\pi_{f_0}(s') = \sum_s q(s'|s,f_0(s))\pi_{f_0}(s) \qquad \text{voor alle } s' \in S$$

en

$$g(f_0) = \sum_s c(s,f_0(s))\pi_{f_0}(s) \,/\, \sum_s \tau(s,f_0(s))\pi_{f_0}(s)$$

waarbij $g(f_0)$ de gemiddelde kosten per tijdseenheid onder politiek f_0 voorstelt.

2. WACHTTIJDPROBLEMEN MET BESTURING

2.0. *Inleiding*

In de eerste paragraaf van dit hoofdstuk zullen wij een overzicht geven van een aantal belangrijke bestuurde wachttijdproblemen die in de literatuur bestudeerd zijn en waarvoor zowel het bestaan als de struktuur van een gemiddeld kosten optimale politiek zijn aangetoond. Veel van deze problemen zijn nogal adhoc geanalyseerd. Niettemin kunnen een aantal hoofdtechnieken worden onderscheiden. Op deze technieken zullen wij nu in het kort ingaan.

1) *De aanpak via de funktionaalvergelijking voor het gemiddelde kosten kriterium*

Soms is het mogelijk om op direkte wijze de struktuur van een gemiddeld kosten optimale stationaire politiek vast te stellen door de funktionaalvergelijking (1.25) uit paragraaf 1.4 te analyseren. Deze aanpak kan met name succesvol zijn in problemen waarin de tijden tussen de toestandsovergangen exponentieel zijn en vanuit toestand i alleen een overgang mogelijk is naar toestand i+1 of toestand i-1. Daarbij beschouwt men soms eerst een "afgeknot" probleem waarvoor de toestandsruimte eindig is zodat de voorwaarden waaronder (1.25) geldt vervuld zijn, waarna men de struktuur van een optimale politiek voor het afgeknotte probleem bewijst en vervolgens aantoont dat deze politiek ook optimaal is voor het oorspronkelijke probleem.

Voor voorbeelden met de aanpak via de funktionaalvergelijking verwijzen wij naar CRABILL (1972), LOW (1971,1974), MILLER (1969) en ROSS (1970).

2) *De aanpak via het verdisconteerde kosten model*

Bij deze aanpak beschouwt men eerst het verdisconteerde model met een eindig aantal beslissingstijdstippen en analyseert men voor dit model de funktionaalvergelijking (1.13). Met volledige induktie probeert men dan de gewenste eigenschappen van de funktie $V_{n+1,\alpha}(s)$ (bijv. convexiteit) en de gewenste struktuur van de minimaliserende politiek $f_{n,\alpha}$ te bewijzen. Vervolgens door n naar ∞ te laten gaan, worden deze resultaten in het algemeen overgedragen op de minimale verdisconteerde kosten $V_\alpha(s)$ en een bijbehorende minimaliserende politiek f_α, zie stelling 1.1 in paragraaf 1.3. Tenslotte kan men dan proberen hetzij via stelling 1.2 uit paragraaf 1.4 hetzij via argumenten analoog aan die in paragraaf 1.4 het bestaan en de struktuur van een gemiddeld kosten optimale stationaire politiek vast te stellen.

Deze aanpak die uiteraard ook toepasbaar is voor problemen met een continue toestandsruimte, is in een groot aantal wachttijdproblemen gevolgd, onder meer in DEB & SERFOZO (1973), LIPPMANN & ROSS (1971) en MITCHELL (1973). Niettemin is deze aanpak in wachttijdtheorie niet zo succesvol gebleken als in voorraadtheorie waar deze aanpak standaard is. De reden hiervan is dat in tegenstelling tot de meeste voorraadproblemen in wachttijdproblemen de tijden tussen opeenvolgende toestandsovergangen niet constant zijn maar afhangen van de laatst waargenomen toestand en de daarin genomen aktie. Dit veroorzaakt dat in veel wachttijdproblemen bovengenoemde induktieaanpak vastloopt doordat bijv. de convexiteit van de funktie $V_{n+1,\alpha}$ niet geldig blijft. Het bezwaar dat de tijden tussen opeenvolgende toestandsovergangen niet onafhankelijk zijn van de laatst waargenomen toestand en de daarin genomen aktie kan ondervangen worden als deze tijden exponentieel verdeeld zijn. In dat geval kan een equivalent probleem geformuleerd worden waarin de tijden tussen opeenvolgende toestandsovergangen ook exponentieel verdeeld zijn maar onafhankelijk zijn van de waargenomen toestanden en de daarin genomen akties. Deze herformulering is afkomstig van LIPPMAN (1973c). In paragraaf 2.2.1 zullen wij de herformulering geven en in paragraaf 2.3 geven wij een toepassing op een M/M/1 wachtrij met variabele bedieningssnelheid. Voor een aantal andere toepassingen van deze techniek verwijzen wij naar LIPPMAN (1973c).

3) *De aanpak via vertaling in een optimaal stopprobleem.*

In wachttijdproblemen waarin alleen "ja of nee" beslissingen mogelijk zijn, kan nadat het bestaan van een gemiddeld kosten optimale stationaire politiek is aangetoond de struktuur van deze politiek soms worden bewezen door de minimalisatie van de gemiddelde kosten te vertalen in een optimaal stopprobleem. Met behulp van de stelling voor het monotoniciteitsgeval in optimaal stoppen kan dan de struktuur van een gemiddeld kosten optimale stationaire politiek worden bewezen. Deze aanpak is afkomstig van BELL (1971,1973). In paragraaf 2.2.2 zullen wij de benodigde theorie uit optimaal stoppen geven, terwijl in paragraaf 2.4 wij een toepassing geven op een wachttijdprobleem met stations waarbij klanten van het ene naar het andere station getransporteerd worden als de rijlengtes bij de stations daartoe aanleiding geven. Dit probleem is afkomstig van IGNALL & KOLESAR (1973) die dit probleem op een andere wijze hebben opgelost door namelijk gebruik te maken van resultaten voor een sterk verwant probleem bestudeerd door DEB & SERFOZO (1973). Voor andere toepassingen van bovenstaande aanpak verwijzen wij naar BELL (1971,1973b).

2.1. *Een overzicht van de literatuur over wachttijdproblemen met besturing.**)

In dit overzicht bespreken wij een aantal van de belangrijkste van de in de literatuur behandelde besturingsproblemen in wachttijdtheorie. Hierbij maken wij ruwweg onderscheid tussen problemen waarin het aankomstproces wordt bestuurd en problemen waarin het bedieningsproces wordt bestuurd.

a) *Besturing van het aankomstproces.*

Een extreem geval van besturing van het aankomstproces is het geval waarin een aankomende klant òf geaccepteerd wordt òf geweigerd wordt. Voorbeelden van deze besturing zijn gegeven in MILLER (1969) en in LIPPMAN & ROSS (1971).

MILLER (1969) beschouwt een M/M/c wachttijdsysteem zonder wachtruimte. De klanten kunnen in m klassen worden onderverdeeld, waarbij de bedieningstijd voor elke klant eenzelfde exponentiële verdeling heeft. Voor elke klant die bediend wordt, is er een opbrengst welke afhangt van de klasse waartoe die klant behoort. Als een klant aankomt terwijl één of meer bedienden vrij zijn dan moet beslist worden of de klant geaccepteerd wordt of niet. MILLER toonde aan dat een gemiddeld kosten optimale stationaire politiek bestaat met de eigenschap dat als een klant van type k wordt geaccepteerd als i bedienden vrij zijn deze klant ook geaccepteerd wordt als $j > i$ bedienden vrij zijn. Uitbreidingen van dit model van Miller zijn behandeld in CRAMER (1971) en LIPPMAN (1973c).

---

*) Een literatuuroverzicht kan ook gevonden worden in PRABHU & STIDHAM (1974).

LIPPMAN & ROSS (1971) beschouwen een G/G/1 wachttijdsysteem zonder wachtruimte. Elke klant biedt een offerte aan waarbij het type van de offerte een gegeven kansverdeling volgt. Een offerte van type x brengt een hoeveelheid werk en een bijbehorende beloning mee die afhangen van x en waarvan de gezamenlijke kansverdeling bekend is. Als een klant aankomt terwijl de bediende vrij is, dan moet de bediende beslissen of de klant geaccepteerd wordt of niet. Onder bepaalde voorwaarden hebben LIPPMAN & ROSS aangetoond dat een gemiddeld kosten optimale stationaire politiek bestaat met een eenvoudige struktuur (namelijk accepteer een klant met een offerte van type x alleen dan als een bepaald quotient afhangend van x een bepaald getal overschrijdt).

LOW (1971,1974) beschouwt een M/M/c wachttijdsysteem waarin het aankomstproces bestuurd wordt door een bepaalde advertentieprijs te voeren. Als een advertentieprijs p wordt gevoerd, dan komen de klanten aan volgens een Poisson proces met parameter λ_p. Elke klant die zich tot het systeem voegt moet bij aankomst de geldende advertentieprijs betalen. Daarnaast zijn er voorraadkosten die een niet-dalende funktie van het aantal aanwezige klanten zijn. De te voeren advertentieprijs mag zowel na aankomst van een klant als na afloop van een bediening herzien worden. Zowel voor het geval met een beperkte wachtruimte als voor het geval met een onbeperkte wachtruimte toont LOW aan dat onder zekere voorwaarden een gemiddeld kosten optimale stationaire politiek bestaat met de eigenschap dat de voorgeschreven advertentieprijs een niet-dalende funktie is van het aantal aanwezige klanten (een ander bewijs van dit resultaat kan gevonden worden in LIPPMAN (1973c)). Daarnaast geeft LOW voor het geval dat er slechts een eindig aantal advertentieprijzen mogelijk zijn een efficient algoritme om de "breekpunten" van de optimale prijzenpolitiek te berekenen.

In dit verband vermelden wij ook het werk van ADIRI & YECHIALI (1973) die een M/M/1 wachttijdsysteem met N wachtrijen beschouwen waarbij de i^{de} rij prioriteit heeft boven de j^{de} rij als $i < j$. Elke aankomende klant verkrijgt alle mogelijke informatie over de aanwezige wachtrijen en op grond van deze informatie moet hij beslissen of hij zich tot het systeem voegt of niet en zo ja tot welke rij, waarbij de klant een bedrag $u \leq \infty$ ontvangt als hij zich tot het systeem voegt. Daartegenover echter staat dat de klant een tol θ_j aan het bedieningsstation moet betalen als hij zich tot rij j voegt.

Voorts loopt de klant wachttijdkosten op die evenredig zijn aan de tijd die hij in het systeem doorbrengt. Tenslotte veroorzaakt een klant die niet tot het systeem toetreedt vaste kosten K voor het bedieningsstation. ADIRI & YECHIALI laten zien hoe bij gegeven waarden van $\theta_1,\ldots,\theta_N$ een optimale beslissing voor een afzonderlijke klant luidt. Daarnaast bepalen zij onder de aanname dat elke afzonderlijke klant voor zich zelf optimaal te werk gaat hoe het bedieningsstation de waarden van $\theta_1,\ldots,\theta_N$ moet kiezen opdat het verschil van de gemiddelde tolopbrengsten minus de gemiddelde kosten veroorzaakt door klanten die zich niet tot het systeem voegen maximaal is.

Tenslotte vermelden wij dat ook in NAOR (1969) en YECHIALI (1971,1972) wachttijdproblemen beschouwd worden waarin het aankomstproces bestuurd wordt door het heffen van een tol aan de klanten.

2) *Besturing van het bedieningsproces.*

Wij zullen eerst een aantal problemen bespreken waarin het bedieningsmechanisme zowel ingeschakeld als uitgeschakeld kan zijn.

HEYMAN (1968) en BELL (1971) beschouwen een M/G/1 wachttijdsysteem waarin de bediende na afloop van een bediening "uit" gedaan kan worden en weer "aan" kan worden gezet bij aankomst van een klant die de bediende "uit" treft. Er zijn vaste kosten verbonden zowel aan het uitzetten als aan het aanzetten van de bediende. Daarnaast worden voor elke klant wachttijdkosten gemaakt die evenredig zijn aan de tijd dat de klant in het systeem doorbrengt. De artikelen van HEYMAN en BELL tesamen (vgl. ook LIPPMAN (1973b)) tonen aan dat een gemiddeld kosten optimale stationaire politiek bestaat van de volgende vorm: doe de bediende aan als het aantal aanwezige klanten groter dan of gelijk aan een bepaalde waarde N is en òf doe de bediende alleen dan uit wanneer het systeem leeg is òf doe de bediende nooit uit.

SOBEL (1970) beschouwde eenzelfde probleem met een algemenere kostenstruktuur voor het G/G/1 wachttijdsysteem en toonde aan dat binnen de klasse van de stationaire politieken een gemiddeld kosten optimale politiek bestaat van de volgende vorm: de bediende is uit als het aantal aanwezige klanten kleiner dan of gelijk aan m is en doe de bediende aan zodra het aantal aanwezige klanten toeneemt tot M > m.

BLACKBURN (1971,1972) behandelt een uitbreiding van bovengenoemd model van Heyman door convexe wachttijdkosten te beschouwen en door toe te laten dat voor het geval de bediende uit is een aankomende klant zich niet tot het systeem voegt en een wachtende klant het systeem verlaat.

BELL (1973) beschouwt een M/G/1 "non-preemptive" prioriteitsmodel met twee typen van klanten. Voor beide klantentypen volgt de bedieningstijd eenzelfde kansverdeling. De bediende kan uitgedaan worden na afloop van een bediening en aangedaan worden bij aankomst van een klant die de bediende uit treft. Er zijn vaste kosten verbonden zowel aan het uitdoen als aan het aandoen van de bediende. Daarnaast zijn er voor elke klant van type i wachttijdkosten h_i voor elke tijdseenheid dat de klant in het systeem is. BELL toont aan dat een gemiddeld kosten optimale stationaire politiek bestaat van de volgende vorm: doe de bediende aan zodra een bepaalde lineaire combinatie van het aantal aanwezige klanten van type 1 en het aantal aanwezige klanten van type 2 een zekere waarde overschrijdt, en doe de bediende alleen dan uit wanneer het systeem leeg is of doe de bediende nooit uit.

Naast dit prioriteitsmodel noemen wij een tweetal prioriteitsmodellen bestudeerd door HARRISON (1972b) en BELL (1973b).

HARRISON beschouwt het klassieke M/G/1 "non-preemptive" prioriteitsmodel met k typen van klanten waarbij voor elke klant van type i wachttijdkosten c_i worden gemaakt voor elke tijdseenheid dat de klant in het systeem is. Het werk van HARRISON / LIPPMAN (1973b) toont aan dat een gemiddeld kosten optimale stationaire politiek bestaat die prioriteit geeft aan de klanten overeenkomstig de afnemende volgorde van de quotienten c_i/μ_i waarbij μ_i de verwachte bedieningstijd van een klant van type i is.

BELL (1973b) beschouwt een M/G/1 "non-preemptive" wachttijdsysteem met klanten van type 1 en van type 2, waarbij voor elke klant van type i wachttijdkosten h_i worden gemaakt voor elke tijdseenheid dat de klant in het systeem is. De bedieningstijd van een klant van type i heeft als verwachting μ_i. Wij nemen aan dat $h_1/\mu_1 > h_2/\mu_2$. De klanten worden in volgorde van aankomst bediend, met dien verstande echter dat na afloop van een bediening de bediende een klant van type 1 voor zijn beurt mag laten gaan tenkoste van een bedrag R. Uit het werk van BELL / LIPPMAN (1973b) blijkt dat een gemiddeld kosten optimale stationaire politiek bestaat van de volgende vorm: als na afloop van een bediening een klant van type 2 als eerste in de rij staat en er is een klant van type 1 in de rij met tenminste n^* klanten van type 2 voor zich, bedien dan als volgende klant de klant van type 1 die het meest achteraan in de rij staat. Hierbij is n^* het kleinste gehele getal met $n^* \geq R/(h_1\mu_2-h_2\mu_1)$.

DEB & SERFOZO (1973) beschouwen een M/G/1 wachttijdsysteem waarin de bediende meer dan één klant tegelijk kan helpen met een maximum van Q klanten waarbij $Q \leq \infty$. De bediende behoeft niet noodzakelijk aan het werk te zijn als er klanten aanwezig zijn. Zowel na afloop van een bediening als bij aankomst van een klant die de bediende "uit" treft, moet de bediende beslissen of een volgende bediening aanvangt of niet en zo ja hoeveel klanten tijdens die bediening tegelijk geholpen worden. De kosten van een bediening waarin j klanten tegelijk geholpen worden, bedragen K + cj. Daarnaast zijn er wachttijdkosten die een stijgende funktie van het aantal aanwezige klanten zijn. DEB & SERFOZO tonen aan dat onder zekere voorwaarden een gemiddeld kosten optimale stationaire politiek bestaat die vastgelegd wordt door een zeker natuurlijk getal M en de volgende vorm heeft: als na afloop van een bediening of na aankomst van een klant die de bediende "uit" treft k klanten in het systeem zijn, vangt geen bediening aan als $k < M$ en vangt een bediening waarin min(k,Q) klanten worden bediend als $k \geq M$.

Vervolgens bespreken wij enkele wachttijdproblemen waarin verschillende bedieningssnelheden mogelijk zijn of waarin het aantal ingezette bedienden gewijzigd kan worden.

CRABILL (1972) beschouwt een M/M/1 wachtrij met een eindig aantal exponentiële bedieningssnelheden waarin zowel bij aankomst van een klant of na afloop van een bediening de gevoerde bedieningssnelheid gewijzigd kan worden. Er zijn geen kosten verbonden aan het wijzigen van de bedieningssnelheid. De wachttijdkosten zijn een niet-dalende funktie van het aantal aanwezige klanten. Daarnaast zijn er bedieningskosten die afhangen van de gevoerde bedieningssnelheid. Onder zwakke voorwaarden heeft CRABILL aangetoond dat binnen de klasse van de stationaire politieken een gemiddeld kosten optimale politiek bestaat met de eigenschap dat de voorgeschreven bedieningssnelheid een niet-dalende funktie is van het aantal aanwezige klanten. Dit resultaat is op een andere wijze ook bewezen door LIPPMAN (1973a,1973c) die tevens aantoonde dat deze politiek ook gemiddeld kosten optimaal is binnen de klasse van alle mogelijke politieken. Evenals ZACKS & YADIN (1970) behandelt LIPPMAN (1973c) ook het geval dat de bedieningssnelheid gekozen wordt uit een gesloten interval.

CRABILL (1973) behandelt het M/M/1 wachttijdsysteem met variabele bedieningssnelheid waarbij omschakelkosten optreden. Voor het geval van twee mogelijke bedieningssnelheden toont hij aan dat een gemiddeld kosten optimale stationaire politiek bestaat van de volgende vorm: schakel van snelheid 1 naar snelheid 2 als het aantal aanwezige klanten aangroeit tot een waarde i_1, en schakel van snelheid 2 naar snelheid 1 als het aantal aanwezige klanten terugvalt tot een waarde $i_2 < i_1$.

CRABILL (1974) beschouwt het eindige machine reparatieprobleem met één reparateur, een tweetal mogelijke exponentiële reparatiesnelheden en geen omschakelkosten. De kosten bestaan uit wachttijdkosten en reparatiekosten. Voor dit probleem toont Crabill aan dat onder zekere voorwaarden een gemiddeld kosten optimale stationaire politiek bestaat met de eigenschap dat de voorgeschreven reparatiesnelheid een niet-dalende funktie van het aantal defekte machines is.

THATCHER (1968) en MITCHEL (1973) hebben M/G/1 wachttijdproblemen beschouwd waarin de bedieningssnelheid gebaseerd wordt op de hoeveelheid werk in het systeem.

THATCHER (1968) beschouwt het geval waar twee bedieningssnelheden 1 en 2 ter beschikking zijn waarbij een hoeveelheid werk σ_i per tijdseenheid verwerkt wordt als snelheid i wordt gebruikt. Op elk gewenst moment kan van bedieningssnelheid veranderd worden waaraan geen omschakelkosten verbonden zijn. De kosten bestaan uit voorraadkosten h per eenheid werk aanwezig in het systeem per tijdseenheid en uit bedieningskosten r_i per tijdseenheid als snelheid i wordt gebruikt. THATCHER toonde aan dat binnen de klasse van de stationaire politieken een gemiddeld kosten optimale politiek bestaat van de volgende vorm: gebruik snelheid 1 resp. snelheid 2 als de hoeveelheid werk in het systeem kleiner dan of gelijk aan resp. groter dan een bepaalde waarde y^* is.

MITCHELL (1973) behandelt een analoog model waarin de bedieningssnelheid gekozen wordt uit een gesloten interval. Onder zekere voorwaarden toonde hij aan dat een gemiddeld kosten optimale stationaire politiek bestaat met de eigenschap dat de voorgeschreven bedieningssnelheid een niet-dalende funktie is van de aanwezige hoeveelheid werk in het systeem.

Tenslotte vermelden wij het door LIPPMAN (1973c) behandelde M/M/c wachttijdprobleem met een variabel aantal bedienden. In dit probleem staan c bedienden ter beschikking en kan zowel bij aankomst van een klant als na afloop van een bediening het aantal ingezette bedienden gewijzigd worden. De kosten bestaan uit wachttijdkosten h per klant per tijdseenheid dat de klant in het systeem is, bedieningskosten die evenredig zijn met het aantal ingezette bedienden, en omschakelkosten $k_1 \cdot (j-i)$ resp. $k_2 \cdot (i-j)$ als het aantal ingezette bedienden gewijzigd wordt van i naar j voor $j > i$ resp. $j < i$. LIPPMAN toonde aan dat een gemiddeld kosten optimale stationaire politiek bestaat die bepaald wordt door paren van gehele getallen $(s(i),S(i))$, $i=0,1,\ldots$ en er als volgt uitziet: als na aankomst van een klant of na afloop van een bediening het aantal aanwezige klanten i is en k bedienden zijn ingezet, voer het aantal ingezette bedienden op tot s(i) als $k < s(i)$, laat het aantal ingezette bedienden ongewijzigd als $s(i) \le k \le S(i)$, en breng het aantal ingezette bedienden terug tot S(i) als $k > S(i)$.*)

2.2. *Enige algemene resultaten.*

2.2.1. *Een equivalente formulering voor het semi-Markov beslissingsmodel met exponentieel verdeelde transitietijden.*

Beschouw eerst een semi-Markov proces met een hoogstens aftelbare toestandsruimte I waarbij voor de overgangskansen q_{ij} en de verdelingsfunkties $F_{ij}(t)$ geldt

$$q_{ii} = 0 \text{ voor alle } i \in I^{**)} \text{ en } F_{ij}(t) = 1-e^{-\lambda_i t} \text{ voor alle } i,j \in I.$$

Een dergelijk semi-Markov proces heet een Markov keten met continue tijdsparameter (vgl. blz. 110 in ROSS (1970)). Definieer $p_{ij}(t)$ als de kans dat het proces op tijdstip t in toestand j is gegeven dat het proces op tijdstip 0 in toestand i is, $i,j \in I$ en $t \ge 0$. Veronderstel dat

$$\sup_{i \in I} \lambda_i < \infty.$$

*) Achteraf blijkt dat LIPPMAN's bewijs een fundamentele fout bevat (het bewijs van stelling 1 is onjuist), zodat dit beroemde probleem nog steeds open is.

**) De aanname $q_{ii} = 0$ voor alle i is niet essentieel.

Laat γ nu een willekeurig eindig getal zijn met

$$\gamma \geq \sup_{i \in I} \lambda_i .$$

Beschouw nu het semi-Markov proces met toestandsruimte I waarvan de overgangskansen $\bar{q}_{ij}$ en de verdelingsfunkties $\bar{F}_{ij}(t)$ voor alle $i,j \in I$ worden gegeven door

$$\bar{q}_{ij} = \begin{cases} (\lambda_i/\gamma)q_{ij} & \text{voor } j \neq i \\ (\gamma-\lambda_i)/\gamma & \text{voor } j = i \end{cases}$$

en

$$\bar{F}_{ij}(t) = 1-e^{-\gamma t}.$$

Definieer voor dit nieuwe semi-Markov proces $\bar{p}_{ij}(t)$ als de kans dat het proces op tijdstip t in toestand j is gegeven dat het proces op tijdstip 0 in toestand i is, $i,j \in I$ en $t \geq 0$. Dan geldt

$$\bar{p}_{ij}(t) = p_{ij}(t) \qquad \text{voor alle } i,j \in I \text{ en alle } t \geq 0, \tag{2.1}$$

m.a.w. het oorspronkelijke en het nieuwe semi-Markov proces zijn kanstheoretisch gezien identiek. Het bewijs van (2.1) is eenvoudig en kan gevonden worden op blz. 563 in KEILSON & WISHART (1964). De volgende physische interpretatie kan aan het nieuwe semi-Markov proces gegeven worden: "In het nieuwe proces worden de tijdstippen van de toestandsovergangen voortgebracht door een Poisson proces met parameter γ dat onafhankelijk is van het verloop van de toestand van het systeem. Onder de toestandsovergangen zijn "schijnovergangen" waarbij de toestand niet verandert. Na een toestandsovergang naar toestand i geldt dat de volgende overgang met kans $(\gamma-\lambda_i)/\gamma$ een "schijnovergang" naar toestand i is, en met kans λ_i/γ een "echte" toestandsverandering is die op zijn beurt met kans q_{ij} een toestandsverandering in toestand j is. Vanuit toestand i is de wachttijd tot de eerstvolgende "echte" toestandsverandering exponentieel verdeeld met verwachting $1/\lambda_i$ en het aantal transities in deze tijd is geometrisch verdeeld met verwachting γ/λ_i."

Wij keren nu weer terug naar het semi-Markov beslissingsmodel uit paragraaf 1.2. Veronderstel dat voor alle $s, s' \in S$ en $a \in A_s$,

$$q(s|s,a) = 0 \quad \text{en} \quad F(t|s,a,s') = 1-e^{-\lambda_s(a)t}.$$

Voor het gemak nemen wij tevens aan dat de direkte kosten $d_1(s,a)$ identiek gelijk aan nul zijn. Veronderstel dat

$$\sup_{s\in S,\, a\in A_s} \lambda_s(a) < \infty.$$

Uit deze veronderstelling volgt dat aanname 1 uit paragraaf 1.2 vervuld is. Neem tevens aan dat de aannames 2 en 3 uit paragraaf 1.3 vervuld zijn.

Laat nu γ een willekeurig eindig getal zijn met

$$\gamma \geq \lambda_s(a) \qquad \text{voor alle } s \in S \text{ en alle } a \in A_s.$$

Beschouw nu het volgende semi-Markov beslissingsmodel waarvan de toestandsruimte S en de verzameling A_s van mogelijke akties hetzelfde zijn als in bovenstaand beslissingsmodel. De overgangskosten, de verdelingsfunkties van de transitietijden en de kostenfunkties worden gegeven door

$$\bar{q}(s'|s,a) = \begin{cases} (\gamma-\lambda_{s,a})/\gamma & \text{voor } s' = s \\ (\lambda_{s,a}/\lambda)q(s'|s,a) & \text{voor } s' \neq s \end{cases}$$

$$\bar{F}(t|s,a,s') = 1-e^{-\gamma t}$$

$$\bar{d}_1(s,a) = 0,^{*)} \quad \text{en} \quad \bar{d}_2(s,a) = d_2(s,a).$$

Definiëren wij voor dit nieuwe model $\bar{g}(s,R)$ en $\bar{V}_\alpha(s,R)$ overeenkomstig (1.5) en (1.6), dan volgt eenvoudig met behulp van (2.1) dat voor elke stationaire politiek f geldt

---

*) Ingeval men uiteindelijk alleen in het gemiddelde kostenkriterium is geïnteresseerd, neem dan $\bar{d}_1(s,a) = d_1(s,a)\lambda_{s,a}/\gamma$ in het nieuwe model als $d_1(s,a)$ niet identiek gelijk aan nul is.

$$\bar{g}(s,f) = g(s,f) \quad \text{en} \quad \bar{V}_\alpha(s,f) = V_\alpha(s,f) \qquad \text{voor alle } s \in S. \tag{2.2}$$

Tevens gaat men direkt nadat voor dit nieuwe model de aannames 1 - 3 uit de paragrafen 1.2 en 1.3 vervuld zijn. Derhalve geldt voor het nieuwe model dat voor elke $\alpha > 0$ een α-optimale stationaire politiek bestaat. Uit (2.2) volgt nu dat een stationaire politiek die α-optimaal resp. gemiddeld kosten optimaal is voor het nieuwe model ook α-optimaal resp. gemiddeld kosten optimaal is voor het oorspronkelijke model. Dit resultaat is zeer bruikbaar om de struktuur van een optimale politiek te bewijzen, aangezien het nieuwe model zich in toepassingen beter leent voor analyse dan het oorspronkelijke model. De reden hiervan is dat in het nieuwe model de transitietijden niet van de toestand en de gekozen aktie afhangen waardoor de funktionaalvergelijking (1.13) een vereenvoudigde vorm verkrijgt die uitermate geschikt is om allerlei prettige eigenschappen zoals bijv. convexiteit met behulp van induktie te bewijzen.

2.2.2. *Optimaal stoppen.*

Beschouw de situatie waarin wij achtereenvolgens op de tijdstippen $n=0,1,\ldots$ waarden observeren van de stochastische variabelen (of stochastische vektoren) $\underline{y}_0, \underline{y}_1, \ldots$ waarvan de gezamenlijke kansverdeling bekend is. Na een eindig aantal waarnemingen moeten wij de observatie van het proces stoppen. Als de observatie van het proces gestopt wordt na tijdstip n nadat achtereenvolgens de waarden $y_0,\ldots,y_n$ zijn waargenomen, dan worden stopkosten $c(y_0,\ldots,y_n)$ gemaakt.

Een stopregel is een regel die na elke waarneming voorschrijft of de verdere observatie van het proces al of niet gestopt wordt waarbij de beslissing om na tijdstip n de verdere observatie van het proces al of niet te stoppen alleen mag afhangen van de tot en met tijdstip n waargenomen waarden $y_0,\ldots,y_n$. Laat D de klasse van stopregels zijn die met kans 1 de observatie van het proces na een begrensd aantal waarnemingen stoppen. Voor elke stopregel $R \in D$, definieer C_R als de verwachte stopkosten bij toepassing van stopregel R.

Veronderstel nu dat de funktie c de volgende monotoniciteitseigenschap bezit: voor alle $n=0,1,\ldots$ geldt dat

$$(2.3)\qquad Ec(y_0,\ldots,y_n,\underline{y}_{n+1}) - c(y_0,\ldots,y_n) \ge 0 \Rightarrow$$

$$Ec(y_0,\ldots,y_{n+1},\underline{y}_{n+2}) - c(y_0,\ldots,y_{n+1}) \ge 0 \qquad \text{voor alle } y_{n+1}.$$

Veronderstel verder dat de stopregel R^* gedefinieerd door: "R^* stopt op het eerste tijdstip n waarvoor

$$Ec(y_0,\ldots,y_n,\underline{y}_{n+1}) - c(y_0,\ldots,y_n) \ge 0"$$

tot de klasse D behoort. Dan geldt

$$C_{R^*} \le C_R \qquad \text{voor alle } R \in D,$$

m.a.w. R^* is een optimale stopregel binnen de klasse D van stopregels.

Dit resultaat is in een algemener verband bewezen op blz. 55 in CHOW, ROBBINS & SIEGMUND (1971).

Tenslotte geven wij het volgende lemma waarmee soms een gemiddeld kosten probleem omgezet kan worden in een optimaal stopprobleem.

LEMMA 2.1. *Laten* $a(x)$ *en* $b(x)$ *twee reële funkties zijn gedefinieerd op een verzameling* X *waarbij* $b(x) \neq 0$ *voor alle* x. *Veronderstel dat*

$$\gamma^* = \min_{x\in X} a(x)/b(x)$$

bestaat en eindig is. Dan geldt dat $\min_{x\in X}\{a(x)-\gamma^* b(x)\} = 0$. *Verder geldt dat elke* x *waarvoor* $a(x) - \gamma^* b(x)$ *minimaal is ook* $a(x)/b(x)$ *minimaliseert, en omgekeerd.*

Bewijs. Stel dat $a(x)/b(x)$ minimaal is voor $x = x^*$. Dan geldt $a(x)-\gamma^* b(x) \ge 0$ voor alle x, waarbij het gelijkheidsteken geldt voor $x = x^*$. Hieruit volgt dat $\min_x\{a(x)-\gamma^* b(x)\} = 0$ en dat $a(x) - \gamma^* b(x)$ minimaal is voor $x = x^*$. Tenslotte stel dat $a(x) - \gamma^* b(x)$ minimaal is voor $x = y^*$. Dan volgt dat $a(x)-\gamma^* b(x) \ge 0$ voor alle x, waarbij het gelijkheidsteken geldt voor $x = y^*$. Hieruit volgt dat $a(x)/b(x) \ge \gamma^*$ voor alle x, waarbij het gelijkheidsteken geldt voor $x = y^*$ waarmee dan aangetoond is dat $a(x)/b(x)$ minimaal is voor $x = y^*$. □

2.3. *Een* M/M/1 *wachtrij met variabele bedieningssnelheid.*

Beschouw een bedieningsinstallatie met 1 bediende waar klanten aankomen volgens een Poisson proces met parameter λ. De bedieningstijden van de klanten zijn onderling onafhankelijke stochastische variabelen. De bediende heeft een M-tal bedieningssnelheden $1,\ldots,M$ tot zijn beschikking. De tijd nodig om een klant te bedienen als bedieningssnelheid k gebruikt wordt is een exponentieel verdeelde stochastische variabele met verwachting $1/\mu_k$, waarbij

$$\mu_1 < \mu_2 < \ldots < \mu_M \quad \text{en} \quad \lambda/\mu_M < 1.$$

De bediende mag zowel bij aankomst van een klant als bij afloop van een bediening van bedieningssnelheid wisselen. Met een wisseling van bedieningssnelheid is geen tijd gemoeid. De klanten worden één voor één bediend en een klant die bij aankomst de bediende bezet treft blijft wachten totdat hij bediend wordt.

De volgende kosten treden op. Als de installatie ingesteld is op bedieningssnelheid k dan worden per tijdseenheid bedieningskosten $r_k \geq 0$ gemaakt, $k=1,\ldots,M$. Daarnaast zijn er wachttijdkosten $h(i)$ voor elke tijdseenheid dat i klanten in het systeem zijn. Wij nemen aan dat $h(i)$ een nietnegatieve, convexe funktie is die begrensd wordt door een polynoom van een eindige orde. Tevens nemen wij aan dat voor alle i geldt $h(i+1)-h(i) \geq h$ voor een $h > 0$.

Voor dit probleem zullen wij nu aantonen dat een gemiddeld kosten optimale stationaire politiek bestaat met de eigenschap dat de voor te schrijven bedieningssnelheid een niet-dalende funktie van het aantal aanwezige klanten is.

Dit model is een semi-Markov beslissingsmodel. De mogelijke toestanden worden gegeven door $i=0,1,\ldots$ waarbij toestand i correspondeert met de situatie waar i klanten aanwezig zijn, en voor elke toestand i worden de mogelijke akties gegeven door $a=1,\ldots,M$ waarbij aktie a het gebruik van bedieningssnelheid a voorschrijft. De overgangskosten $q(j|i,a)$, de verdelingsfunkties $F(t|i,a,j)$ en de kostenfunkties $d_1(i,a)$ en $d_2(i,a)$ worden gegeven door

$$q(i+1|i,a) = \lambda/(\lambda+\mu_a),\ q(i-1|i,a) = \mu_a/(\lambda+\mu_a) \text{ voor } i \geq 1,\ q(1|0,a) = 1,$$

$$F(t|i,a,j) = 1-e^{-(\lambda+\mu_a)t} \quad \text{voor } i \geq 1,\ F(t|0,a,j) = 1-e^{-\lambda t},$$

$$d_1(i,a) = 0,\quad d_2(i,a) = h(i)+r_a.$$

Men gaat direkt na dat de aannames 1 - 3 uit hoofdstuk 1 vervuld zijn. Om het bestaan en de struktuur van een gemiddeld optimale stationaire politiek te bewijzen, beschouwen wij het gewijzigde probleem zoals geformuleerd in paragraaf 2.2. Kiezen wij $\gamma = \lambda+\mu_M$, dan geldt voor het gewijzigde probleem (met weglating van de bovenstrepen)

$$q(i+1|i,a) = \lambda/\gamma,\ q(i-1|i,a) = \mu_a/\gamma,\ q(i|i,a) = (\mu_M-\mu_a)/\gamma \text{ voor } i \geq 1,$$

$$q(1|0,a) = \lambda/\gamma,\ q(0|0,a) = \mu_M/\gamma,\ F(t|i,a,j) = 1-e^{-\gamma t},\ d_1(i,a) = 0,$$
$$d_2(i,a) = h(i)+r_a.$$

Zoals reeds gesteld in paragraaf 2.2 voor dit gewijzigde probleem gelden ook de aannames 1 - 3 uit hoofdstuk 1. Tevens geldt dat een stationaire politiek die gemiddeld kosten optimaal is voor het gewijzigde probleem ook gemiddeld kosten optimaal is voor het oorspronkelijke probleem.

Wij zullen nu eerst het beslissingsmodel met een eindig aantal beslissingstijdstippen bestuderen. Gebruikmakend van

$$\beta_\alpha(i,a,j) = \gamma/(\alpha+\gamma) \quad \text{en} \quad \int_0^\infty \gamma e^{-\gamma t}\, dt \int_0^t e^{-\alpha y}\, dy = 1/(\alpha+\gamma)$$

vinden wij dat (1.13) overgaat in

$$\text{(2.4)} \qquad V_{n+1,\alpha}(i) = \frac{1}{\alpha+\gamma} \min_a [h(i)+r_a+\lambda V_{n,\alpha}(i+1)+\mu_a V_{n,\alpha}(i-1) + $$
$$+(\mu_M-\mu_a)V_{n,\alpha}(i)], \qquad i \geq 0 \text{ en } n \geq 0,$$

waarbij $V_{0,\alpha} \equiv 0$ en $V_{n,\alpha}(-1) = V_{n,\alpha}(0)$. Definiëren wij

$$\text{(2.5)} \qquad v_{n,\alpha}(i) = V_{n,\alpha}(i)-V_{n,\alpha}(i-1) \qquad \text{voor } i \geq 0 \text{ en } n \geq 0,$$

dan kunnen wij (2.4) voor $i \geq 0$ en $n \geq 0$ schrijven als

$$(2.6) \qquad V_{n+1,\alpha}(i) = \frac{1}{\alpha+\gamma}[h(i)+\lambda V_{n,\alpha}(i+1)+\mu_M V_{n,\alpha}(i)+\min_a\{r_a-\mu_a v_{n,\alpha}(i)\}].$$

Definieer nu $f_{n,\alpha}(i)$ als de kleinste waarde van a waarvoor het rechterlid van (2.4) minimaal is.

LEMMA 2.2. *Voor elke* $\alpha > 0$ *en* $n \geq 0$ *geldt dat de funktie* $V_{n,\alpha}(i)$ *convex in* i *is en geldt dat* $f_{n,\alpha}(i+1) \geq f_{n,\alpha}(i)$ *voor alle* $i \geq 0$.

Bewijs. Kies α vast. Wij zullen eerst met behulp van volledige induktie aantonen dat de funktie $V_{n,\alpha}(i)$ convex in i is voor alle $n \geq 0$. De funktie $V_{0,\alpha}$ is identiek gelijk aan nul en derhalve convex. Stel nu dat $V_{m,\alpha}(i)$ convex is voor een $m \geq 0$. Om te bewijzen dat $V_{m+1,\alpha}(i)$ convex in i is, moeten wij aantonen dat (zie (2.5))

$$(2.7) \qquad v_{m+1,\alpha}(i+1) \geq v_{m+1,\alpha}(i) \qquad\qquad \text{voor alle } i \geq 1.$$

Schrijven wij ter afkorting $f_{m+1,\alpha}(i) = a(i)$, dan volgt uit (2.6) en de definitie van a(i) dat voor alle $i \geq 1$ geldt

$$(\alpha+\gamma)v_{m+1,\alpha}(i+1) \geq h(i+1)+\lambda V_{m,\alpha}(i+2)+\mu_M V_{m,\alpha}(i+1)+r_{a(i+1)}-\mu_{a(i+1)}v_{m,\alpha}(i+1) +$$
$$- \{h(i)+\lambda V_{m,\alpha}(i+1)+\mu_M V_{m,\alpha}(i)+r_{a(i+1)}-\mu_{a(i+1)}v_{m,\alpha}(i)\},$$

en

$$(\alpha+\gamma)v_{m+1,\alpha}(i) \leq h(i)+\lambda V_{m,\alpha}(i+1)+\mu_M V_{m,\alpha}(i)+r_{a(i-1)}-\mu_{a(i-1)}v_{m,\alpha}(i) +$$
$$- \{h(i-1)+\lambda V_{m,\alpha}(i)+\mu_M V_{m,\alpha}(i-1)+r_{a(i-1)}-\mu_{a(i-1)}v_{m,\alpha}(i-1)\}.$$

Uit deze twee ongelijkheden volgt

$$(\alpha+\gamma)\{v_{m+1,\alpha}(i+1)-v_{m+1,\alpha}(i)\} \geq h(i+1)-h(i) - \{h(i)-h(i-1)\} +$$
$$+ \lambda\{v_{m,\alpha}(i+2)-v_{m,\alpha}(i+1)\} + \{\mu_M-\mu_{a(i+1)}\}\{v_{m,\alpha}(i+1)-v_{m,\alpha}(i)\}$$
$$+ \mu_{a(i-1)}\{v_{m,\alpha}(i)-v_{m,\alpha}(i-1)\} \qquad\qquad \text{voor } i \geq 1.$$

Het rechterlid van deze ongelijkheid is niet-negatief aangezien de funkties $h(i)$ en $V_{m,\alpha}(i)$ convex zijn en $\mu_M \geq \mu_a > 0$ voor alle a. Derhalve geldt (2.7) waarmee dan bewezen is dat de funktie $V_{n,\alpha}$ convex is voor alle n. De tweede bewering uit het lemma volgt nu eenvoudig. Kies n en i vast, en schrijf $f_{n,\alpha}(i) = k$. Aangezien k de kleinste waarde van a is waarvoor het rechterlid van (2.4) minimaal is, volgt uit (2.6) dat

$$r_a - \mu_a v_{n,\alpha}(i) > r_k - \mu_k v_{n,\alpha}(i) \qquad \text{voor } a < k.$$

Op grond van de convexiteit van $V_{n,\alpha}$ geldt dat $v_{n,\alpha}(i+1) \geq v_{n,\alpha}(i)$. Voorts geldt $\mu_k - \mu_a > 0$ voor $a < k$. Derhalve vinden wij

$$r_a - r_k + (\mu_k - \mu_a) v_{n,\alpha}(i+1) \geq r_a - r_k + (\mu_k - \mu_a) v_{n,\alpha}(i) > 0 \qquad \text{voor } a < k,$$

oftewel, $r_a - \mu_a v_{n,\alpha}(i+1) > r_k - \mu_k v_{n,\alpha}(i+1)$. Hieruit volgt $f_{n,\alpha}(i+1) \geq k$, waarmee het lemma bewezen is. □

In het volgende lemma zullen wij aantonen dat als i en n voldoende groot zijn en α voldoende klein is de politiek $f_{n,\alpha}$ in toestand i de snelste bedieningssnelheid voorschrijft.

LEMMA 2.3. *Er bestaan een getal* $\alpha^* > 0$ *en natuurlijke getallen* i^* *en* N^* *zodat* $f_{n,\alpha}(i) = M$ *voor alle* $0 < \alpha < \alpha^*$, $i \geq i^*$ *en* $n \geq N^*$.

Bewijs. Wij tonen eerst aan dat

$$(2.8) \qquad v_{n+1,\alpha}(i+1) \geq \frac{h}{\alpha+\gamma} \sum_{j=0}^{n \wedge i} \left(\frac{\gamma}{\alpha+\gamma}\right)^j \qquad \text{voor alle } \alpha > 0,\ i \geq 0 \text{ en } n \geq 0,$$

waarbij $n \wedge i = \min(n,i)$ en h een positieve constante is met $h(i+1)-h(i) \geq h$ voor alle i. Uit (2.8) volgt dat voor elke $\alpha_0 > 0$ wij een i_0 en een N_0 kunnen vinden zodat $v_{n,\alpha}(i) \geq h/2\alpha_0$ voor alle $0 < \alpha < \alpha_0$, $i \geq i_0$ en $n \geq N_0$. Om (2.8) te bewijzen, kies α, i en n vast, en stel $f_{n,\alpha}(i+1) = k$. Gebruikmakend van de convexiteit van $V_{n,\alpha}$ en het feit dat $\mu_M \geq \mu_k$ volgt uit (2.3),

$$v_{n+1,\alpha}(i+1) \geq \frac{1}{\alpha+\gamma}\,[h(i+1)-h(i)+\lambda v_{n,\alpha}(i+2)+\mu_M v_{n,\alpha}(i+1) +$$

$$- \mu_k\{v_{n,\alpha}(i+1)-v_{n,\alpha}(i)\}] \geq \frac{1}{\alpha+\gamma}[h+\lambda v_{n,\alpha}(i)+(\mu_M-\mu_k)v_{n,\alpha}(i) +$$

$$+\mu_k v_{n,\alpha}(i)] = \frac{1}{\alpha+\gamma}[h+\gamma v_{n,\alpha}(i)].$$

Door deze gelijkheid herhaald toe te passen, vinden wij (2.8).

Vervolgens definieer voor elke vaste $n \geq 1$ de volgende politiek $R_{n,\alpha}$ voor het $(n+1)$-stapsbeslissingsmodel waarin alleen beslissingen worden genomen op de eerste $n+1$ beslissingstijdstippen. Op het eerste beslissingstijdstip $\underline{T}_0 = 0$ neemt politiek $R_{n,\alpha}$ in elke toestand de aktie M en op het $(k+1)^{ste}$ beslissingstijdstip neemt politiek $R_{n,\alpha}$ de beslissing $f_{k+1,\alpha}(i)$ in toestand i voor $k=1,\ldots,n$. Definieer $W_{n+1,\alpha}(i)$ als de totale verwachte verdisconteerde kosten tot het $(n+2)^{de}$ beslissingstijdstip wanneer de begintoestand i is en politiek $R_{n,\alpha}$ wordt toegepast. Dan geldt

$$W_{n+1,\alpha}(i) = \frac{1}{\alpha+\gamma}\{h(i)+r_M+\lambda V_{n,\alpha}(i+1)+\mu_M V_{n,\alpha}(i-1)+(\mu_M-\mu_M)V_{n,\alpha}(i)\}.$$

Uit (2.1) volgt dat

$$V_{n+1,\alpha}(i) = \frac{1}{\alpha+\gamma}\{h(i)+r_{f_{n,\alpha}(i)}+\lambda V_{n,\alpha}(i+1)+\mu_{f_{n,\alpha}(i)}V_{n,\alpha}(i-1) +$$

$$+(\mu_M-\mu_{f_{n,\alpha}(i)})V_{n,\alpha}(i)\}.$$

Derhalve geldt

$$V_{n+1,\alpha}(i)-W_{n+1,\alpha}(i) = \frac{1}{\alpha+\gamma}\{r_{f_{n,\alpha}(i)}-r_M+(\mu_M-\mu_{f_{n,\alpha}(i)})v_{n,\alpha}(i)\}.$$

Uit deze gelijkheid, relatie (2.8) en het feit dat α_0 zo klein gekozen kan worden dat $(\mu_M-\mu_a)h/2\alpha_0 > r_a-r_M$ voor alle $a \neq M$ volgt nu het lemma, aangezien altijd geldt dat $V_{n+1,\alpha}(i) \leq W_{n+1,\alpha}(i)$ is, zie opmerking 1.1 in paragraaf 1.3. □

STELLING 2.1. *Er is een gemiddeld kosten optimale stationaire politiek* f^* *met de eigenschap dat* $f(i+1) \geq f(i)$ *voor alle* $i \geq 0$ *en* $f(i) = M$ *voor alle* i *voldoende groot.*

Bewijs. Kies α^*, i^* en N^* zoals in lemma 2.3. Uit de lemma's 1.1 en 1.2 en stelling 1.1(c) in paragraaf 1.3 volgt dat voor elke α met $0 < \alpha < \alpha^*$ een α-optimale stationaire politiek g_α bestaat met de eigenschap dat $g_\alpha(i+1) \geq$ $\geq g_\alpha(i)$ voor alle $i \geq 0$ en $g_\alpha(i) = M$ voor alle $i \geq i^*$. Aangezien het aantal mogelijke akties eindig is en $g_\alpha(i) = M$ voor alle $0 < \alpha < \alpha^*$ en $i \geq i^*$, volgt dat een stationaire politiek f^* en een rij $\{\alpha_k, k\geq 1\}$ met $\alpha_k \to 0$ als $k \to \infty$ bestaan, zodat $g_{\alpha_{n_k}} = f^*$ voor alle $k \geq 1$. Derhalve is aanname 4(a) uit paragraaf 1.4 vervuld. Op grond van het feit dat $f^*(i) = M$ voor alle $i \geq i^*$ en de aanname $\lambda/\mu_M < 1$ volgt nu eenvoudig dat ook aanname 4(b) vervuld is. De stelling volgt nu uit stelling 1.2(a). □

2.4. *Een wachttijdsysteem met transport tussen twee stations.*

Beschouw een wachttijdsysteem met twee stations 1 en 2. Bij station 1 komen klanten aan volgens een Poisson proces met parameter λ_1, en onafhankelijk daarvan komen bij station 2 klanten aan volgens een Poisson proces met parameter λ_2. Elke klant blijft wachten totdat hij wordt getransporteerd van het station waar hij aankwam naar het andere station. Voor het transport van de klanten is één transportwagen aanwezig. Deze wagen rijdt tussen de stations op en neer en heeft een onbeperkte capaciteit. Als de wagen bij station 2 aankomt, dan neemt de wagen direkt alle daar eventueel aanwezige klanten mee en vertrekt naar station 1. In tegenstelling tot station 2 kan de wagen bij station 1 blijven wachten. Als de wagen bij station 1 aankomt of als een klant aankomt terwijl de wagen bij station 1 staat te wachten, dan moet beslist worden of de wagen blijft staan bij station 1 of vertrekt naar station 2 met meeneming van alle klanten aanwezig bij station 1. Op elk moment is bekend hoeveel klanten bij elk van de stations aanwezig is. Zowel de tijd nodig om van station 1 naar station 2 te rijden als de tijd nodig om van station 2 naar station 1 te rijden is een constante $T \geq 0$. De volgende kosten worden beschouwd. Voor elke klant bij station 1 (2) zijn er wachttijdkosten $h_1 > 0$ ($h_2 > 0$) per tijdseenheid dat de klant op transport wacht. De kosten van een rit van station 1 naar station 2 en terug zijn gelijk aan $K > 0$.

Wij zullen aantonen dat voor dit probleem een gemiddeld kosten optimale stationaire politiek bestaat van de volgende vorm: de wagen vertrekt van station 1 als een lineaire combinatie van het aantal klanten aanwezig bij station 1 en het aantal klanten aanwezig bij station 2 een bepaalde waarde overschrijdt.

Dit probleem is een semi-Markov beslissingsmodel waarvan de toestand wordt waargenomen op de momenten dat de wagen bij station 1 aankomt en de momenten dat een klant bij één van de stations aankomt terwijl de wagen bij station 1 staat. De toestand kan worden gegeven door een paar (n_1,n_2) waarbij n_k het aantal klanten aangeeft dat bij station $k=1,2$ aanwezig is. Voor elke toestand zijn er twee mogelijke akties g en w, waarbij aktie g voorschrijft naar station 2 te gaan en aktie w voorschrijft te blijven wachten. Het is eenvoudig na te gaan dat

$$q((n_1+1,n_2)\,|\,(n_1,n_2),w) = \lambda_1/(\lambda_1+\lambda_2) = 1-q((n_1,n_2+1)\,|\,(n_1,n_2),w),$$

$$q((a_1,a_2)\,|\,(n_1,n_2),g) = e^{-2\lambda_1 T}\frac{(2\lambda_1 T)^{a_1}}{a_1!}\; e^{-\lambda_2 T}\frac{(\lambda_2 T)^{a_2}}{a_2!},$$

$$F(t\,|\,(n_1,n_2),w,(n_1+1,n_2)) = F(t\,|\,(n_1,n_2),w,(n_1,n_2+1)) = 1-e^{-(\lambda_1+\lambda_2)t},$$

$$F(t\,|\,(n_1,n_2),w,(a_1,a_2)) = \begin{cases} 1 & \text{voor } t \geq 2T, \\ 0 & \text{voor } t < 2T, \end{cases}$$

$$d_1((n_1,n_2),w) = 0, \quad d_2((n_1,n_2),w) = h_1n_1+h_2n_2,$$

$$d_1((n_1,n_2),g) = K+h_2n_2T+h_1\lambda_1(2T)^2/2+h_2\lambda_2T^2/2. \quad ^{*)}$$

Bij bovenstaande keuze van de toestandsruimte zijn echter de aannames 2 en 3 uit paragraaf 1.3 niet vervuld. Dit kan echter in orde gemaakt worden door de toestandsruimte te herdefinieren. Vervang elke toestand (n_1,n_2) door de toestand (i,ξ), waarbij $i = n_1$ en $\xi = (n_2,1)$ als $n_1 \geq n_2$, en $i = n_2$ en $\xi = (n_1,2)$ als $n_2 > n_1$. Men gaat nu direkt na dat aanname 2 vervuld is met $K = 1$ en dat aanname 3 vervuld is met $d = \max(1,2\lambda_1T+\lambda_2T)$. Aanname 1 is

---

*) Gebruik de volgende eigenschap van het Poisson proces (zie ROSS (1970)): gegeven dat in $(0,t)$ n aankomsten plaatsvinden, dan heeft elk van de n aankomsttijdstippen als verwachting $t/2$.

triviaal vervuld. Derhalve geldt voor elke $\alpha > 0$ dat een α-optimale stationaire politiek f_α bestaat, zie stelling 1.1.

Om voorwaarde 4 te verifiëren, zullen wij aantonen dat een $\alpha^* > 0$ en een natuurlijk getal M bestaan zodat

$$(2.9) \qquad f_\alpha(i,\xi) = g \text{ voor alle } (i,\xi) \text{ met } n_1+n_2 \geq M \text{ en alle } 0 < \alpha < \alpha^*.$$

Veronderstel dat (2.9) niet juist is. Dan bestaan een rij $\{\alpha_k\}$ met $\alpha_k \to 0$ als $k \to \infty$ en een rij $\{i(k),\xi(k)\}$ met $n_1(k)+n_2(k) \to \infty$ als $k \to \infty$ zodat

$$f_{\alpha_k}((i(k),\xi(k))) = w \qquad \text{voor alle } k.$$

Beschouw nu voor vaste k de volgende politiek π_{α_k}. Deze politiek laat de wagen continu rijden tot het tijdstip $\underline{t}_k$ waarbij $\underline{t}_k$ het eerste tijdstip is waarop de wagen bij aankomst bij station 1 de toestand (0,0,1) aantreft terwijl in $[\underline{t}_k-4T,\underline{t}_k]$ geen klanten bij één van de stations zijn aangekomen. Vanaf tijdstip $\underline{t}_k$ gaat politiek π_{α_k} te werk als politiek g_{α_k}. Aangezien bewezen kan worden dat politiek g_{α_k} de aktie wachten voorschrijft als geen klanten in het systeem zijn, geldt dat onder politiek g_{α_k} op tijdstip $\underline{t}_k$ de wagen stilstaat bij station 1. Derhalve zijn vanaf tijdstip $\underline{t}_k$ de kosten onder politiek g_{α_k} tenminste zo groot als onder politiek π_{α_k}. Vervolgens tonen wij aan dat voor k voldoende groot het verschil in kosten tot tijdstip $\underline{t}_k$ tussen politiek π_{α_k} en g_{α_k} negatief is. Dit geeft een tegenspraak met het feit dat politiek g_{α_k} α_k-optimaal is, waarmee dan (2.5) bewezen is. Als p de kans is dat in een tijd 4T geen klanten aankomen, dan is onder politiek π_{α_k} het verwachte aantal ritten tot tijdstip $\underline{t}_k$ begrensd door 2/p. Derhalve is het verschil in verwachte verdisconteerde kosten tot tijdstip $\underline{t}_k$ tussen politiek π_{α_k} en g_{α_k} bij begintoestand $((i(k),\xi(k))$ kleiner dan of gelijk aan

$$2K/p-\{h_1n_1(k)+h_2n_2(k)\}/(\alpha_k+\lambda_1+\lambda_2).$$

Deze grootheid is kleiner dan nul als k voldoende groot is, waarmee de tegenspraak gevonden is.

Uit (2.9) en het feit dat er slechts eindig veel toestanden (i,ξ) zijn met $n_1+n_2 < M$ volgt dat een rij $\{\alpha_k\}$ met $\alpha_k \to 0$ als $k \to \infty$ en een stationaire politiek f^* met $f^*((i,\xi)) = g$ als $n_1+n_2 \geq M$ bestaan zodat f^* α_k-optimaal is voor alle k. Derhalve is aanname 4 vervuld (kies $s^* = (0,0,1)$). Uit stelling 1.2 volgt nu dat politiek f^* gemiddeld kosten optimaal is.

Om de structuur van een gemiddeld kosten optimale politiek te bewijzen, beschouwen wij de klasse C^* van stationaire politieken f met $f((i,\xi)) = g$ voor alle (i,ξ) met $n_1+n_2 \geq N$, waarbij N zo groot gekozen wordt dat C^* een gemiddeld kosten optimale politiek bevat en dat $h_1n_1+h_2n_2 < g^*-h_2\lambda_2T$ impliceert $n_1+n_2 < N$. Hierbij is g^* de minimale gemiddelde kosten. Wij merken nu op dat op elk tijdstip waarop de wagen bij station 1 terugkeert er $\underline{a}_1$ klanten bij station 1 en $\underline{a}_2$ klanten bij station 2 staan te wachten, waarbij $\underline{a}_1$ en $\underline{a}_2$ Poisson verdeeld zijn met als verwachting $2\lambda_1T$ resp. λ_2T. Op grond van een standaard resultaat uit de theorie van de regeneratieve processen geldt voor elke politiek $f \in C^*$ dat de gemiddelde kosten g(f) gelijk zijn aan (zie bijv. stelling 3.16 in ROSS (1970)),

$$g(f) = K(f)/T(f),$$

waarbij K(f) de verwachting is van de kosten te maken onder politiek f gedurende het tijdsinterval tussen twee opeenvolgende tijdstippen waarop de wagen bij station 1 terugkomt en T(f) de verwachte lengte van dit tijdsinterval is. Aangezien C^* een gemiddeld kosten optimale politiek bevat, geldt

$$g^* = \min_{f\in C^*} K(f)/T(f).$$

Uit lemma 2.1 in paragraaf 2.2 volgt dat een politiek $f^* \in C$ gemiddeld kosten optimaal is wanneer f^* de grootheid

$$K(f) - g^*T(f)$$

onder $f \in C^*$ minimaliseert. Het probleem van de minimalisatie van deze grootheid kan vertaald worden in een optimaal stopprobleem. Om dit te doen, gaan wij uit van de situatie dat op tijdstip 0 de wagen bij station 1 terugkomt terwijl $\underline{a}_1$ klanten bij station 1 en $\underline{a}_2$ klanten bij station 2 staan te wachten. Definieer $\underline{\tau}_n$ als het tijdstip van aankomst van de n^{de} na tijdstip 0 aankomende klant, en definieer $\underline{\eta}_n = 1$ als deze klant bij station 1 aankomt en definieer $\underline{\eta}_n = 0$ als deze klant bij station 2 aankomt. Voor n=0,1,... definieer de stochastische vektor $\underline{y}_n$ door

$$\underline{y}_n = (\underline{y}_{n1}, \underline{y}_{n2}, \underline{y}_{n3})$$

waarbij $\underline{y}_{01} = \underline{a}_1$, $\underline{y}_{02} = \underline{a}_2$, $\underline{y}_{03} = 0$, en

$$\underline{y}_{n1} = \underline{y}_{n-1,1} + \underline{\eta}_n,\ \underline{y}_{n2} = \underline{y}_{n-1,2} + 1 - \underline{\eta}_n \text{ en } \underline{y}_{n3} = \underline{\tau}_n \qquad \text{voor } n \geq 1.$$

Derhalve het proces $\{\underline{y}_n\}$ geeft het verloop van de aankomsten van de klanten bij de stations. Associeer nu met elke politiek $f \in C^*$ de volgende stopregel R_f met betrekking tot het proces $\{\underline{y}_n\}$: De regel R_f schrijft na achtereenvolgens de waarden $y_0,\dots,y_k$ te hebben waargenomen stoppen voor op het k^{de} waarnemingstijdstip dan en slechts dan als politiek f de aktie gaan voorschrijft in de toestand waarin y_{k1} de klanten bij station 1 en y_{k2} klanten bij station 2 aanwezig zijn. Als op het k^{de} waarnemingstijdstip na het waarnemen van de waarden $(y_0,\dots,y_k)$ gestopt wordt, dan kennen wij hieraan de volgende stopkosten toe:

$$c(y_0,\dots,y_k) = \sum_{j=0}^{k-1} (h_1 y_{j1} + h_2 y_{j2})(y_{j+1,3} - y_{j,3}) + K + h_2 y_{k2} T +$$
$$+ h_1\lambda_1 (2T)^2/2 + h_2\lambda_2 T^2/2 - g^*(y_{k3} + 2T),$$

m.a.w. $c(y_0,\dots,y_k)$ is de som van de wachttijdkosten tot tijdstip τ_k plus de verwachte kosten van tijdstip τ_k tot het eerstvolgende tijdstip waarop de wagen bij station 1 terugkomt minus g^* maal de som van τ_k plus de tijd van τ_k tot het eerstvolgende tijdstip waarop de wagen bij station 1 terugkomt. Het zal duidelijk zijn dat de verwachte stopkosten onder stopregel R_f gelijk zijn aan

$$K_f - g^* T_f.$$

De structuur van een gemiddeld optimale stationaire politiek is derhalve gelijk aan de structuur van een optimale stopregel. Uit de keuze van

$c(y_0,\dots,y_k)$ volgt dat

$$Ec(y_0,\dots,y_k,\underline{y}_{k+1})-c(y_0,\dots,y_k) = (h_1y_{k1}+h_2y_{k2})/\lambda+h_2\lambda_2T/\lambda-g^*/\lambda.$$

Aangezien $\underline{y}_{k+2,1} \geq \underline{y}_{k+1,1}$ en $\underline{y}_{k+2,2} \geq \underline{y}_{k+1,2}$ volgt nu dat de monotoniciteitseigenschap (2.3) uit paragraaf 2.1.2 vervuld is. Een optimale stopregel heeft dus de volgende vorm: stop zodra $h_1y_{k1}+h_2y_{k2} \geq g^*-h_2\lambda_2T$. Vertaald betekent dit dat een gemiddeld optimale stationaire politiek van de volgende vorm bestaat: laat de wagen van station 1 vertrekken zodra

$$h_1n_1+h_2n_2 \geq g^*-h_2\lambda_2T$$

waarbij n_j het aantal klanten aanwezig bij station $j=1,2$ aangeeft.

APPENDIX

In de appendix geven wij een aantal algemene stellingen.

De vaste punt stelling.

Beschouw een willekeurige verzameling V. Stel dat er een funktie ρ is die aan elk tweetal punten $u,v \in V$ een reëel getal $\rho(u,v)$ toevoegt zodat
(a) $\rho(u,v) = 0$ dan en slechts dan als $u = v$,
(b) $\rho(u,v) = \rho(v,u)$ voor alle $u,v \in V$,
(c) $\rho(u,w) \le \rho(u,v)+\rho(v,w)$ voor alle $u,v,w \in V$.
Dan heet V een metrische ruimte met metriek ρ. De verzameling V heet een compleet metrische ruimte als voor elke rij $\{v_n, n\ge 1\}$ van punten uit V met de eigenschap dat voor elke $\varepsilon > 0$ er een $n_0(\varepsilon)$ is zodat $\rho(v_n, v_{n+m}) < \varepsilon$ voor alle $n \ge n_0(\varepsilon)$ en $m \ge 1$, geldt dat er een punt $v \in V$ is met $\rho(v_n, v) \to 0$ als $n \to \infty$.

Laat A een afbeelding zijn die aan elk punt $v \in V$ een punt $Av \in V$ toevoegt. De afbeelding A heet continu in $v_0 \in V$ als voor elke $\varepsilon > 0$ er een $\delta > 0$ zodat $\rho(Av, Av_0) < \varepsilon$ voor alle v met $\rho(v,v_0) < \delta$. De afbeelding A heet continu als deze continu is in elk punt $v_0 \in V$. De afbeelding A^n voor $n=1,2,\ldots$ wordt gedefinieerd door $A^n v = A(A^{n-1}v)$ voor $n \ge 2$ waarbij $A^1 v = Av$.

STELLING 1. *Laat* V *een compleet metrische ruimte met metriek* ρ *zijn, en laat* A *een continue afbeelding van* V *in zichzelf zijn. Veronderstel dat er een natuurlijk getal* N *en een getal* β *met* $0 \le \beta < 1$ *bestaan zodat**)

$$\rho(A^N u, A^N v) \le \beta\rho(u,v) \qquad \textit{voor alle } u,v \in V.$$

Dan is er een uniek punt $v^* \in V$ *met*

$$Av^* = v^*$$

---

*) De afbeelding A^N heet een contractie afbeelding.

Verder geldt dat

$$\lim_{n\to\infty} \rho(A^n v_0, v^*) = 0 \qquad \textit{voor alle } v_0 \in V.$$

Voor een bewijs van deze stelling verwijzen wij naar blz. 50 in KOLMOGOROV & FOMIN (1957).

Een Abel stelling.

STELLING 2. *Laat* $\phi(t)$ *een niet-negatieve, niet-dalende funktie van* $t \geq 0$ *zijn. Veronderstel dat* $\int_0^\infty e^{-\alpha t} d\phi(t)$ *convergeert voor alle reële* $\alpha > 0$. *Dan geldt*

$$\limsup_{t\to\infty} t^{-1}\phi(t) \geq \limsup_{\alpha\downarrow 0} \alpha\int_0^\infty e^{-\alpha t} d\phi(t).$$

Als $\lim_{t\to\infty} t^{-1}\phi(t)$ *bestaat en eindig is, dan geldt*

$$\lim_{t\to\infty} t^{-1}\phi(t) = \lim_{\alpha\downarrow 0} \alpha\int_0^\infty e^{-\alpha t} d\phi(t).$$

Voor een bewijs van deze stelling verwijzen wij naar blz. 181-182 in WIDDER (1946).

Een selektie stelling.

STELLING 3. *Laat* I *een aftelbare of eindige verzameling zijn waarop een rij* $f_1, f_2, \ldots$ *van reëelwaardige funkties is gedefinieerd. Veronderstel dat voor elke vaste* $i \in I$ *de rij* $\{f_n(i), n\geq 1\}$ *begrensd is. Dan is er een eindige funktie* f *op* I *en een deelrij* $\{n_k, k\geq 1\}$ *van de natuurlijke getallen met* $n_k \to \infty$, *zo dat*

$$\lim_{k\to\infty} f_{n_k}(i) = f(i) \qquad \textit{voor alle } i \in I.$$

Voor een bewijs van deze stelling, zie bijvoorbeeld blz. 179 in ROYDEN (1968).

LITERATUUR

ADIRI, I. & U. YECHIALI (1973), *Optimal priority purchasing and pricing decisions in I. non-monopoly and II. monopoly queues*, Operations Research, Statistics and Economics Mimeograph Series No. 137, Technion-Israel Institute of Technology.

BELL, C. (1971), *Characterization and computation of optimal policies for operating an* M/G/1 *queuing system with removable server*, Operations Res. 19, 208-218.

BELL, C. (1973a), *Efficient operation of optional priority queueing systems*, Operations Res. 21, 777-786.

BELL, C. (1973b), *Optimal average-cost operating policy for an* M/G/1 *queuing system with removable server and several priority classes*, Operations Res. 21, 1281-1290.

BLACKBURN, J. (1971), *Optimal control of queueing systems with intermittent service*, Technical Report No. 23, Department of Operations Research, Stanford University.

BLACKBURN, J. (1972), *Optimal control of a single-server queue with balking and regening*, Management Sci. 19, 297-313.

CHOW, Y.S., ROBBINS, H. & D. SIEGMUND (1971), *Great Expectations: the theory of optimal stopping*, Houghton Mifflin Company, Boston.

CHUNG, K.L. (1960), *Markov chains with stationary transition probabilities*, Springer-Verlag, Berlin.

CRABILL, T.B. (1972), *Optimal control of a service facility with variable exponential service times and constant arrival rate*, Management Sci. 18, 560-566.

CRABILL, T.B. (1973), *Optimal hysteretic control of a stochastic service system with variable service rates and fixed switch-over costs*, University of North Carolina.

CRABILL, T.B. (1974), *Optimal control of a maintenance system with variable service rates*, Operations Res. 22, 736-745.

CRAMER, M. (1971), *Optimal customer selection in exponential queues*, ORC 71-24, Operations Research Center, University of California, Berkeley.

DEB, R.K. & R.F. SERFOZO (1973), *Optimal control of bath service queues*, Advances in Appl. Prob. 5, 340-361.

DERMAN, C. (1970), *Finite state Markovian decision processes*, Academic Press, New York.

HARRISON, M. (1970), *Countable state discounted Markovian decision processes*, Technical Report No. 17, Department of Operations Research, Stanford University.

HARRISON, M. (1972a), *Discrete dynamic programming with unbounded rewards*, Annals of Math. Statist. 43, 636-644.

HARRISON, M. (1972b), *Dynamic scheduling of a multi-class queue I & II*, Technical Reports Nos. 36 and 37, Department of Operations Research, Stanford University.

HEYMAN, D.P. (1968), *Optimal operating policies for* M/G/1 *queueing systems*, Operations Res. 16, 362-382.

HORDIJK, A. (1974), *Dynamic programming and Markov potential theory*, Mathematical Centre Tract No. 51, Mathematisch Centrum, Amsterdam.

IGNALL, E. & P. KOLESAR (1974), *Optimal dispatching of an infinite capacity shuttle: control at a single terminal*, Operations Res. 22, 1008-1024.

KEILSON, J. & D.M.G. WISHART (1964), *A central limit theorem for processes defined on a finite Markov chain*, Proceedings Cambridge Phil. Soc. 60, 547-567.

KOLMOGOROV, A.N. & S.V. FOMIN (1957), *Elements of the theory of functions and functional analysis*, Vol. 1, Graylock.

LIPPMAN, S.A. & S.M. ROSS (1971), *The streetwalkers dilemma: a job shop model*, Siam Journal of Appl. Math. 20, 336-344.

LIPPMAN, S.A. (1973a), *Semi-Markov decision processes with unbounded rewards*, Management Sci. 19, 717-731.

LIPPMAN, S.A. (1973b), *On dynamic programming with unbounded rewards*, Working Paper No. 212, Western Management Science Institute, University of California, Los Angeles.

LIPPMAN, S.A. (1973c), *A new technique in the optimization of exponential queueing systems*, Working Paper No. 211, Western Management Science Institute, University of California, Los Angeles.

LOW, D. (1971), *Optimal dynamic operating policies for an* M/M/s *queue with variable arrival rate*, Working Paper No. 178, Western Management Science Institute, University of California, Los Angeles.

LOW, D. (1974), *Optimal dynamic pricing policies for an* M/M/s *queue*, Operations Res. 22, 545-561.

MILLER, B.L. (1969), *A queueing reward system with several customer classes*, Management Sci. 16, 234-245.

MITCHELL, B. (1973), *Optimal service-rate selection in an* M/G/1 *queue*, Siam Journal of Appl. Math. 24, 19-35.

NAOR, P. (1969), *On the regulation of queue size by levying tolls*, Econometrica 37, 15-24.

PRABHU, N.U. & S. STIDHAM Jr. (1974), *Optimal control of queueing systems*, in: Mathematical Methods in Queueing Theory, Lecture Notes in Economics and Mathematical Systems, No. 98, Springer-Verlag, Berlin.

ROSS, S.M. (1970), *Applied probability models with optimization applications*, Holden-Day, Inc., San Francisco.

ROYDEN, H.L. (1968), *Real Analysis (2nd. ed.)*, MacMillan, New York.

SOBEL, M. (1969), *Optimal average-cost policy for a queue with start-up and shut-down costs*, Operations Res. 17, 145-162.

THATCHER, R.M. (1968), *Optimal single-channel service policies for stochastic arrivals*, ORC 68-16, Operations Research Center, University of California, Berkeley.

TIJMS, H.C. (1974), *Regeneratieve processen met toepassingen in wachttijdtheorie*, Rapport BC 11/74, Mathematisch Centrum, Amsterdam.

WIDDER, D.V. (1946), *The Laplace transform,* Princeton University Press.

YECHIALI, U. (1971), *On optimal balking rules and toll charges in a* GI/M/1 *queuing process,* Operations Res. 19, 349-370.

YECHIALI, U. (1972), *Customers' optimal joining rules for the* GI/M/s *queue,* Management Sci. 18, 434-443.

ZACKS, S. & M. YADIN (1970), *Analytic characterization of the optimal control of a queueing system,* Journal of Appl. Prob. 7, 617-633.