MATHEMATISCH CENTRUM
2e BOERHAAVESTRAAT 49
AMSTERDAM

STATISTISCHE AFDELING

Leiding: Prof. Dr D. van Dantzig
Chef van de Statistische Consultatie: Prof. Dr J. Hemelrijk

Rapport S 156(C7)

Syllabus van het college Markof-ketens 1953-'54 van Prof.Dr D. van Dantzig bewerkt door G. Zoutendijk met medewerking van J. Hirsch.

October 1954-Juni 1956

The Mathematical Centre at Amsterdam, founded the 11th of February 1946, is a non-profit institution aiming at the promotion of pure mathematics and its applications, and is sponsored by the Netherlands Government through the Netherlands Organization for Pure Research (Z.W.O.) and the Central National Council for Applied Scientific Research in the Netherlands (T.N.O.), by the Municipality of Amsterdam and by several industries.

Inhoudsopgave

Deel I: Markof-ketens

# 1. Inleiding

## 1.1. Notatie en afkortingen

Wij zullen geregeld de volgende afkortingen gebruiken:

wh(n) = waarschijnlijkheid (heden),
whr = waarschijnlijkheidsrekening,
el(n) = element(en),
vz(n) = verzameling(en),
fct(s) = functie(s),
o.o. = onderling onafhankelijk,
(o)l.o. = (onderling) lineair onafhankelijk.

De vz van alle natuurlijke getallen geven wij aan met $N$ als het getal 0 hierbij inbegrepen is; is dit niet het geval, dan schrijven wij $N'$. Verder zal $R$ de vz van alle reële getallen, $\mathcal{I}$ die van alle gehele getallen, $P$ die van alle positieve getallen zijn.

De vz van de getallen $1,2,3,\dots,n$ zullen wij aangeven met $\{1,2,\dots,n\}$ of korter met $N'_n$. Geregeld zullen wij gebruik maken van de tensornotatie van Einstein. Hierbij is $a_i b^i$ of $a^i b_i$ een andere schrijfwijze voor $\Sigma^i a_i b_i$. Het sommatieteken kan bij deze notatie weggelaten worden. Wij sommeren steeds stilzwijgend over die en slechts die indices, die precies éénmaal als boven- en éénmaal als benedenindex voorkomen, dus bijvoorbeeld nooit over indices, die slechts als boven- of benedenindex of tweemaal als boven- en eenmaal als benedenindex voorkomen. Een nadeel van deze notatie is, dat bovenindices niet van exponenten onderscheiden kunnen worden. Immers $a^i$ kan zowel de $i$-de component aanduiden van een vector $a$, als ook de $i$-de macht van een getal $a$. Men zij op deze mogelijkheid tot verwarring bedacht. Waar in de tekst verwarring dreigt, zullen wij deze door het gebruik van haakjes of alleen maar door een waarschuwend woord trachten te voorkomen.

De formules worden aangegeven door cijfers tussen ronde haken. Zo betekent (2.3.4) de vierde formule van hoofdstuk 2, paragraaf 3. Cijfers tussen vierkante haken verwijzen naar de literatuurlijst achterin.

## 1.2. Praedicaten en verzamelingen

Om dit collegedictaat te kunnen lezen is enige wiskundige voorkennis onontbeerlijk. Zo zal zonder verdere explicatie een enkele maal van de differentiaal- en integraalrekening gebruik gemaakt worden, terwijl ook de grondbeginselen van de vector- en matrixrekening bekend ondersteld worden. Wel zullen wij sommige elementaire begrippen alvorens ze te gebruiken nog definiëren om ze in de herinnering van de lezer terug te brengen. De eenvoudigste relaties tussen deze begrippen zullen dan zonder bewijs genoemd worden.

De grondbeginselen van de praedicatenrekening en de verzamelingsleer zullen in deze paragraaf nog genoemd worden. In de appendix, hoofdstuk 9, zullen de stellingen uit de theorie der aftelbaar additieve verzamelingsfuncties, voor zover wij ze nodig hebben, genoemd, doch veelal niet bewezen worden. Ook een zekere kennis van de elementaire waarschijnlijkheidsrekening kan niet ontbeerd worden. De enkele opmerkingen, die wij in de komende paragraaf over dit onderwerp maken, zijn dan ook alleen bedoeld ter oriëntatie, doch vormen zeker geen voldoende inleiding voor een lezer, die met de gedachtengang en grondbegrippen der whr onbekend is.

Wij beschouwen in deze paragraaf een vz van uitspraken (proposities), gekenmerkt door hun eigenschap "waar" of "niet waar" te kunnen zijn. Deze uitspraken zullen wij aangeven met sierletters. Met ieder paar uitspraken $\mathcal{A}$ en $\mathcal{B}$ onderstellen wij, dat ook de volgende uitspraken in de vz opgenomen zijn:

$\neg\mathcal{A}$ (non $\mathcal{A}$), de negatie van $\mathcal{A}$, de uitspraak, die waar is, als $\mathcal{A}$ niet waar is, en niet waar, als $\mathcal{A}$ waar is.

$\mathcal{A} \cap \mathcal{B}$ ($\mathcal{A}$ en $\mathcal{B}$), de conjunctie van $\mathcal{A}$ en $\mathcal{B}$, de uitspraak, die waar is, als $\mathcal{A}$ en $\mathcal{B}$ waar zijn, en anders niet waar.

$\mathcal{A} \cup \mathcal{B}$ ($\mathcal{A}$ of $\mathcal{B}$), de disjunctie van $\mathcal{A}$ en $\mathcal{B}$, de uitspraak, die waar is als van $\mathcal{A}$ en $\mathcal{B}$ minstens één van beide waar is en niet waar, als beide niet waar zijn.

$\mathcal{A} \rightarrow \mathcal{B}$ (als $\mathcal{A}$, dan $\mathcal{B}$), de uitspraak, die slechts dan niet waar is, als $\mathcal{A}$ waar is en $\mathcal{B}$ niet waar en die anders waar is.

$\mathcal{A} \equiv \mathcal{B}$ ( $\mathcal{A}$ dan en slechts dan als $\mathcal{B}$), de uitspraak, die waar is, als $\mathcal{A}$ en $\mathcal{B}$ beide waar of beide niet waar zijn en die anders onwaar is.

Nu gelden bijvoorbeeld de volgende relaties:

(1)

1. $\neg\neg\mathcal{A} \equiv \mathcal{A}$,
2. $\mathcal{A} \cap \mathcal{A} \equiv \mathcal{A}$ , idempotente eigenschap van $\cap$,
3. $\mathcal{A} \cap \mathcal{B} \equiv \mathcal{B} \cap \mathcal{A}$ , commutativiteit van $\cap$,
4. $(\mathcal{A} \cap \mathcal{B}) \cap \mathcal{C} \equiv \mathcal{A} \cap (\mathcal{B} \cap \mathcal{C})$ , associativiteit van $\cap$,
5. $\mathcal{A} \cup \mathcal{A} \equiv \mathcal{A}$, idempotente eigenschap van $\cup$,
6. $\mathcal{A} \cup \mathcal{B} \equiv \mathcal{B} \cup \mathcal{A}$ , commutativiteit van $\cup$,
7. $(\mathcal{A} \cup \mathcal{B}) \cup \mathcal{C} \equiv \mathcal{A} \cup (\mathcal{B} \cup \mathcal{C})$ , associativiteit van $\cup$,
8. $\mathcal{A} \cap (\mathcal{B} \cup \mathcal{C}) \equiv (\mathcal{A} \cap \mathcal{B}) \cup (\mathcal{A} \cap \mathcal{C})$ , distributiviteit van $\cap$ t.o.v. $\cup$,
9. $\mathcal{A} \cup (\mathcal{B} \cap \mathcal{C}) \equiv (\mathcal{A} \cup \mathcal{B}) \cap (\mathcal{A} \cup \mathcal{C})$ , distributiviteit van $\cup$ t.o.v. $\cap$,
10. $\neg(\mathcal{A} \cap \mathcal{B}) \equiv \neg\mathcal{A} \cup \neg\mathcal{B}$
11. $\neg(\mathcal{A} \cup \mathcal{B}) \equiv \neg\mathcal{A} \cap \neg\mathcal{B}$

$\Big\}$ , de wetten van de Morgan.

Een uitspraak, die van één of meer variabelen afhangt, en die dus pas een echte uitspraak wordt, als wij voor deze variabelen namen van de objecten, waarover wij spreken, substitueren, noemen wij een <u>praedicaat</u> of <u>relatie</u>: $\mathcal{A}(x)$ , $\mathcal{R}(x, y, z)$.
De uitspraken (1) zijn dus in deze zin praedicaten of relaties van proposities.
Uit een praedicaat kan door voorvoegen van een "quantificator" weer een uitspraak ontstaan. Wij zullen de volgende twee quantificatoren geregeld gebruiken:
1. <u>De universele quantificator</u> $\forall$, die aan het praedicaat $\mathcal{A}(x)$ de uitspraak $\forall_x \mathcal{A}(x)$ toevoegt, welke dan en slechts dan waar is, zo $\mathcal{A}(x)$ waar is voor elk object, waarover wij spreken.
2. <u>De existentiële quantificator</u> $\exists$, die aan het praedicaat $\mathcal{A}(x)$ de uitspraak $\exists_x \mathcal{A}(x)$ toevoegt, welke dan en slechts dan waar is, als er minstens een object bestaat, waarvoor $\mathcal{A}(x)$ een ware uitspraak is.
Door een bovenindex geven wij verder aan, wat de vz van mogelijke objecten is. Bijvoorbeeld:

$\forall_i^N \mathcal{A}(i)$ betekent: voor ieder natuurlijk getal $i$ geldt, dat $\mathcal{A}(i)$ waar is.

$\exists_a^R \mathcal{A}(a)$ betekent: er is (minstens) een reëel getal $a$, waarvoor geldt, dat $\mathcal{A}(a)$ waar is.

Praedicaten en proposities, die deel uitmaken van een relatie, zullen wij zo nodig tussen verticale streepjes (onder aan de regel) zetten. (In (1) geschiedde dit door het plaatsen van haakjes). Men maakt zodoende bijvoorbeeld onderscheid tussen:

$\forall_{x\,\|}\, \mathcal{A}(x) \cap \mathcal{B}(x)_{|} \cup_{|} \mathcal{C}(x)_{\|}$ en $\forall_{x\,\|}\, \mathcal{A}(x)_{|} \cap_{|} \mathcal{B}(x) \cup \mathcal{C}(x)_{\|}$.

Een <u>exclusief systeem</u> van uitspraken, $\mathcal{B}_1, \ldots, \mathcal{B}_n$ , is een systeem, waarbij ieder tweetal uitspraken elkaar uitsluiten (niet tegelijk waar kunnen zijn).
Een <u>volledig systeem</u> van uitspraken is een systeem, waarbij minstens één der uitspraken waar is.
Een <u>kategorisch systeem</u> van uitspraken is een systeem, dat zowel exclusief als volledig is.
Aan een praedicaat $\mathcal{A}(x)$ kan de verzameling $A$ van alle objecten $x$ , waarvoor $\mathcal{A}(x)$ waar is, toegevoegd worden.
Notatie: $A = \text{Ens}\{x \mid \mathcal{A}(x)\}$. (Ens is de afkorting van ensemble=vz)
Heeft men anderzijds een vz $A$ , dan kan men aan elke $x \in A$ [1]
de uitspraak $\mathcal{A}(x) \overset{\text{def}}{=}$[2] $x \in A$, toevoegen. Wij noemen de vz van alle objecten, waarvoor wij spreken, (de gehele vz) $\mathcal{I}$ .
$A$ is dus een deelvz van $\mathcal{I}$. Dan zal dus, als een element $x \in \mathcal{I}$ het praedicaat $\mathcal{A}$ bezit, d.w.z. als $\mathcal{A}(x)$ geldt, $x \in A$ zijn.
Op deze manier is aan elk praedicaat (aan te duiden met sierletters) één deelvz (aan te duiden met de overeenkomstige drukletters) toegevoegd, en omgekeerd.
Wij definiëren nu voor vzn:

$A \subset B \overset{\text{def}}{=} \forall_x,\ x \in A \rightarrow x \in B$, , <u>subsumptie</u> van $A$ en $B$, $A$ is een deelvz van $B$ .

$B \supset A \overset{\text{def}}{=} A \subset B$ .

Dan geldt:

$$(2)\begin{cases} 1.\ A \subset A, \\ 2.\ A \subset B \wedge B \subset C \rightarrow A \subset C, \\ 3.\ A \subset B \wedge B \subset A \equiv A = B. \end{cases}$$

De <u>doorsnede</u> van twee vzn $A$ en $B$ wordt gedefiniëerd door:

$A \cap B \overset{\text{def}}{=} Ens\{x \mid x \in A \wedge x \in B\}$ , de vz van de eln, die zowel tot $A$ als tot $B$ behoren.

De <u>vereniging</u> van twee vzn $A$ en $B$ wordt gedefiniëerd door:

$A \cup B \overset{\text{def}}{=} Ens\{x \mid x \in A \vee x \in B\}$ , de vz van die eln, die of tot $A$ of tot $B$ of tot beide vzn behoren.

De eigenschappen van de doorsnede- en verenigingsrelatie zijn volkomen aequivalent aan de eigenschappen van de conjunctie, resp. disjunctie uit de logica (eigenschappen 2 t/m 9 van (1)).
Deze relaties zijn dus <u>idempotent</u>, <u>commutatief</u>, <u>associatief</u>, alsmede <u>distributief</u> t.o.v. elkaar.

--------

1) $x \in A$ wil zeggen: $x$ is een element van de vz $A$.
2) De toevoeging def boven het gelijkheidsteken betekent: het rechterlid definieert het linkerlid.

Dus:

$$A \cap A = A\,;\ A \cap B = B \cap A\,;\ (A \cap B) \cap C = A \cap (B \cap C),$$

(3) $$A \cup A = A\,;\ A \cup B = B \cup A\,;\ (A \cup B) \cup C = A \cup (B \cup C),$$

$$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)\,;\ A \cup (B \cap C) = (A \cup B) \cap (A \cup C).$$

Het verschil (relatief complement) van een vz $B$ t.o.v. een vz $A$ definiëren wij als:

$A -- B \overset{def}{=} Ens\{x \mid x \in A \wedge \neg\, x \in B\}$ , de vz van al die eln, die wel tot $A$, maar niet tot $B$ behoren. Wij zullen hiervoor ook wel $A - B$ schrijven, wanneer daar althans nog geen andere betekenis aan is gegeven.

Wij noemen $\mathcal{I} -- A$ het complement van $A$. Dan geldt:

$$\mathcal{I} -- (\mathcal{I} -- A) = A,$$

(4) $$A -- (A -- B) = A \cap B,$$

$$\mathcal{I} -- (A \cap B) = (\mathcal{I} -- A) \cup (\mathcal{I} -- B),$$

$$\mathcal{I} -- (A \cup B) = (\mathcal{I} -- A) \cap (\mathcal{I} -- B).$$

Behoort bij iedere index $i$ uit een willekeurige vz $K$ (die niet eindig, noch aftelbaar behoeft te zijn) een vz $A_i$, dan kunnen wij doorsnede en vereniging van alle $A_i$ met $i \in K$ als volgt definiëren:

$$\bigcap_{i \in K} A_i \overset{def}{=} Ens\{x \mid \forall_i^K ,\, x \in A_i ,\},$$

$$\bigcup_{i \in K} A_i \overset{def}{=} Ens\{x \mid \exists_i^K ,\, x \in A_i ,\}.$$

## 1.3. Oriënterende opmerkingen over waarschijnlijkheidsrekening

De klassieke definitie van het begrip wh, afkomstig van Laplace, - de wh van een gebeurtenis is de verhouding van het aantal voor de gebeurtenis gunstige gevallen tot het totaal aantal gelijkelijk mogelijke gevallen - , is lang niet altijd bruikbaar [1]. Wij zullen in de komende hoofdstukken uitgaan van de axiomatische opzet van de whr (afkomstig van Kolmogorof [46]). Hierbij wordt een functie gedefinieerd op deelvzn van een bepaalde vz. Voldoet deze functie aan zekere voorwaarden, dan is zij als wh-functie te gebruiken. Is de functiewaarde voor een bepaalde deelvz gedefinieerd, dan is de wh van deze deelvz en van

---------

1) Omstreeks 1780 betoogde d'Alembert herhaaldelijk, dat de kans om in 2 worpen met een munt minstens eenmaal munt (M) te gooien niet 3/4 maar 2/3 was. Er zijn, zo betoogde hij, 3 gevallen: M, KM en KK (als bij de 1e worp M valt, is er geen 2e worp nodig). Van deze 3 mogelijkheden zijn er 2 gunstig. Necker, zoon van de minister van Financiën van Lodewijk XVI heeft zonder veel succes getracht d'Alembert te overtuigen, dat deze redenering fout was, omdat de 3 mogelijkheden niet "également possibles" zijn. Daarom heeft Laplace de voorwaarde van gelijke waarschijnlijkheid der mogelijke gevallen in zijn definitie opgenomen, hetgeen (tenzij men nadere voorzorgen neemt) een circulaire redenering inhoudt.

de ermee corresponderende bewering per definitie gelijk aan deze functiewaarde. Wij zullen dit alles hieronder nog preciseren. Op die manier krijgt men een zuiver mathematische waarschijnlijkheidstheorie, die, zoals in hoofdstuk 5 blijken zal, een onderdeel is van de maattheorie.
Een heel ander probleem is dan, hoe het zo gedefiniëerde wh-begrip in verband gebracht kan worden met de uit de waarneming bekende frequentiequotiënten en het intuïtieve wh-begrip, dus hoe de theorie in verband gebracht kan worden met de waarneming. Met dit probleem der empirische grondslagen zullen wij ons hier niet bezighouden. Men zie hiervoor [14] en [17] . Onze theorie is dus zuiver mathematisch. Bij de toepassingen, die wij geven, zal stilzwijgend een beroep gedaan worden op het intuïtieve wh-begrip van de lezer, dat echter alleen ter illustratie dient, evenals bijvoorbeeld de intuïtieve ruimtelijke aanschouwing voor de analytische of axiomatische meetkunde.

## 1.4. Eindige waarschijnlijkheidsvelden

Wij denken ons algemeen een eindige vz $\mathcal{I}$ gegeven, waarvan wij de eln aangeven met $1,2,\ldots, r$ , dus $\mathcal{I} = \{1,\ldots,r\}$.
Het stelsel van alle deelvzn van $\mathcal{I}$ vormt nu een eindig wh-veld, wanneer op dit stelsel gedefiniëerd is een verzamelingsfct $p$ met de volgende eigenschappen:

(1) $\forall_{A \subset \mathcal{I}} \; p(A) \geqq 0 \;, \; p(\mathcal{I}) = 1,$

(Dus voor alle deelvzn $A$ van $\mathcal{I}$ geldt, dat $p(A) \geqq 0$ is)
en

(2) $p(\bigcup_1^n A_j) = \sum_1^n p(A_j)$ mits $A_j \cap A_k = 0$ als $j \neq k$.

Een dergelijke fct kunnen wij krijgen door aan ieder element $i \in \mathcal{I}$ een getal $p_i$ toe te voegen met de eigenschappen:

(3) $p_i \geqq 0 \;, \sum_{i \in \mathcal{I}} p_i = 1.$

Wij definiëren dan het wh-veld door aan iedere deelvz $A \subset \mathcal{I}$ het getal

(4) $p(A) \stackrel{\text{def}}{=} \sum_{i \in A} p_i$

toe te voegen. Het is direct duidelijk, dat deze $p$ aan de eisen (1) en (2) voldoet. Een uitbreiding tot oneindige wh-velden zal gegeven worden in 5.6.
Wij kunnen nu de wh, dat de bewering $\mathcal{A}$ juist is, $\mathcal{P}(\mathcal{A})$, definiëren door:

(5) $\mathcal{P}(\mathcal{A}) \stackrel{def}{=} p(A)$, 1)

waarbij dus $A \stackrel{def}{=} \mathcal{E}_{ns}\{i \in \mathcal{I} \mid \mathcal{A}(i)\}$.
De <u>voorwaardelijke wh</u> van $\mathcal{A}$ onder de voorwaarde $\mathcal{B}$ wordt gedefiniëerd door:

(6) $$\mathcal{P}\{\mathcal{A} \mid \mathcal{B}\} \stackrel{def}{=} \frac{\mathcal{P}\{\mathcal{A} \cap \mathcal{B}\}}{\mathcal{P}\{\mathcal{B}\}}.$$

Deze definitie is alleen maar mogelijk als $\mathcal{P}\{\mathcal{B}\} \neq 0$ is.
$\mathcal{A}$ en $\mathcal{B}$ heten <u>onderling onafhankelijk</u> als:

$$\mathcal{P}\{\mathcal{A} \cap \mathcal{B}\} = \mathcal{P}\{\mathcal{A}\}.\mathcal{P}\{\mathcal{B}\}.$$

Dan is dus ook

$$\mathcal{P}\{\mathcal{A} \mid \mathcal{B}\} = \mathcal{P}\{\mathcal{A}\}.$$

Voor een <u>kategorisch systeem</u> van uitspraken $\mathcal{B}_1, \ldots, \mathcal{B}_m$ (zie 1.2) geldt verder, als $\mathcal{A}$ een willekeurige bewering is:

(7) $$\mathcal{P}\{\mathcal{A}\} = \sum_{1}^{m} \mathcal{P}\{\mathcal{B}_i\}.\mathcal{P}\{\mathcal{A} \mid \mathcal{B}_i\}.$$

Bij toepassing van de whr heeft men vaak te doen met experimenten waarvan de vz der mogelijke uitkomsten gegeven is. Deelvzn van deze vz noemt men <u>eventualiteiten</u>. Indien de uitkomst van het experiment in een (of meer) getallen wordt uitgedrukt, zoals bijvoorbeeld het geval is bij tellingen, metingen en rangschikkingen weergegeven in rangnummers, dan komen de eventualiteiten dus overeen met getallenvzn (bijvoorbeeld intervallen). Deze uitkomst is dan een variabele zodanig dat met iedere deelvz der waarden die zij kan aannemen een wh correspondeert dat de variabele tot die vz behoort. Een dergelijke grootheid noemt men een <u>stochastische variabele</u>. Wij zullen in dit rapport symbolen die een stochastische variabele representeren onderstrepen. Een willekeurige, door deze grootheid aangenomen waarde, wordt vaak aangegeven door hetzelfde symbool, maar dan niet onderstreept. De symbolen $P[\underline{x} = x]$ en $P[\underline{x} \leq x]$ betekenen dan dus: de wh dat de stochastische grootheid $\underline{x}$ de getallen-waarde $x$ aanneemt resp. de waarde $x$ niet overschrijdt. Indien men de wh $P[\underline{x} \leq x]$ voor iedere $x$ uit de vz der mogelijke uitkomsten van het experiment kent, dan kan men de wh van iedere eventualiteit betreffende $\underline{x}$ berekenen. Men zegt dan, dat de <u>wh-verdeling</u> van $\underline{x}$ beke[illegible]
<u>functie</u> $F(x) = P[\underline{x} \leq x]$ noemt men de <u>verdelingsfunctie</u> [illegible]

1) $\mathcal{P}$ is geen praedicaat maar een "functor" [illegible]
ken getallen toevoegt). We [illegible]
leen voor uitspraken, doch [illegible]
den.

Wij zullen aan de hand van enkele voorbeeldende hierboven ingevoerde begrippen toelichten.

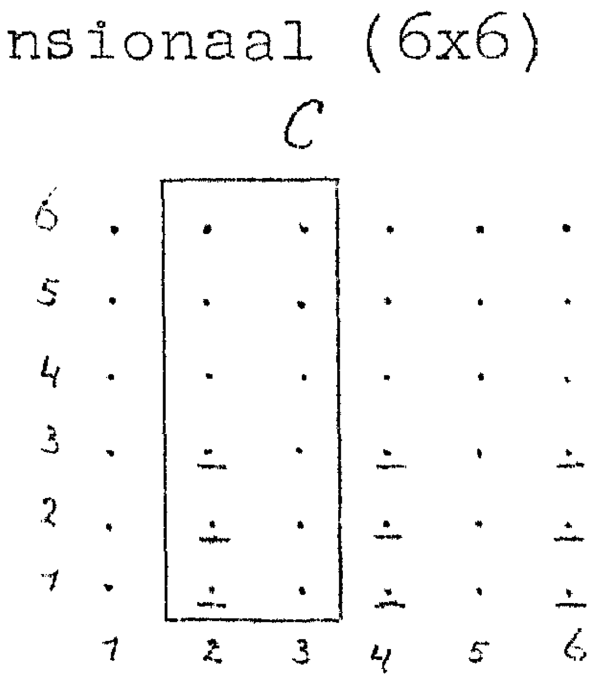
Ons eerste voorbeeld heeft betrekking op aantallen ogen geworpen met een zuivere dobbelsteen. Onder een zuivere dobbelsteen verstaat men gewoonlijk een dobbelsteen, die een zuivere kubusvorm heeft en die vervaardigd is van een homogeen materiaal. In deze zin is geen dobbelsteen volkomen zuiver. Bovendien zal men een dobbelsteen niet vertrouwen, waarbij men in lange reeksen worpen systematische verschillen vindt tussen de frequenties der mogelijke aantallen ogen. Wij zullen daarom een dobbelsteen zuiver noemen, als blijkt dat dergelijke systematische verschillen niet optreden. Bij gebruik in de whr zullen wij aan ieder der mogelijke worpresultaten: 1,2,...,6 een kans $\frac{1}{6}$ toekennen. De wh-verdeling van het aantal ogen $x$ wordt dus gegeven door:

$$P[\underline{x} = x] = \frac{1}{6} \qquad \text{voor } x = 1,2,\ldots,6.$$

Vragen wij nu naar de kans om een even aantal ogen te gooien, dan volgt uit de klassieke definitie direct, dat deze $\frac{1}{2}$ is. (3 van de 6 mogelijke gevallen zijn gunstig). Wij kunnen echter ook de volgende redenering houden: er zijn 3 mogelijkheden om een even aantal ogen te werpen; ieder van deze 3 mogelijkheden heeft de wh $\frac{1}{6}$; de totale wh van de gebeurtenis "een even aantal ogen komt boven" is dus $\frac{1}{6} + \frac{1}{6} + \frac{1}{6} = \frac{1}{2}$. Wij maken dan gebruik van de additiviteitsregel. Deze geldt echter niet altijd. Berekenen wij bijvoorbeeld de kans, dat een aantal ogen geworpen wordt, dat òf even is, òf kleiner dan 3 is. Men zou nu als volgt kunnen redeneren: de kans op een even aantal ogen is $\frac{1}{2}$, die op een aantal ogen, kleiner dan 3 (1 of 2) is $\frac{1}{3}$, dus de totale kans is $\frac{1}{2} + \frac{1}{3} = \frac{5}{6}$. Deze redenering is echter fout, zoals uit de klassieke definitie volgt (er zijn 4 gunstige gevallen: 1,2,4 en 6, dus de kans is $\frac{4}{6}$). De fout is gelegen in het feit, dat de beschouwde gebeurtenissen elkaar niet uitsluiten, doch gezamenlijk voor kunnen komen. Sluiten de gebeurtenissen elkaar wel uit, dan vertonen de waargenomen frequentiequotiënten deze additieve eigenschap en dus is het logisch deze additiviteit ook bij whn te verwachten of te eisen.

De beschouwde eventualiteiten zijn wel onafhankelijk. Weten wij namelijk reeds, dat een even aantal ogen geworpen is, dan blijft de kans op een 1 of 2 ongewijzigd gelijk $\frac{1}{3}$ (er zijn nu 3 mogelijkheden, 2,4 of 6; gunstig hiervan is alleen 2).

De voorwaardelijke wh van de gebeurtenis "er wordt een 1 of 2 geworpen" onder de voorwaarde "er is een even aantal ogen geworpen" is dus gelijk aan de wh van deze gebeurtenis zonder beperkende voorwaarde. Is dit het geval, dan heten de gebeurtenissen onafhankelijk. De kans, dat beide eventualiteiten tegelijk plaatsvinden is dan gelijk aan het product van de afzonderlijke kansen (in dit geval $\frac{1}{2} \times \frac{1}{3} = \frac{1}{6}$, zoals ook uit de klassieke definitie volgt; alleen de 2 is namelijk gunstig). Niet onafhankelijk zijn bijvoorbeeld de eventualiteiten "er wordt een even aantal ogen geworpen" en "er wordt een aantal ogen, kleiner dan 4, geworpen (1,2 of 3)". De tweede gebeurtenis heeft een kans $\frac{1}{2}$ (evenals de eerste), doch weten wij reeds, dat de eerste gebeurtenis heeft plaats gevonden, dan is de kans op de tweede gebeurtenis $\frac{1}{3}$ (3 mogelijke gevallen, 2,4 of 6; één gunstig, 2). De kans, dat beide gebeurtenissen tegelijk optreden is $\frac{1}{6}$ en niet $\frac{1}{4}$, dus niet gelijk aan het product der afzonderlijke kansen.
Werpen wij een zuivere dobbelsteen tweemaal, dan zijn er 36 mogelijke uitkomsten, die ieder de wh $\frac{1}{36}$ hebben. (Wij maken dus onderscheid tussen de uitkomsten $a, b$ en $b, a$ $(a \neq b)$ ).
Beweringen omtrent het aantal ogen van de eerste worp zijn nu onafhankelijk [1] van beweringen omtrent het aantal ogen van de tweede worp. Door vermenigvuldiging van de ermee corresponderende whn kunnen wij dus de wh van de bewering "beide beweringen zijn juist" berekenen. Zo is de wh, dat de eerste worp een even aantal ogen oplevert en de tweede een aantal, kleiner dan 4, gelijk aan $\frac{1}{2} \times \frac{1}{2} = \frac{1}{4}$ .
Wij kunnen de 36 mogelijke uitkomsten aangeven door de punten van een tweedimensionaal (6x6) rooster:




---------

1) Men neemt tegenwoordig algemeen aan, dat verschillende worpen met bijvoorbeeld een munt onafhankelijk van elkaar zijn. d'Alembert nam dit niet aan evenals trouwens tegenwoordig nog vele leken dit ook niet doen. Als vaak achtereen kruis gevallen is, meent men, dat de kans op munt groter geworden is. (Als bij de roulette het balletje enige malen op "rou gekomen is, zet men meer op "noir" dan op "rouge" in).

De eventualiteiten corresponderen nu met deelvzn van deze puntenvz. Zo correspondeert de eventualiteit "de eerste worp geeft 2 of 3 ogen" met de vz $C$ der ptn, die omlijnd zijn. De eventualiteit "de eerste worp geeft een even aantal ogen en de tweede een aantal, kleiner dan 4, correspondeert met de vz $D$ der punten, aangegeven door onderstreping.
Wij kunnen nu een wh-fct $p(A)$ invoeren, die aan iedere vz $A$ dat getal toevoegt, dat de wh voorstelt van de met deze vz corresponderende eventualiteit. In ons geval is $p(A) = \frac{1}{36}$ als $A$ een vz, bestaande uit een punt is, $p(C) = \frac{12}{36}$ ( $C$ bevat 12 punten), $p(D) = \frac{9}{36} = \frac{1}{4}$. Ook is duidelijk, dat $p(A \cup B) = p(A) + p(B)$ als $A \cap B = 0$, d.w.z. als de met $A$ en $B$ overeenkomende eventualiteiten elkaar uitsluiten.

Wij beschouwen nu een voorbeeld, waarbij aselecte trekkingen worden verricht uit gegeven eindige vzn (b.v.: een vaas met een gegeven aantal ballen). De term "aselect", die door ons in het Nederlands is ingevoerd als een equivalent voor het Engelse woord "random", heeft dezelfde functie als het woord "zuiver" bij het voorbeeld van de dobbelsteen. Wij noemen een trekking "aselect" als zij het resultaat is van een "aselecte trekkingsmethode", waarbij deze laatste gedefiniëerd is als een methode tot het verkrijgen of aanwijzen van één element ener collectie, welke methode niet afhangt van de kenmerken van het aan te wijzen element. Men kan een aselecte trekking bij benadering realiseren door er zorg voor te dragen, dat de persoon die of het instrument dat de trekking verricht de kenmerken niet kan waarnemen resp. niet op de kenmerken reageert. Ook hier zal men slechts door middel van experimenten kunnen nagaan of een trekkingsmethode voldoende aselect is. Dan zullen alle eln van de vz in lange waarnemingsreeksen ongeveer even vaak voorkomen. Bij aselecte trekking van één el uit een vz van de uitgebreidheid $N$ zullen wij aan ieder trekkingsresultaat de wh $\frac{1}{N}$ toekennen.

Het werpen met een zuivere dobbelsteen is in dit opzicht equivalent met het aselect trekken uit een vaas met 6 ballen, genummerd 1,...,6. Indien men b.v. geconstateerd heeft, dat de dobbelsteen voldoende zuiver is, dan kan men aselect trekken uit de vaas, door de bal te nemen waarvan het nummer overeenstemt met het aantal ogen dat men werpt met de dobbelsteen.
In het algemeen kan men uit iedere eindige vz aselect eln trekken, indien men een aselecte trekkingsmethode voor een bepaalde vz gevonden heeft.

Het voorbeeld, aan de hand waarvan wij formule (7) zullen toelichten heeft betrekking op $N+1$ vazen, die ieder $N$ ballen bevatten. De $i$-de vaas bevat $i$ rode en $N-i$ zwarte ballen. Wij kiezen aselect een vaas en uit de gekozen vaas trekken wij $n$ maal aselect een bal, waarbij de getrokken bal steeds teruggelegd wordt. Als nu blijkt, dat alle $n$ keren een rode bal getrokken wordt, hoe groot is dan de kans, dat de $n+1^{e}$ trekking uit de gekozen vaas ook een rode bal oplevert.

Oplossing:

Wij noemen:

$\mathcal{A} \overset{\text{def}}{\equiv}$ ,de eerste $n$ trekkingen uit de gekozen vaas geven rode ballen, en

$\mathcal{B} \overset{\text{def}}{\equiv}$ ,de eerste $n+1$ trekkingen uit de gekozen vaas geven rode ballen,.

Wanneer de gekozen vaas de $i$-de is (kans hierop $\frac{1}{N+1}$ ), dan is de voorwaardelijke kans op $n$ rode ballen achter elkaar volgens de productregel $\left(\frac{i}{N}\right)^{n}$. (De verschillende trekkingen zijn onderling onafhankelijk daar wij trekkingen met teruglegging beschouwen en steeds goed "schudden", d.w.z. er zorg voor dragen, dat bij iedere trekking alle ballen met gelijke wh gekozen kunnen worden).

Uit (7) volgt nu dus, als wij $C_i \equiv$ ,de $i$-de vaas wordt getrokken, invoeren, zodat

$$\mathcal{P}\{C_i\} = \frac{1}{N+1} \quad \text{en} \quad \mathcal{P}\{A \mid C_i\} = \left(\frac{i}{N}\right)^{n}:$$

(8) $$\mathcal{P}\{\mathcal{A}\} = \frac{1}{N+1} \sum_{1}^{N} \left(\frac{i}{N}\right)^{n}.$$

Evenzo

(9) $$\mathcal{P}\{\mathcal{B}\} = \frac{1}{N+1} \sum_{1}^{N} \left(\frac{i}{N}\right)^{n+1}.$$

Nu is de gevraagde kans gelijk aan:

$$\mathcal{P}\{\mathcal{B} \mid \mathcal{A}\} = \frac{\mathcal{P}\{\mathcal{A} \cap \mathcal{B}\}}{\mathcal{P}\{\mathcal{A}\}} = \frac{\mathcal{P}\{\mathcal{B}\}}{\mathcal{P}\{\mathcal{A}\}}$$ , daar de bewering $\mathcal{B}$ de bewering $\mathcal{A}$ impliceert (dus $B \subset A$, $A \cap B = B$).

Als $N$ groot is, geldt ongeveer:

$$\frac{1}{N} \sum_{1}^{N} \left(\frac{i}{N}\right)^{n} \approx \int_{0}^{1} x^{n}\, dx = \frac{1}{n+1},$$

zodat voor grote $N$ bij benadering geldt:

(10) $$\mathcal{P}\{\mathcal{B} \mid \mathcal{A}\} \approx \frac{n+1}{n+2} .$$

Is een vaas gevuld met een groot aantal rode of zwarte ballen, zodanig dat iedere samenstelling van de vaas even waarschijnlijk is (waarmee bedoeld wordt, dat de kansen, dat de vaas 0, resp. 1,2,..., $N$ rode ballen bevat, alle aan elkaar gelijk zijn) en hebben wij bij $n$ trekkingen met teruglegging steeds een rode bal verkregen, dan volgt uit het bovenstaande, dat de kans op een rode bal bij de $n+1$-de trekking ongeveer gelijk is aan $\frac{n+1}{n+2}$ . Dit is de zogenaamde " loi de succession" van Laplace, die zeer dikwijls op onjuiste wijze is toegepast. Zo geeft Laplace zelf de volgende toepassing: Gedurende de hele geschiedenis, b.v. 5000 jaar of 1.826.213 dagen is waargenomen, dat iedere dag de zon opging. Men kan dus 1.826.214 tegen 1 wedden, dat de zon ook morgen op zal gaan. Deze conclusie is natuurlijk zinloos, o.a. daar de onderstelde onafhankelijkheid der "experimenten" nergens op gebaseerd is.

In de elementaire whr heeft men meestal te maken met whn van conjuncties van o.o. praedicaten. Hebben wij bijvoorbeeld een vaas met $N$ ballen, $a$ rode en $N-a$ zwarte, en doen wij $n$ trekkingen met teruglegging (wij leggen dus steeds een bal van dezelfde kleur terug), dan is de kans om $k$ rode ballen te trekken gelijk aan:

$$\binom{n}{k}\left(\frac{a}{N}\right)^{k}\left(1-\frac{a}{N}\right)^{n-k}$$ , de zg. verdeling van Bernoulli of binomiale verdeling.

Wij kunnen echter ook, als wij een rode bal trekken, er een zwarte voor in de plaats in de vaas leggen en omgekeerd, dus steeds een bal van de andere kleur terugleggen. Het probleem wordt hierdoor ingewikkelder; de kans op een rode bal bij de $k$-de trekking, gegeven de uitkomsten van de $k-1$ voorafgaande trekkingen, hangt nu van deze uitkomsten af. Wij krijgen afhankelijke praedicaten. Dit probleem zal voor het geval $N = 4$, $a = 1$ en $n = 5$ aan het eind van 1.5 worden behandeld.
Wij zullen zien, dat de praedicaten, welke in de theorie van de stochastische processen beschouwd worden, meestal afhankelijk zijn.

## 1.5. Eenvoudig voorbeeld van een stochastisch proces

Twee personen 1 en 2 spelen een kansspel. Een van de spelers is bankhouder. Bij ieder spel wordt geloot of de bankhouder bankhouder blijft (kans hierop $p$) of dat de bank in handen van de andere speler overgaat (kans hierop $q = 1 - p$). Bij het begin van het spel zij speler 1 bankhouder.
Gevraagd wordt de kans, dat na $n$ spelen speler $i$, $i \in \{1,2\}$, bankhouder is. Hiertoe definiëren wij het praedicaat $\mathcal{B}(i) \overset{def}{=}$ $_{\text{l}}i$ is bankhouder$_{\text{l}}$, $i \in \{1,2\}$.
De wh, dat na $n$ spelen de speler $i$ bankhouder is, noemen wij $p_i^{(n)}$, dus $p_i^{(n)} \overset{def}{=} \mathcal{P}\{$ na $n$ spelen $\mathcal{B}(i)\}$.
Daar het spel begint met speler 1 als bankhouder geldt:

$$p_1^{(o)} = 1 \;,\; p_2^{(o)} = 0 \qquad \text{(beginvoorwaarden).}$$

Verder:

$$\forall_n^N \; p_i^{(n)} \geqq 0 \;,\; p_1^{(n)} + p_2^{(n)} = 1 .$$

Met behulp van de regel (1.4.7) vinden wij voor de $p_i^{(n)}$ onmiddellijk de betrekking:

$$\begin{cases} p_1^{(n)} = p_1^{(n-1)} p + p_2^{(n-1)} q \\ p_2^{(n)} = p_1^{(n-1)} q + p_2^{(n-1)} p . \end{cases}$$

De getallen $p_1^{(n)}$ en $p_2^{(n)}$ kunnen voor iedere $n$ beschouwd worden als de componenten van een vector met 2 componenten. Dus de vector $p_i^{(n)}$ ontstaat uit $p_i^{(n-1)}$ door een lineaire transformatie van de vorm:

$$p_j^{(n)} = \Sigma_i \, p_i^{(n-1)} P_{ij} \;, \text{ waarbij } i,j \in \{1,2\} \quad , \text{ terwijl de } P_{ij} \text{ 's}$$

de elementen zijn van de transformatiematrix $\mathcal{P} = (P_{ij}) = \begin{pmatrix} p & q \\ q & p \end{pmatrix}$.
In matrixnotatie:

$$p^{(n)} = p^{(n-1)} \mathcal{P} .$$

Hieruit volgt onmiddellijk vanwege de associativiteit van de matrixvermenigvuldiging [1]:

$$p^{(n)} = p^{(n-1)} \mathcal{P} = (p^{(n-2)} \mathcal{P}) \mathcal{P} = p^{(n-2)} \mathcal{P}^2 = p^{(n-k)} \mathcal{P}^k = p^{(o)} \mathcal{P}^n .$$

-------

1) Het product van twee $r$ bij $r$ matrices $M$ en $N$ wordt als volgt gedefiniëerd: $(MN)_j^i = \sum_1^r M_k^i N_j^k$ . Dit product is associa maar in het algemeen niet commutatief. Men kan op deze manier ook machten van matrices definiëren: $M^n = M^{n-1} M$ . Hiervoor geldt de eigenschap $M^n = M^a M^{n-a}$ voor elke gehele $a \leqq n$ . Zie ook 8.1.

Uitrekenen geeft:

$$P^2 = \begin{pmatrix} p^2+q^2 & 2pq \\ 2pq & p^2+q^2 \end{pmatrix}, \quad P^3 = \begin{pmatrix} p^3+3pq^2 & 3p^2q+q^3 \\ 3p^2q+q^3 & p^3+3pq^2 \end{pmatrix} \quad \text{enz.}$$

In het algemeen is:

$$P^n = \begin{pmatrix} \frac{1}{2}+\frac{1}{2}(p-q)^n & \frac{1}{2}-\frac{1}{2}(p-q)^n \\ \frac{1}{2}-\frac{1}{2}(p-q)^n & \frac{1}{2}+\frac{1}{2}(p-q)^n \end{pmatrix}, \text{ hetgeen te bewijzen is}$$

door volledige inductie:
Neem $n = 1$, dan is het beweerde juist (want $p+q=1$).
Stel het is goed voor een zekere $n$, dan rekenen wij dus uit:

$$P^{n+1} = P^n P = \begin{pmatrix} \frac{1}{2}+\frac{1}{2}(p-q)^n & \frac{1}{2}-\frac{1}{2}(p-q)^n \\ \frac{1}{2}-\frac{1}{2}(p-q)^n & \frac{1}{2}+\frac{1}{2}(p-q)^n \end{pmatrix} \begin{pmatrix} p & q \\ q & p \end{pmatrix} =$$

$$= \begin{pmatrix} \frac{1}{2}p+\frac{1}{2}p(p-q)^n+\frac{1}{2}q-\frac{1}{2}q(p-q)^n & \frac{1}{2}q+\frac{1}{2}q(p-q)^n+\frac{1}{2}p-\frac{1}{2}p(p-q)^n \\ \frac{1}{2}p-\frac{1}{2}p(p-q)^n+\frac{1}{2}q+\frac{1}{2}q(p-q)^n & \frac{1}{2}q-\frac{1}{2}q(p-q)^n+\frac{1}{2}p+\frac{1}{2}p(p-q)^n \end{pmatrix}.$$

Voor het rechterlid kunnen wij schrijven:

$$\begin{pmatrix} \frac{1}{2}+\frac{1}{2}(p-q)^{n+1} & \frac{1}{2}-\frac{1}{2}(p-q)^{n+1} \\ \frac{1}{2}-\frac{1}{2}(p-q)^{n+1} & \frac{1}{2}+\frac{1}{2}(p-q)^{n+1} \end{pmatrix}, \text{ q.e.d.}$$

Tenzij $pq = 0$ (dus $p = 0$ of $q = 0$, weinig interessant) geldt $|p-q| < 1$, dus $\lim\limits_{n\to\infty} (p-q)^n = 0$, zodat $\lim\limits_{n\to\infty} P^n = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$.
Daar $p^{(n)} = p^{(0)} P^n$ en $p^{(0)} = (1, 0)$ is dus

$$p^{(n)} = \left(\tfrac{1}{2}+\tfrac{1}{2}(p-q)^n, \tfrac{1}{2}-\tfrac{1}{2}(p-q)^n\right) \text{ en}$$

$$\lim_{n\to\infty} p_1^{(n)} = \lim_{n\to\infty} p_2^{(n)} = \tfrac{1}{2}.$$

Voor $n \to \infty$ heeft de vector $p_i^{(n)}$ dus een verdeling, die niet meer van de begintoestand afhangt. De wh. om na een groot aantal spelen bankhouder te zijn wordt op den duur voor beide spelers gelijk.
Wij hebben hiermee een eenvoudig voorbeeld behandeld van de stationaire enkelvoudige Markof-keten.

1.5. Markof-ketens

Gegeven is een systeem, dat op elk willekeurig tijdstip in één van $r$ verschillende toestanden verkeert. Deze toestanden worden aangegeven met behulp van de getallen $1,2,\ldots,r$. De vz van alle $r$ toestanden noemen wij $\mathcal{I}$, $\mathcal{I}=\{1,\ldots,r\}$. Wij onderstellen nu, dat op bepaalde discrete tijdstippen, de overgangstijdstippen, het systeem in een andere toestand kan overgaan en dat er voor elk overgangstijdstip een voorwaardelijke wh gedefiniëerd is, dat het systeem overgaat van de toestand, waarin het zich bevindt, in een willekeurige toestand uit $\mathcal{I}$. Deze whn noemen wij de overgangswhn. Zij kunnen afhangen van alle voorafgaande toestandsveranderingen.
Noemen wij $\underline{i}_k$ de stochastische grootheid, die de toestanden aangeeft, waarin het systeem na $k$ toestandsveranderingen kan verkeren en die dus de waarden $1,2,\ldots,r$ kan aannemen, dan bestaat volgens de veronderstelling:

$$\mathcal{P}\{\underline{i}_n = i_n \mid \underline{i}_0 = i_0, \ldots, \underline{i}_{n-1} = i_{n-1}\} \text{ voor alle } i_0, i_1, \ldots, i_n \in \mathcal{I}.$$

Hiervoor kunnen wij schrijven:

$$P_{(n)}{}^{i_0,\ldots,i_{n-1}}_{i_n}\ .$$

(Hierbij geeft de onderindex $(n)$ aan, dat deze wh de $n$-de toestandsverandering betreft).
Nu geldt natuurlijk:

$$P_{(n)}{}^{i_0,\ldots,i_{n-1}}_{i_n} = \frac{\mathcal{P}\{\underline{i}_0 = i_0, \ldots, \underline{i}_n = i_n\}}{\mathcal{P}\{\underline{i}_0 = i_0, \ldots, \underline{i}_{n-1} = i_{n-1}\}}.$$

Wij spreken dan van een hereditaire of meervoudige Markofketen (A.A. Markof, 1904).
Het eenvoudigste bijzondere geval is dat, waarbij

$$\mathcal{P}\{\underline{i}_n = i_n \mid \underline{i}_0 = i_0, \ldots, \underline{i}_{n-1} = i_{n-1}\} = \mathcal{P}\{\underline{i}_n = i_n\},$$

dus als de overgangswh bij de $n$-de toestandsverandering van geen enkele voorafgaande toestandsverandering afhangt. Voor verschillende $n$ worden de toestandsveranderingen onafhankelijke gebeurtenissen, de erop betrekking hebbende praedicaten zijn onafhankelijk, waardoor de elementaire whr kan worden toegepast.
Een ander belangrijk bijzonder geval krijgen wij, wanneer:

$$(1) \qquad \mathcal{P}\{\underline{i}_n = i_n \mid \underline{i}_0 = i_0, \ldots, \underline{i}_{n-1} = i_{n-1}\} = \mathcal{P}\{\underline{i}_n = i_n \mid \underline{i}_{n-1} = i_{n-1}\},$$

dus als

$$P_{(n)}{}^{i_0,\ldots,i_{n-1}}_{i_n} = P_{(n)}{}^{i_{n-1}}_{i_n}$$

voor $n \geqq 2$ niet van $i_0, i_1, \ldots, i_{n-2}$ afhangt.
Wij hebben dan een enkelvoudige Markof-keten. De overgangswhn bij de $n$-de toestandsverandering hangen dus - behalve van $n$ - alleen nog af van $i_{n-1}$, de toestand na $n-1$ overgangen, doch niet meer van de daaraan voorafgaande toestanden.
Wanneer wij de wh, dat het systeem na $n$ overgangen in de toestand $i$ verkeert, aangeven met $p_i^{(n)}$, wanneer dus:

(2) $$p_i^{(n)} = \mathcal{P}\{\underline{i}_n = i\},$$

dan geldt:

(3) $$p_j^{(n)} = p_i^{(n-1)} P_{(n)}{}^{i}_{j}$$ , (sommatie over $i$, tensornotatie!)

of in matrixnotatie:

$$p^{(n)} = p^{(n-1)} P_{(n)},$$

waarbij dus $P_{(n)} =$ matrix $(P_{(n)}{}^{i}_{j})$ 1) en $p^{(n)}, p^{(n-1)}$ vectoren zijn met $r$ componenten.
Hieruit volgt:

(4) $$p^{(n)} = p^{(0)} P_{(1)} P_{(2)} \ldots P_{(n)} .$$ 2)

Wij krijgen dus een gedurig product van matrices.
Noemen wij verder de wh, dat het systeem na $n$ overgangen in de toestand $i$ is, wanneer het na $m\ (< n)$ overgangen in de toestand $j$ verkeerde, $P_{(m,n)}{}^{i}_{j}$, dus

$$P_{(m,n)}{}^{i}_{j} \overset{def}{=} \mathcal{P}\{\underline{i}_n = i \mid \underline{i}_m = j\}$$, dan geldt:

(5) $$P_{(m,n)} = P_{(m+1)} P_{(m+2)} \ldots P_{(n)}$$ en

$$p^{(n)} = p^{(m)} P_{(m,n)},$$ zoals door inductie te bewijzen is. Verder $P_{(n)} = P_{(n-1,n)}$.

--------

1) In verband met het gebruik van de tensornotatie, onderscheiden wij de eln van een matrix door middel van een boven- en een beneden-index. De bovenindex correspondeert dan met het nummer van de rij, de benedenindex met het nummer van de kolom waarin het el voorkomt.
2) Bedenk hierbij, dat de matrixvermenigvuldiging in het algemeen niet commutatief is, zodat de volgorde der vermenigvuldigingen essentiëel is.

Voor al deze matrices geldt:

$$(6)\quad \begin{cases} \forall_n^{N'} \forall_i^{\mathcal{Y}} \forall_j^{\mathcal{Y}} \; P_{(n)\,j}^{\;i} \geqq 0 \\ \forall_n^{N'} \forall_i^{\mathcal{Y}} \sum_1^{\tau} P_{(n)\,j}^{\;i} = 1 \;. \end{cases}$$

(Dus voor elk natuurlijk getal $n > 0$, voor elke $i$ en $j$ uit $\mathcal{Y}$ geldt, dat $P_{(n)\,j}^{\;i} \geqq 0$ , het is immers een wh).

Aangaande de begintoestand kan men onderstellen:

a) Het systeem is in een bepaalde toestand $i_0$.

b) De begintoestand is niet bepaald, maar bezit een zekere wh-verdeling.

De tweede methode is algemener. Deze zullen wij kiezen. De wh-verdeling $p_i^{(0)}$ wordt daarbij bekend ondersteld. Methode a) is hiervan uiteraard een bijzonder geval, waarbij $p_i^{(0)} = 1$ is als $i = i_0$ , de bepaalde begintoestand, en $p_i^{(0)} = 0$ als $i \neq i_0$.

Een numerieke berekening is alleen mogelijk als $P_{(n)}$ op eenvoudige wijze van $n$ afhangt.

Het eenvoudigste geval is wel:

$P_{(n)}$ hangt niet van $n$ af, dus

$$P_{(1)\,j}^{\;i} = P_{(2)\,j}^{\;i} = \ldots = P_{(n)\,j}^{\;i} = P_j^i \;, \text{ dus } \forall_n^{N'} P_{(n)\,j}^{\;i} = P_j^i \,.$$

Wij spreken dan van een <u>stationaire (enkelvoudige) Markof-keten.</u>

Uit (3), (4), (5) en (6) volgt nu:

$$(7)\qquad p_j^{(n)} = p_i^{(n-1)} P_j^i \;,$$

$$(8)\qquad p^{(n)} = p^{(0)} P^n \quad \text{(het gedurig product wordt een macht van } P\text{)},$$

$$(9)\qquad p^{(n)} = p^{(m)} P^{n-m} \quad (m < n),$$

$$(10)\qquad \forall_i^{\mathcal{Y}} \forall_j^{\mathcal{Y}} \; P_j^i \geqq 0 \,, \qquad \forall_i^{\mathcal{Y}} \sum_1^{\tau} P_j^i = 1 \,.$$

Voor de beginvoorwaarden geldt:

$$(11)\qquad \forall_i^{\mathcal{Y}} \; p_i^{(0)} \geqq 0 \;, \; \sum_1^{\tau} p_i^{(0)} = 1 .$$

Nu geldt:

$$(12)\qquad \forall_i^{\mathcal{Y}} \; p_i^{(n)} \geqq 0 \;, \; \forall_n^{N} \sum_1^{\tau} p_i^{(n)} = 1 .$$

Immers voor $n = 0$ is dit waar. Stel, dat (12) reeds bewezen is met $n-1$ , in plaats van $n$ , dan volgt uit (7) en (10), dat ook $p_j^{(n)} \geqq 0$ is en dat $\sum_1^{\tau} p_j^{(n)} = \sum_1^{\tau} \sum_1^{\tau} p_i^{(n-1)} P_j^i = \sum_1^{\tau} p_i^{(n-1)} \sum_1^{\tau} P_j^i = \sum_1^{\tau} p_i^{(n-1)} = 1$ volgens onderstelling, q.e.d.

Stationaire Markof-ketens zijn dus bepaald door de matrix $P$ en d vector $p^{(0)}$.

De matrix $P$ noemen wij de matrix van overgangswhn of Markof-matrix. Elke vierkante matrix met $r$ rijen en kolommen, die de eigenschappen (10) bezit, is als matrix van overgangswhn bij een stationaire Markof-keten te beschouwen. Van dit soort matrices zullen wij dus de eigenschappen moeten bestuderen.
De overgangswhn hangen bij deze ketens dus alleen af van de toestanden, waartussen de overgang plaats vindt, doch niet van voorafgaande overgangen of overgangstijdstippen.

Voorbeeld

In een vaas bevinden zich 4 ballen, 1 rode en 3 zwarte. Men trekt aselect een bal uit de vaas en legt er een bal van de andere (niet getrokken) kleur voor in de plaats. Deze handeling wordt vijf maal uitgevoerd. Gevraagd wordt de wh, dat bij de vijfde trekking de getrokken bal rood is.

Oplossing:

Bevinden zich na een aantal trekkingen $i$ rode ballen in de vaas, dan is het systeem in toestand $i$. De verzameling van mogelijke toestanden is dan $\mathcal{I} = \{0, 1, \ldots, 4\}$.

De overgangswhn zijn:

$P^{i}_{i-1} = \frac{i}{4}$,

$P^{i}_{i+1} = \frac{4-i}{4}$,

$P^{i}_{j} = 0$ als $j \neq i-1$ en $j \neq i+1$.

Wij hebben dus te maken met een stationaire Markof-keten.
De matrix der overgangswhn ziet er, uitgeschreven, als volgt uit:

$$P = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ \frac{1}{4} & 0 & \frac{3}{4} & 0 & 0 \\ 0 & \frac{2}{4} & 0 & \frac{2}{4} & 0 \\ 0 & 0 & \frac{3}{4} & 0 & \frac{1}{4} \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}.$$

De wh, dat zich na $n$ trekkingen $j$ rode ballen in de vaas bevinden, kunnen wij, gezien (2), aangeven met $p^{(n)}_{j}$.
Nu is volgens (8):

$$p^{(4)} = p^{(0)} P^{4}.$$

Nu is:

$$p^{(0)} = (0, 1, 0, 0, 0) \quad \text{en} \quad P^{4} = \begin{pmatrix} \frac{5}{32} & 0 & \frac{3}{4} & 0 & \frac{3}{32} \\ 0 & \frac{17}{32} & 0 & \frac{15}{32} & 0 \\ \frac{1}{8} & 0 & \frac{3}{4} & 0 & \frac{1}{8} \\ 0 & \frac{15}{32} & 0 & \frac{17}{32} & 0 \\ \frac{3}{32} & 0 & \frac{3}{4} & 0 & \frac{5}{32} \end{pmatrix},$$

zoals eenvoudig na te rekenen is.

Voor het product $p^{(0)}\ P^4$ vinden wij dus:

$$p^{(4)} = (0 \quad \tfrac{17}{32} \quad 0 \quad \tfrac{15}{32} \quad 0).$$

De kans, dat zich na 4 trekkingen 1 rode bal in de vaas bevindt is dus $\frac{17}{32}$. (De kans bij de vijfde trekking een rode bal te krijgen is in dat geval $\frac{1}{4}$). De kans op 3 rode ballen na 4 trekkingen is $\frac{15}{32}$ (bij de vijfde trekking is de kans op een rode bal dan $\frac{3}{4}$). Na 4 trekkingen kunnen niet 0,2 of 4 rode ballen in de vaas voorkomen, zoals natuurlijk ook uit de opgave direct blijkt.
De gevraagde wh, dat bij de vijfde trekking een rode bal getrokken wordt, is nu:

$$p = \frac{17}{32} \cdot \frac{1}{4} + \frac{15}{32} \cdot \frac{3}{4} = \frac{31}{64}\ .$$

In dit eenvoudige geval was deze uitkomst natuurlijk ook sneller te krijgen zonder matrixvermenigvuldigingen:
Uit de opgave blijkt direct, dat het aantal rode ballen in de vaas na een even aàntal trekkingen oneven moet zijn, dus 1 of 3.

$$p_1^{(0)} = 1, \qquad p_i^{(0)} = 0 \text{ als } i \neq 1.$$

$$p_1^{(2)} = \frac{1}{4} \cdot \frac{4}{4} + \frac{3}{4} \cdot \frac{2}{4} = \frac{5}{8}\ .$$

Immers na de eerste trekking is het aantal rode ballen in de vaas 0 of 2. De kansen hierop zijn resp. $\frac{1}{4}$ en $\frac{3}{4}$. Om dan na de tweede trekking weer één rode bal te hebben in de vaas moet dus bij de tweede trekking een zwarte, resp. rode bal getrokken worden. De kansen hierop zijn $\frac{4}{4}$ , resp. $\frac{2}{4}$ . Hieruit volgt de betrekking voor $p_1^{(2)}$.

$$p_3^{(2)} = 1 - p_1^{(2)} = \tfrac{3}{8},$$

$$p_1^{(4)} = \tfrac{5}{8} \cdot p_1^{(2)} + \tfrac{3}{8} \cdot p_3^{(2)} = \left(\tfrac{5}{8}\right)^2 + \left(\tfrac{3}{8}\right)^2 = \tfrac{17}{32},$$

$$p_3^{(4)} = \tfrac{5}{8} \cdot p_3^{(2)} + \tfrac{3}{8} \cdot p_1^{(2)} = 2 \cdot \tfrac{3}{8} \cdot \tfrac{5}{8} = \tfrac{15}{32}\ .$$

Waaruit weer volgt:

$$p = \frac{31}{64}.$$

Wij willen ook nog de wh berekenen, dat bij 5 trekkingen 3 rode ballen getrokken worden. Uit de opgave is duidelijk, dat 3 rode ballen in 5 trekkingen dan en slechts dan getrokken worden als er na de vijfde trekking 0 rode ballen in de vaas zijn. De kans hierop is dus $(P^5)_0^1$.
Nu is:

$$P^5 = P . P^4 = \begin{pmatrix} 0 & \frac{17}{32} & 0 & \frac{15}{32} & 0 \\ \frac{17}{128} & 0 & \frac{3}{4} & 0 & \frac{15}{128} \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ \frac{15}{128} & 0 & \frac{3}{4} & 0 & \frac{17}{128} \\ 0 & \frac{15}{32} & 0 & \frac{17}{32} & 0 \end{pmatrix} .$$

Dus is $\left(P^5\right)^1_0 = \frac{17}{128}$ .
Noemen wij de kans op $k$ rode ballen in 5 trekkingen $(k \leq 5)$ $p_k$ ,
dan vinden wij zo: $p_0 = p_4 = p_5 = 0$,

$$p_1 = \left(P^5\right)^1_4 = \frac{15}{128} , \quad p_2 = \left(P^5\right)^1_2 = \frac{3}{4} , \quad p_3 = \left(P^5\right)^1_0 = \frac{17}{128} .$$

1.7. Generalisaties

a) In 1.6 onderstelden wij steeds, dat de vz van alle toestanden eindig was. Wij kunnen natuurlijk ook ketens beschouwen met een aftelbaar oneindig aantal mogelijke toestanden. Wij krijgen dan te maken met oneindige matrices.

b) Het aantal toestanden kunnen wij ook niet aftelbaar oneindig nemen. Het matrixbegrip moet dan verder gegeneraliseerd worden. Als vz van mogelijke toestanden kan dan ook een willekeurige abstracte vz genomen worden.

c) Tenslotte kunnen wij nog de onderstelling laten vallen, dat de toestandsveranderingen op discrete tijdstippen plaatsvinden. In plaats van $P^{(m,n)}_{ij}$ moeten wij dan beschouwen $P^{(t_1,t_2)}_{ij}$ , waarin $t_1$ en $t_2$ reële getallen zijn met $t_1 \leq t_2$.
Hiermee hebben wij het meest algemene geval gekregen van een stochastisch proces. Zijn de overganstijdstippen niet meer discreet, dan spreken wij liever van een proces dan van een keten.

Wij zullen in hoofdstuk 2 en 3 enkelvoudige stationaire Markof-ketens met een eindig aantal mogelijke toestanden bestuderen. In hoofdstuk 4 worden vervolgens als toepassing enige problemen uit de whr besproken (ruïnerings- en absorptieproblemen). In hoofdstuk 5,6 en 7 wordt als generalisatie een abstracte vz toegelaten als vz van mogelijke toestanden. De daar te behandelen ketens zijn echter ook enkelvoudig stationair en in de tijd discreet (de toestandsveranderingen blijven op discrete tijdstippen plaatsvinden). Van deze algemene Markof-ketens worden echter slechts enkele bijzondere onderwerpen beschouwd, namelijk absorptie- en terugkeer (lussen)-problemen. In een appendix (hoofdstuk 8 en 9) worden tenslotte enkele zuiver mathematische problemen behandeld die niet algemeen bekend ondersteld mochten worden en in de tekst gebruikt worden.

## 1.8. Toepassingsgebieden

De theorie der stochastische processen vindt toepassing op vele gebieden. Wij noemen hier:

Statistiek: b.v. in de sequente analyse,

Natuurkunde: b.v. beweging van Brown, ergodentheorie, turbulentietheorie, kernphysica,

Biologie:
Demografie: b.v. geboorte en sterfte, groeiprocessen,

Geneeskunde: b.v. besmettingstheorie,

Economie: b.v. tijdreeksen,

enz.

## 2. Eigenschappen van eindige Markof-matrices.

### 2.1. Markof-matrices.

Een Markof-matrix $\underline{P}$ is een vierkante matrix met reële elementen, waarvoor geldt:

$$\forall_i^{\mathfrak{J}} \forall_j^{\mathfrak{J}} \; P_j^i \geqq 0$$

(1.6.10)
$$\forall_i^{\mathfrak{J}} \sum_1^{r} {}^j P_j^i = 1$$

$\mathfrak{J}$ stelt hierbij weer de vz van mogelijke toestanden voor, $\mathfrak{J} = \{1,\ldots,r\}$, dus een eindige vz. Wij herinneren eraan, dat $P_j^i$ in een stationaire Markof-keten de voorwaardelijke wh voorstelt, dat het systeem, dat zich in de toestand $i$ bevindt, overgaat naar de toestand $j$. Hieruit volgt, dat de wh, dat het systeem, uitgaande van $i$, na één overgang in $k$ en na twee overgangen in $j$ zal zijn, gelijk is aan $P_k^i P_k^j$ (geen sommatie over $k$).
De voorwaardelijke wh, dat het systeem vanuit $i$ na twee overgangen in $j$ zal zijn is dus

$$\sum_1^{r} {}^k P_k^i P_j^k = (P^2)_j^i \; .$$

Evenzo geeft $(P^n)_j^i$ de voorwaardelijke wh, dat het systeem vanuit $i$ na $n$ overgangen in $j$ zal zijn.
Elke macht van een Markof-matrix is dus weer een Markof-matrix.
Voor ieder natuurlijk getal $n$ geldt derhalve:

$$\forall_n^{N'} \forall_i^{\mathfrak{J}} \forall_j^{\mathfrak{J}} \; (P^n) \geqq 0$$

(1)
$$\forall_n^{N'} \forall_i^{\mathfrak{J}} \sum_1^{r} {}^j (P^n)_j^i = 1 \; .$$

Deze betrekking kunnen wij natuurlijk ook rechtstreeks door volledige inductie bewijzen:

Stel dat (1) geldt voor een bepaalde $n$, dan zal $(P^{n+1})_j^i = (P^n)_k^i P_j^k \geqq 0$ zijn en $\sum_1^{r} {}^j (P^{n+1})_j^i = \sum^j \sum^k (P^n)_k^i P_j^k = \sum^k (P^n)_k^i \sum^j P_j^k = \sum^k (P^n)_k^i = 1$ krachtens inductieonderstelling. Daar (1) voor $n=1$ geldt, is hiermee het bewijs geleverd.

De algemene theorie van de vierkante matrices is ook op Markofmatrices van toepassing. Enkele stellingen en eigenschappen uit deze algemenetheorie zullen wij in de appendix (hoofdstuk 8) bewijzen en in dit hoofdstuk zonder bewijs toepassen.

## 2.2. Eigen waarden en eigen vectoren van Markof-matrices.

Het stelsel vergelijkingen in $f$ :

(8.2.1) $$\forall_i^J P_j^i f^j = \lambda f^i$$

of in matrixnotatie $Pf = \lambda f$ is alleen oplosbaar voor speciale waarden van $\lambda$ , namelijk die, welke voldoen aan de karakteristieke vergelijking $D(\lambda)=0$ en de eigen waarden worden genoemd. Hierin is $D(\lambda) = |P - \lambda I|$, dat is de determinantwaarde van de matrix $(P_j^i - \lambda I_j^i)$ , waarbij $I$ (iota) de eenheidsmatrix voorstelt (dus $I_j^i = \left\{\begin{matrix} 1 \text{ als } i=j \\ 0 \text{ als } i \neq j \end{matrix}\right)$ . Er zijn hoogstens $\tau$ verschillende eigen waarden. Wij kunnen schrijven:

(8.2.5) $$D(\lambda) = (-1)^\tau \prod_{\rho=0}^{s-1} (\lambda - \lambda_\rho)^{n_\rho} \quad \text{met} \quad \sum_0^{s-1} n_\rho = \tau ,$$

als de eigen waarde $\lambda_\rho$ de multipliciteit $n_\rho$ heeft en er $s$ verschillende eigen waarden zijn.

Is $\lambda$ nu een eigen waarde, dan heeft het stelsel (8.2.1) dus een oplossing, die een eigen vector wordt genoemd, behorende bij de eigen waarde $\lambda$ . Het aantal l.o. eigen vectoren, behorende bij eenzelfde eigen waarde $\lambda$ zal hoogstens gelijk aan de multiplciteit van $\lambda$ zijn.

Is $\lambda$ een eigen waarde, dan heeft ook het stelsel

(8.2.6) $\forall_j^J g_i P_j^i = \lambda g_j$ of $g' P = \lambda g'$ een oplossing, een eigen covector $g'$ .

(Wij onderscheiden covectoren door een accent van gewone vectoren; dit accent wordt echter ook wel weggelaten, wanneer er geen mogelijkheid tot verwarring is)

Enkele stellingen (zie voor de bewijzen 8.2. en 8.3):

1. Is $f$ een eigen vector van $P$ bij een eigen waarde $\lambda$ , dan zal $f$ tevens een eigen vector van $P^n$ zijn bij de eigen waarde $\lambda^n$ Dus als $P_j^i f^j = \lambda f^i$ geldt ook $(P^n)_j^i f^j = \lambda^n f^i$
2. Is $f$ een eigen vector en $g'$ een eigen covector van $P$ , behorende bij verschillende eigen waarden, dan geldt $g'f = 0$ .

3. Als geen twee eigen waarden $\lambda_\alpha$, $\alpha \in \{0,\dots,z\text{-}1\}$ van $P$ aan elkaar gelijk zijn, als $f_\alpha$ de eigen vector en $g_\alpha'$ de eigen covector voorstelt bij $\lambda_\alpha$, dan geldt na normering van de $g_{\alpha i}$ :

$$g_{\alpha i} f_\beta^i = \delta_{\alpha\beta} \qquad 1)$$

4. Onder de voorwaarden van stelling 3 geldt na normeren:

$$\forall_k^{N'} (P^k)^i_j = \Sigma^\alpha \lambda_\alpha^k f_\alpha^i g_{\alpha j} \quad \text{, dus als } k=1:\ P^i_j = \Sigma^\alpha \lambda_\alpha f_\alpha^i g_{\alpha j}.$$

Tenslotte bewijzen wij de stellingen:

5. De eigen waarden van een Markof-matrix zijn in absolute waarde kleiner dan of gelijk aan 1.

   Bewijs:

   $P^i_j f^j = \lambda f^i$ dus $|\lambda|.|f^i| = |P^i_j.f^j| \leq \Sigma^j P^i_j .\max\limits_{k\in J}|f^k| = \|f\|$

   wanneer wij (1) $\|f\| \overset{def}{=} \max\limits_{k\in J}|f^k|$ invoeren (zie voor deze normdefinitie 8.8)

   Dus $\forall_i^J |\lambda|.|f^i| \leq \|f\|$, dus ook $|\lambda|.\|f\| \leq \|f\|$, waaruit volgt, daar altijd $\|f\| > 0$, dat $|\lambda| \leq 1$, q.e.d.

6. Elke Markof-matrix heeft een eigen waarde $\lambda = 1$. Een bij $\lambda = 1$ behorende eigen vector is de eenheidsvector $\begin{pmatrix}1\\ \vdots\\ 1\end{pmatrix}$ (waarvan wij de $i$-de component aangeven met $1^i$).

   Bewijs: In $P^i_j f^j = \lambda f^i$ vullen wij $\lambda = 1$ en $f^i = 1^i = 1$ in. Dan krijgen wij: $\Sigma^j P^i_j . 1^j = 1.1^i$ dus $\Sigma^j P^i_j = 1$, hetgeen juist is. Dus is inderdaad $f^i = 1^i$ een eigen vector bij de eigen waarde 1, q.e.d.

   De eigen waarde 1 zullen wij soms aangeven met $\lambda_0$.

2.3. <u>Markof-matrices met enkelvoudige eigen waarde 1 en geen andere eigen waarden met modulus 1.</u>

Wij beschouwen nu een Markof-matrix, die de eigen waarde 1 enkelvoudig heeft en waarvan alle andere eigen waarden een modulus kleiner dan 1 hebben. Als stelling 4 van 2.2. toegepast mag worden, geldt:

$$(P^n)^i_j = \Sigma^\alpha \lambda_\alpha^n f_\alpha^i g_{\alpha j} = \lambda_0^n f_0^i g_{0j} + \sum_{\alpha\neq 0}^{\alpha} \lambda_\alpha^n f_\alpha^i g_{\alpha j} = g_{0j} + \sum_{\alpha\neq 0} \lambda_\alpha^n f_\alpha^i g_{\alpha j},$$

---------

1) $\delta_{\alpha\beta} = \begin{cases} 1 \text{ als } \alpha = \beta \\ 0 \text{ als } \alpha \neq \beta \end{cases}$ (Kronecker-symbool)

dus

(1) $\lim\limits_{n\to\infty}(P^n)^i_j = g_{oj}$ , waarbij $g_o$ de genormeerde eigen covector is, behorende bij de eigen waarde 1.

Uit de betrekkingen (2.1.1) volgt nu direct:

(2) $\forall^{J}_{j}\, g_{oj} \geqq 0\,,\quad \sum^{J} g_{oj} = 1.$

Dus de $g_{oj}$ zijn whn.

Verder is

(1.6.8) $p_j^{(n)} = p_i^{(o)}(P^n)^i_j$

dus

(3) $\lim\limits_{n\to\infty} p_j^{(n)} = g_{oj}.$

De wh $p_j^{(n)}$ , dat is de wh, dat het systeem na $n$ overgangen in de toestand $j$ verkeert, heeft een limiet, die onafhankelijk is van de beginverdeling (in een speciaal eenvoudig geval lieten wij dit reeds zien in 1.5).

Men zou kunnen denken, dat de betrekkingen (1), (2) en (3) alleen gelden als alle eigen waarden met een modulus kleiner dan 1 ook enkelvoudig zijn. Immers anders mag stelling 4 van 2.2 niet toegepast worden. Later (2.7) zullen wij echter zien, dat in ons geval, waar behalve 1 geen eigen waarden met modulus 1 bestaan, de multipliciteit van de eigen waarden met modulus kleiner dan 1, niet ter zake doet.

Voorbeeld

Gegeven zijn twee vazen, $n$ witte ballen en $n$ zwarte ballen; in ieder van de vazen bevinden zich $n$ van deze ballen.

Bij de aanvang van het experiment bevinden zich in de eerste vaas $a$ witte ballen. Daarna wordt uit ieder van beide vazen aselect één bal getrokken en in de andere vaas gelegd. Deze handeling wordt herhaald.

Gevraagd wordt de Markof-matrix van het systeem te bepalen en de limiet voor $k\to\infty$ van de kans, dat er zich na $k$ verwisselingen $j$ witte ballen in de 1e vaas bevinden.

Oplossing:

$J = \{0,\ldots,n\}.$

Wanneer $i$ het aantal witte ballen in de 1e vaas voorstelt, dan zijn de overgangswhn

$$(4)\quad \begin{cases} P^i_{i-1} = \dfrac{i^2}{n^2} \\ P^i_i = \dfrac{2i(n-i)}{n^2} \\ P^i_{i+1} = \dfrac{(n-i)^2}{n^2} \end{cases} \qquad \text{en } P^i_j = 0 \quad \text{als } j \neq i-1,\ i, \text{ of } i+1.$$

Controleer dat $\sum_0^n{}_j\, P^i_j = 1$.

Om de eigen covector $g_0$ te vinden zoeken wij de oplossing van $g'P = \lambda g'$ voor $\lambda = \lambda_0 = 1$. De index 0 van $g_0$ zullen wij hierbij weglaten.

$\sum^i g_i P^i_j = g_j$ dus

$$g_j = g_{j-1}\frac{(n-j+1)^2}{n^2} + g_j\frac{2j(n-j)}{n^2} + g_{j+1}\frac{(j+1)^2}{n^2}.$$

Deze betrekking is geldig voor $j = 0, \dots\dots\dots, n$ als wij toelaten

$g_{-1} = g_{n+1} = 0.$

Wij vinden:

$g_0 = g_1 \cdot \frac{1}{n^2}$, dus $g_1 = n^2 g_0 = \binom{n}{1}^2 g_0$.

Algemeen geldt:

$$g_j = \binom{n}{j}^2 g_0.$$

Immers het gestelde is juist voor $j=0$ en $j=1$. Stel dat het geldt voor $j = k-1$ en $j = k$, dan volgt uit:

$g_k = g_{k-1}\frac{(n-k+1)}{n^2} + g_k\frac{2k(n-k)}{n^2} + g_{k+1}\frac{(k+1)^2}{n^2}$, dus volgens inductieonderstelling:

$$\binom{n}{k}^2 g_0 = \binom{n}{k-1}^2\frac{(n-k+1)^2}{n^2}g_0 + \binom{n}{k}^2\frac{2k(n-k)}{n^2}g_0 + g_{k+1}\frac{(k+1)^2}{n^2}, \text{ dat}$$

$$g_{k+1} = \frac{n^2}{(k+1)^2}\left\{\binom{n}{k}^2 - \binom{n}{k-1}^2\frac{(n-k+1)^2}{n^2} - \binom{n}{k}^2\frac{2k(n-k)}{n^2}\right\}g_0 = \binom{n}{k+1}^2 g_0, \text{ q.e.d.}$$

Wij moeten de covector $g$ zo normeren, dat $\sum_0^n{}_j\, g_j = 1$ is.

Dus: $$g_j = \frac{\binom{n}{j}^2}{\sum_0^n{}_i\binom{n}{i}^2} = \frac{\binom{n}{j}^2}{\binom{2n}{n}}. \quad 1)$$

---

1) Dat $\sum_0^n{}_i\binom{n}{i}^2 = \binom{2n}{n}$ is, ziet men eenvoudig door $(a+b)^{2n}$ en $(a+b)^n.(a+b)^n$ in reeksen te ontwikkelen en de coëfficiënt van $a^n b^n$ in beide gevallen te beschouwen.

Als nu inderdaad de eigen waarde 1 enkelvoudig is en de andere eigen waarden een modulus kleiner dan 1 hebben, dan zal dus:

(5) $$\lim_{k\to\infty} p_j^{(k)} = \frac{\binom{n}{j}^2}{\binom{2n}{n}}$$

zijn, waarmee de gevraagde limiet gevonden is. Wij zullen nog aantonen, dat bij de eigen waarde 1 slechts één eigen vector gevonden kan worden en dat eigen waarden met modulus 1, die ongelijk 1 zijn, niet kunnen voorkomen als $n>1$ is. Later zullen wij bewijzen, dat als bij de eigen waarde 1 van een Markof-matrix slechts één eigen vector gevonden kan worden, de eigen waarde 1 ook enkelvoudig is. De differentie vergelijking voor de eigen vectoren luidt, zoals uit (4) volgt:

(6) $$i^2 f^{i-1} + 2i(n-i)f^i + (n-i)^2 f^{i+1} = \lambda n^2 f^i \ , \quad i \in \{0,\dots,n\} \text{ als } f^{-1} \overset{\text{def}}{=} 0 \text{ en } f^{n+1} \overset{\text{def}}{=} 0.$$

a) $\lambda = 1$.

Wij kunnen (6) schrijven in de vorm:

$(n-i)^2(f^{i+1}-f^i)-i^2(f^i-f^{i-1})=0$, dus uit $f^i=f^{i-1}$ volgt $f^{i+1}=f^i$.

Neem $i=0$ dan zien wij $f^1=f^0$ dus $f^0=f^1=f^2=f^3=\dots\dots=f^n$, zodat geen met de eenheidsvector l.o. vector gevonden kan worden.

b) $\lambda \neq 1$, $|\lambda|=1$.

Uit (6) volgt nu:

$$n^2|f^i| \leqq i^2|f^{i-1}| + 2i(n-i)|f^i| + (n-i)^2|f^{i+1}| \qquad \text{of}$$

$$(n-i)^2\{|f^{i+1}|-|f^i|\} - i^2\{|f^i|-|f^{i-1}|\} \geqq 0 \ , \qquad \text{dus :}$$

uit $|f^i| \geqq |f^{i-1}|$ volgt $|f^{i+1}| \geqq |f^i|$ en

uit $|f^{i+1}| \leqq |f^i|$ volgt $|f^i| \leqq |f^{i-1}|$.

Neem $i=0$, dan zien wij: $f^1=\lambda f^0$, dus $|f^1|=|f^0|$, dus ook $|f^1|\geq|f^0|$

Neem $i=n$, dan zien wij: $f^{n-1}=\lambda f^n$ dus $|f^{n-1}|=|f^n|$ dus ook $|f^{n-1}|\geq|f^n|$

Hieruit volgt direct, dat $|f^0|=|f^1|=|f^2|=\dots\dots=|f^n|$ moet gelden.

Neem $i=1$ :

$$f^0+2(n-1)f^1+(n-1)^2f^2=\lambda n^2 f^1, \text{dus } (n-1)^2 f^2 = (\lambda^2 n^2 - 2(n-1)\lambda - 1)f^0.$$

Nu is $|\lambda^2 n^2 - 2(n-1)\lambda - 1| \geqq n^2 - 2(n-1) - 1 = (n-1)^2$.
Het gelijkheidsteken is alleen mogelijk als in het complexe vlak $\lambda^2 n^2$, $2(n-1)\lambda$ en 1 op een rechte lijn door de oorsprong liggen. Dit is alleen zo als $\lambda = \pm 1$ is. Wij sloten reeds uit $\lambda = 1$
Stel $\lambda = -1$ , dan wordt dus $(n-1)^2 f^2 = (n^2 + 2(n-1) - 1) f^0 = \{(n-1)^2 + (4n-4)\} f^0$, dus $|f^2| > |f^0|$ als $n > 1$ is. Het geval $n = 1$ zullen wij apart bekijken. Als $\lambda \neq \pm 1$ is, geldt dus, zoals uit het bovenstaande direct volgt:

$$(n-1)^2 |f^2| > (n-1)^2 |f_0| , \quad \text{dus ook } |f^2| > |f^0| .$$

Wij bewezen echter reeds $|f^2| = |f^0|$ , zodat hieruit volgt, dat als $n > 1$ is, eigen waarden $\neq 1$ met modulus 1 inderdaad niet voorkomen zodat uit a) en b) volgt, dat (5) geldt, als $n > 1$ is.

Apart behandelen wij het geval $n = 1$ , d.w.z. elke vaas bevat één bal, een witte of een zwarte. Uit (4) volgt nu:

$$P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

Uitrekenen geeft: $P^2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, $P^3 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, algemeen $P^n = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ als $n$ oneven en $P^n = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ als $n$ even. Dus $P^n$ heeft geen limiet. Verder is duidelijk: $g_0 = (\frac{1}{2}, \frac{1}{2})$ , want $(\frac{1}{2}, \frac{1}{2})\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = (\frac{1}{2}, \frac{1}{2})$.
Echter heeft $P^n$ geen limiet, zodat (1) nu niet geldt, evenmin als (3). Inderdaad is het aantal witte ballen in de eerste vaas nu niet onafhankelijk van de begintoestand na een groot aantal trekkingen, maar daar juist volkomen door bepaald.
Na een even aantal trekkingen zal in de eerste vaas precies die bal zitten, die er in het begin inzat, na een oneven aantal precies de andere. Het geval $n = 1$ neemt dus een uitzonderingspositie in; $g_0$ heeft hier geen wh-theoretische betekenis.

2.4. <u>Markof-matrices met meervoudige eigen waarde 1.</u>

Wij beschouwen nu het geval, dat de eigen waarde 1 meervoudig is. Dan is er, zoals wij in 2.6 zullen bewijzen bij de eigen waarde 1 nog minstens één eigen vector $f$ , niet evenredig met de eenheidsvector. Voor deze vector geldt dus:

$$(1) \quad P^i_j f^j = f^i , \; \|f\| > 0 \qquad \text{en} \qquad \exists^{\mathfrak{I}}_i f^i \neq f^1 .$$

Zonder beperking kunnen wij aannemen, dat $\|f\| = 1$ is, dus, daar $\mathfrak{I}$ eindig ondersteld is:

$$\exists^{\mathfrak{I}}_i |f^i| = 1 .$$

Ook houdt het geen beperking in voor deze $i$ te eisen: $f^i = 1$
Definieer nu:

(2) $$S \overset{def}{=} \mathcal{E}ns\,\{i \in \mathfrak{I} \mid f^i = 1\},$$

dan is $S$ een echte, niet lege deelverzameling van $\mathfrak{I}$.
Nu geldt dus:

(3) $$\begin{cases} i \in S \leftrightarrows f^i = 1 \\ i \in \mathfrak{I} - S \rightleftarrows f^i \neq 1,\ |f^i| \leq 1 \end{cases}$$

Kies $i \in S$:

$$\sum_1^r{}^j P_j^i = 1 = f^i = P_j^i f^j = \sum_1^r{}^j P_j^i f^j, \qquad \text{dus}$$

(4) $$\sum{}^j P_j^i (1 - f^j) = 0.$$

Nodig hiervoor is, dat het reële deel van (4) 0 is, dus:

(5) $$\sum{}^j P_j^i (1 - \mathrm{Re}\, f^j) = 0.$$

Nu is $\mathcal{R}e\, f^j = 1$ als $j \in S$ is, want dan is $f^j = 1$ en $\mathcal{R}e\, f^j < 1$ als $j \in \mathfrak{I} - S$, want dan is $|f^j| \leq 1$ en $f^j \neq 1$. (5) stelt dus een som voor van niet negatieve getallen, die 0 moet zijn, dus iedere term moet 0 zijn. Nu is van elke term van (5) de tweede factor dan en slechts dan 0 als $j \in S$ is, zodat als $j \notin S$ moet gelden $P_j^i = 0$. Waarmee dus bewezen is:

(6) $$\forall_i^S \forall_j^{\mathfrak{I}-S}\ P_j^i = 0.$$

Definiëren wij, als $A \subset \mathfrak{I}$ is,

(7) $$P_A^i \overset{def}{=} \sum_{j \in A}{}^j P_j^i$$

dus $P_A^i$ is de wh, dat een systeem, dat zich in de toestand $i$ bevindt, na eén overgang in een toestand van de vz $A$ is gekomen, dan volgt uit (6)

(8) $$\begin{cases} \forall_i^S\ P_{\mathfrak{I}-S}^i = 0, \text{ dus} \\ \forall_i^S\ P_S^i = 1, \quad \text{daar} \quad P_{\mathfrak{I}-S}^i + P_S^i = \sum_1^r{}^j P_j^i = 1. \end{cases}$$

De matrix van overgangswhn $P$ ziet er nu als volgt uit, eventueel na verwisseling van rijen en kolommen:

$$P = \begin{array}{c} \\ S \\ J-S \end{array} \!\!\begin{array}{c} \begin{array}{cc} \overbrace{\quad}^{S} & \overbrace{\quad}^{J-S} \end{array} \\ \left( \begin{array}{c|c} * & 0 \\ \hline * & * \end{array} \right) \end{array} ,$$

dus alleen de deelmatrices, aangeduid met een $*$ kunnen elementen $\neq 0$ hebben. De verwisseling van rijen en kolommen komt neer op een andere nummering der mogelijke toestanden. Dit is natuurlijk niet essentieel.

Uit het voorgaande volgt, dat een systeem, dat zich eenmaal in een der toestanden van $S$ bevindt, nooit meer in een niet tot $S$ behorende toestand kan komen, (nooit meer, d.w.z. behoudens een wh 0). Een deelvz van $J$ met een dergelijke eigenschap noemen wij een <u>stochastisch afgesloten verzameling</u> of een <u>fuik</u>. $S$ is dus een fuik. De door $S$ bepaalde deelmatrix van $P$ noemen wij gemakshalve eveneens een fuik. Triviale fuiken zijn natuurlijk $J$ en de lege vz.

Definieer nu:

(9) $\qquad h^i = 1^i - f^i ,$

waarbij $f$ aan (1) voldoet en zodanig genormeerd is, dat $\|f\| = 1$ en $\exists_j^J f^j = 1$. $\quad h^i = 0$ als $i \in S$ en $\exists_i^J h^i \neq 0$ , daar $\exists_i^J f^i \neq 1$, dus $h^i$ is noch de nulvector noch evenredig met de eenheidsvector. $h$ is weer een eigen vector van $P$ bij de eigen waarde 1, daar het verschil van twee eigen vectoren bij dezelfde eigen waarde weer een eigen vector is bij die eigen waarde. $h$ is weer zo te normeren, dat $\|h\| = 1$ en $\exists_i^J h^i = 1$ is.

Definiëer:

(10) $\qquad T \overset{\text{def}}{=} \mathcal{E}ns\{ i \in J \mid h^i = 1 \} ,$

dan is $T$ volgens het voorgaande weer een fuik, dus:

(11) $\forall_i^T P_{J-T}^i = 0 \quad$ en

$\qquad \forall_i^T P_T^i = 1 .$

Verder geldt:

(12) $\quad S \cap T = 0$,

daar als $i \in S$ uit (9) volgt $h^i = 0$ , dus $i \notin T$.
Nummeren wij de mogelijke toestanden geschikt, dan krijgt $P$ dus de vorm:

$$(13) \quad P = \begin{array}{c} \\ S \\ T \\ J-S-T \end{array} \begin{array}{c} \begin{array}{ccc} S & T & J-S-T \end{array} \\ \left(\begin{array}{c|c|c} * & 0 & 0 \\ \hline 0 & * & 0 \\ \hline * & * & * \end{array}\right) \end{array}$$

waarbij $J-S-T = 0$ natuurlijk mogelijk is en alleen de deelmatrices, aangeduid met een $*$ elementen $\neq 0$ kunnen hebben.
Bewezen is dus de

<u>Stelling</u>

Een Markof-matrix, die de eigen waarde 1 meervoudig heeft, bezit minstens twee disjuncte niet triviale fuiken.

De door $S$ en $T$ bepaalde deelmatrices van $P$ zijn weer Markof-matrices. De matrix $P$ en ook de Markof-keten zelf zijn in dit geval <u>reducibel</u>. Wij noemen een Markof-matrix namelijk reducibel als hij minstens een niet-triviale fuik bevat. Is dit niet het geval, dan spreken wij van een <u>irreducibele</u> Markof-keten of -matrix. Een reducibele Markof-matrix kan de eigen waarde 1 enkelvoudig hebben. Neem bijvoorbeeld:

$$P = \begin{pmatrix} \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 1 \end{pmatrix}$$

dan is de toestand 3 een fuik, doch de eigen waarde 1 is enkelvoudig. Wij mogen de in deze paragraaf gegeven theorie dus niet omkeren. Wel is direct duidelijk, dat, wanneer er twee disjuncte niet-triviale fuiken zijn, dat dan de eigen waarde 1 meervoudig is. Dit volgt namelijk direct uit (13): de door $(S,S)$ en $(T,T)$ bepaalde matrices zijn weer Markof-matrices en hun karakteristieke veelterm bevat de factor $1-\lambda$ dus minstens éénmaal. De karakteristieke veelterm van $P$ is verder het product van die van $(S,S)$, $(T,T)$ en $(J-S-T\,,\,J-S-T)$ , waaruit het gestelde volgt.

Als $J - S - T = o$ , als dus zowel $S$ als $J-S$ een fuik is, dan noemen wij de Markof-keten en de Markof-matrix <u>ontbindbaar</u> (<u>decomposabel</u>). Nu zal de vorm van $P$ zijn:

$$P = \begin{array}{c|c} & S \quad\; J-S \\ \begin{array}{c} S \\ J-S \end{array} & \left(\begin{array}{c|c} * & o \\ \hline o & * \end{array}\right) \end{array}$$

Het proces is gesplitst in twee deelprocessen, die niet met elkaar te maken hebben.

2.5. <u>Markof-matrices met minstens één eigen waarde, ongelijk 1, met modulus 1.</u>

Wij beschouwen een Markof-matrix $P$ , die een eigen waarde $\lambda$ heeft met de eigenschappen:

(1) $\quad |\lambda| = 1 \, , \; \lambda \neq 1 .$

Een eigen vector, behorende bij $\lambda$ , geven wij aan met $f$ .
Evenals in 2.4 normeren wij $f$ weer zo, dat $\|f\| = 1$ en $\exists_i^J f^i = 1$.
Definieer weer:

(2) $\quad S_o \stackrel{def}{=} \mathcal{E}ns \{ i \in J \,|\, f^i = 1 \}$,

dan is $S_o$ dus niet leeg en geldt:

(3) $\quad \begin{cases} i \in S_o \Leftrightarrow f^i = 1 \\ i \in J - S_o \Leftrightarrow f^i \neq 1 , |f^i| \leq 1 . \end{cases}$

Kies $i \in S_o$ :
$\lambda = \lambda f^i = P^i_j f^j$ , dus na deling door $\lambda$ :

$$1 = \frac{1}{\lambda} \sum_1^? P^i_j f^j$$

Ook geldt: $\sum P^i_j = 1$ , waaruit volgt:

(4) $\quad \sum P^i_j \left(1 - \lambda^{-1} f^j\right) = o .$

Nodig hiervoor is weer, dat het reële deel van (4) nul is:

(5) $\quad \sum P^i_j \left(1 - \mathcal{R}e \, \lambda^{-1} f^j\right) = o .$

Nu is $\mathcal{R}e \, \lambda^{-1} f^j \leq |\lambda^{-1} f^j| = |f^j| \leq 1$, dus wij hebben weer te maken met een som van niet negatieve getallen, die 0 moet zijn, zodat iedere term van (5) gelijk 0 moet zijn.

Dus geldt:

$$\forall_i^{S_0} \forall_j^{J} \quad \text{òf } P_j^i = 0 \text{ òf } 1 - \mathcal{R}e\, \lambda^{-1} f^j = 0 .$$

Daar $\mathcal{R}e\, \lambda^{-1} f^j \leqq 1$ is, zal $1 - \mathcal{R}e\, \lambda^{-1} f^j = 0$ dan en slechts dan als $\lambda^{-1} f^j = 1$, dus als $f^j = \lambda$.
Definieer nu:

(6) $$S_1 \stackrel{\text{def}}{=} \mathcal{E}ns \{ j \in J \mid f^j = \lambda \},$$

dan is $S_1$ niet leeg, daar anders uit (5) volgt, dat $\forall_j^{J} P_j^i = 0$ als $i \in S_0$, dus ook $\sum P_j^i = 0$ in strijd met (1.5.10).
Uit het voorgaande volgt nu, dat

(7) $\forall_i^{S_0} \forall_j^{J-S_1} P_j^i = 0$, dus

(8) $\forall_i^{S_0} P_{J-S_1}^i = 0$ en $P_{S_1}^i = 1$.

Wij definiëren nu algemeen:

(9) $$S_\alpha \stackrel{\text{def}}{=} \mathcal{E}ns \{ j \in J \mid f^j = \lambda^\alpha \},\ \alpha \text{ geheel}.$$

De reeds gegeven definities voor $S_0$ en $S_1$ zijn hiermee in overeenstemming. Stel $S_\alpha$ niet leeg voor een vaste $\alpha \in N$.
Kies $i \in S_\alpha$, dan is $f^i = \lambda^\alpha$.

$\lambda^{\alpha+1} = \lambda f^i = P_j^i f^j = \lambda^{\alpha+1} \sum P_j^i$, dus $\sum P_j^i \left(1 - \frac{f^j}{\lambda^{\alpha+1}}\right) = 0$, dus

$\sum P_j^i \left(1 - \mathcal{R}e\, \frac{f^j}{\lambda^{\alpha+1}}\right) = 0$. Weer is $\mathcal{R}e\, \frac{f^j}{\lambda^{\alpha+1}} \leq |f^j| \leqq 1$, dus

$\forall_i^{S_\alpha} \forall_j^{J} P_j^i \left(1 - \mathcal{R}e\, \frac{f^j}{\lambda^{\alpha+1}}\right) = 0$ (geen sommatie)

$\mathcal{R}e\, \frac{f^j}{\lambda^{\alpha+1}} = 1$ dan en slechts dan als $j \in S_{\alpha+1}$, waaruit volgt:

(10) $\forall_i^{S_\alpha} P_{J-S_{\alpha+1}}^i = 0$ dus $P_{S_{\alpha+1}}^i = 1$.

Hieruit volgt direct, dat als $S_\alpha \neq 0$, ook $S_{\alpha+1} \neq 0$. Wij bewezen reeds, dat $S_0$ en $S_1 \neq 0$, dus:

(11) $$\forall_\alpha^{N} S_\alpha \neq 0 .$$

Tevens is duidelijk:

(12) Als $\lambda^\alpha \neq \lambda^\beta$ zal $S_\alpha \cap S_\beta = 0$.

Daar J eindig ondersteld is, moeten tenslotte twee machten van $\lambda$ met verschillende gehele exponenten aan elkaar gelijk zijn. Dus:

(13) $\exists_m \lambda^m = 1$, dus $S_m = S_0$, $m > 0$ minimaal.

Definiëren wij $m$ volgens (13), dan is dus $S_\alpha \cap S_\beta = 0$ als $\alpha \neq \beta$ en $\alpha, \beta < m$. Het is dus een bijzonderheid van eindige Markof-matrices, dat de eigen waarden met modulus 1 eenheidswortels zijn, dus als argument een rationaal veelvoud van $2\pi$ hebben.
Door de nummering der mogelijke toestanden geschikt te kiezen, kunnen wij ervoor zorgen, dat als $\alpha < \beta < m$ geldt $\forall_i^{S_\alpha} \forall_j^{S_\beta}\, i < j$.

Als $m = 4$ is, zal P er nu als volgt uitzien:

(14) $$P = \begin{array}{c} \\ S_0 \\ S_1 \\ S_2 \\ S_3 \\ \\ \end{array} \begin{array}{ccccc} S_0 & S_1 & S_2 & S_3 & \\ \end{array}$$
$$P = \begin{array}{c|c} & \begin{array}{ccccc} S_0 & S_1 & S_2 & S_3 & \end{array} \\ \begin{array}{c} S_0 \\ S_1 \\ S_2 \\ S_3 \\ \phantom{0} \end{array} & \left( \begin{array}{c|c|c|c|c} 0 & * & 0 & 0 & 0 \\ \hline 0 & 0 & * & 0 & 0 \\ \hline 0 & 0 & 0 & * & 0 \\ \hline * & 0 & 0 & 0 & 0 \\ \hline \multicolumn{5}{c}{} \end{array} \right) \end{array}$$

Behalve in het gearceerde gebied, waarover niets te zeggen valt, en de gebieden, aangegeven met een $*$ , zullen de elementen van P nul zijn, hetgeen direct uit (10) volgt.
Noem $S = S_0 + S_1 + \ldots + S_{m-1}$, dan zal het systeem, eenmaal aan gekomen in een tot $S$ behorende toestand, nooit meer een buiten $S$ gelegen toestand kunnen bereiken. $S$ is dus een fuik (zie 2.4)
Bevindt het systeem zich in $S_0$ , dan kan het alleen overgaan naar een toestand van $S_1$ , vandaar alleen naar een toestand van $S_2$ , enz. Wij geven dit kort aan met:

$S_0 \longrightarrow S_1 \longrightarrow S_2 \longrightarrow \ldots\ldots \longrightarrow S_{m-1} \longrightarrow S_m$ en spreken van een <u>kringfuik</u> van de orde $m$.

<u>Definitie</u>:

Een kringfuik van de orde $m$ is een fuik, bestaande uit een geordend $m$-tal deelvzn, waarbij het systeem steeds met wh 1 van een deelvz in de eropvolgende overgaat.

Zijn er bij $\lambda$ $(|\lambda| = 1, \lambda \neq 1)$ nog andere eigen vectoren, lineair onafhankelijk van $f$ , d.w.z. is $\lambda$ meervoudig, dan zullen er meer kringfuiken zijn, alle van dezelfde orde $m$ .

Stel $\{S_0, \ldots\ldots, S_{m-1}\}$ en $\{S'_0, \ldots\ldots, S'_{m-1}\}$ zijn kringfuiken, dan definiëren wij voor $\alpha, \beta \in \{0, \ldots\ldots, m-1\}$ :

(15) $\quad S_{\alpha,\beta} \stackrel{def}{=} S_\alpha \cap S'_\beta$

Stel $S_{\alpha\beta}$ niet leeg, dan is er dus een $i \in S_{\alpha\beta}$

$$i \in S_\alpha \longrightarrow P^{\,i}_{J - S_{\alpha+1}} = 0$$

$$i \in S'_\beta \dashrightarrow P^{\,i}_{J - S'_{\beta+1}} = 0 \qquad \text{dus}$$

(16) $\quad i \in S_{\alpha,\beta} \longrightarrow P^{\,i}_{J - (S_{\alpha+1} \cap S_{\beta+1})} = 0$, dus $P^{\,i}_{J - S_{\alpha+1,\beta+1}} = 0$.

Een systeem, dat zich in een toestand van $S_{\alpha,\beta}$ bevindt, kan alleen maar overgaan naar een toestand van $S_{\alpha+1,\beta+1}$ , dus als $S_{\alpha,\beta}$ niet leeg is, dan is $S_{\alpha+1,\beta+1}$ ook niet leeg. Pas dit toe op $S_{0,j}$ met $j \in \{0, \ldots\ldots, m-1\}$ en $S_{0,j} \neq 0$ :

$$S_{0,j} \longrightarrow S_{1,j+1} \longrightarrow S_{2,j+2} \longrightarrow \ldots\ldots \longrightarrow S_{m-1,j+m-1} \longrightarrow S_{m,j+m} = S_{0,j}.$$

Na $m$ stappen zijn wij weer terug in $S_{0,j}$. Dit geldt voor iedere $j \in \{0, \ldots\ldots, m-1\}$ waarvoor $S_{0,j}$ niet leeg is. Hieruit volgt, dat er hoogstens $m$ niet lege kringfuiken zijn, alle van de orde $m$

Wij hebben steeds $\lambda \neq 1$ ondersteld, doch nergens van deze voorwaarde gebruik gemaakt. Al het bewezene geldt dus ook voor $\lambda = 1$

Dus:

Is $\lambda = 1$ een enkelvoudige eigen waarde, dan is er één kringfuik, bestaande uit één verzameling $(m=1)$ . De eigen vector, bij 1 behorend, is de eenheidsvector, dus $S_0 = J$. Alles dus triviaal. Is $\lambda = 1$ een meervoudige eigen waarde, dan is er een kringfuik, bestaande uit één verzameling $(m=1)$ , waarvoor geldt $S_0 \neq J$. Dit is een niet-trviale fuik. Wij behandelden dit geval in 2.4, alwaar wij tevens bewezen, dat er in dat geval minstens twee niet-triviale fuiken zijn.

Tenslotte behandelen wij het geval, dat $P$ nog een tweede eigen waarde $\mu$ heeft, $\neq 1$ , met modulus 1. De bij $\lambda\,(\lambda \neq 1, |\lambda| = 1)$ behorende eigen vector noemen wij $f$ , de bij $\mu\,(\mu \neq 1, |\mu| = 1)$ behorende $h$ . Wij definiëren:

(17) $\quad \|f\|^S \stackrel{def}{=} \max_{i \in S} |f^i| \qquad \|f\|^J = \|f\|$.

De kringfuik, behorende bij $\lambda$ en $f$ is weer $\{S_0, S_1, \ldots, S_{m-1}\}$ (dus $\lambda^m = 1$)
Kies $i \in S_\alpha$ , $\alpha \in \{0, \ldots\ldots, m-1\}$ :

$\mu h^i = \sum_{j \in J} P^i_j h^j = \sum_{j \in S_{\alpha+1}} P^i_j h^j$, dus

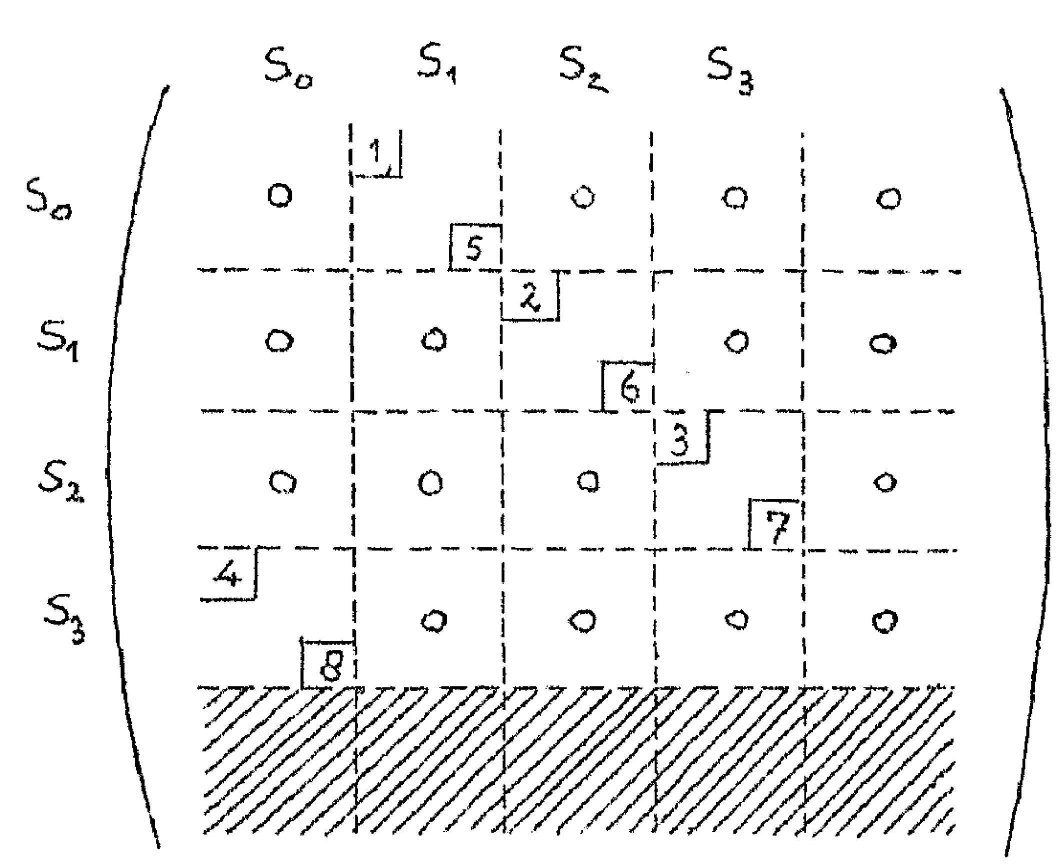
$\forall_i^{S_\alpha} |\mu h^i| = |h^i| \leq \max_{j \in S_{\alpha+1}} |h^j| = \| h \|^{S_{\alpha+1}}$, waaruit volgt:

$\| h \|^{S_\alpha} \leq \| h \|^{S_{\alpha+1}} \qquad \| h \|^{S_0} \leq \| h \|^{S_1} \leq \ldots \leq \| h \|^{S_m} = \| h \|^{S_0}$,

zodat overal het gelijkheidsteken geldt:

(18) $\quad \| h \|^{S_0} = \| h \|^{S_1} = \| h \|^{S_2} = \ldots = \| h \|^{S_{m-1}}$.

Hoewel $h$ dus bij de eigen waarde $\mu$ behoort, zijn toch - mits $|\mu| = 1$ is - de normen van $h$ met betrekking tot de vzn $S_\alpha$, behorende bij de eigen waarden $\lambda$ $(\lambda \neq \mu, |\lambda| = 1)$ gelijk.
Daar $|\mu| = 1$ is, hoort bij $h$ ook een kringfuik.
Is $\| h \|^{S_0} < \| h \|$, dan is deze kringfuik disjunct met $S_0, \ldots, S_{m-1}$.
Als $\| h \|^{S_0} > 0$, bevindt zich in $S_0, \ldots, S_{m-1}$ een kleinere kringfuik.
Immers de door $S = S_0 + S_1 + \ldots + S_{m-1}$ bepaalde deelmatrix van $P$ is weer een Markof-matrix, die weer de eigen waarde $\mu$ heeft met als eigen vector die vector, die als componenten heeft de componenten $h^i$ van $h$, waarvoor $i \in S$.
Bij deze eigen vector behoort nu weer een kringfuik in $S$. De orde van deze kringfuik moet een veelvoud zijn van de orde van de kringfuik, behorende bij eigen vector $f$. Voor $m = 4$ krijgen wij bijvoorbeeld:

(19) $\quad P =$ 

## 2.6. Nadere beschouwingen over eigen waarden met modulus 1.

Reeds enige malen maakten wij gebruik van de stelling, dat bij een Markof-matrix het aantal lineair onafhankelijke eigen vectoren, dat bij de eigen waarde 1 gevonden kan worden, gelijk is aan de multipliciteit van 1 in de karakteristieke veelterm van de matrix. (Wij gebruikten deze stelling namelijk in het voorbeeld van 2.3 en in 2.4.)

In deze paragraaf zullen wij bewijzen, dat deze stelling voor alle eigen waarden met een modulus 1 geldt. Echter alleen voor Markof-matrices, waartoe wij ons, zoals steeds in dit hoofdstuk, beperken.

Stelling 1

Is $P$ een Markof-matrix, $\lambda$ een eigen waarde van $P$ met modulus 1, en zijn $f$ en $g$ vectoren, dan geldt:

$$(1) \qquad \left.\begin{array}{l} Pg = \lambda g + f \\ Pf = \lambda f \\ |\lambda| = 1 \end{array}\right\} \longrightarrow f = 0 .$$

Bewijs: $P^n g = \lambda^n g + n\lambda^{n-1} f$. Dit zien wij in door volledige inductie : $P^{n+1} g = P^n (Pg) = P^n(\lambda g + f) = \lambda P^n g + P^n f =$
$= (\lambda^{n+1} g + n\lambda^n f) + \lambda^n f = \lambda^{n+1} g + (n+1)\lambda^n f$, terwijl het gestelde voor $n=1$ geldt.

Dus $\forall_i^J \forall_n^{N'} (P^n)^i_j\, g^j = \lambda^n g^i + n\lambda^{n-1} f^i$

$$f_i = \frac{(P^n)^i_j\, g^j - \lambda^n g^i}{n\lambda^{n-1}}$$

dus $\forall_i^J \forall_n^{N'} |f^i| \leqq \dfrac{(P^n)^i_j |g^j| + |\lambda|^n |g^i|}{n|\lambda|^{n-1}} \leqq \dfrac{2\|g\|}{n}$,

daar $|\lambda| = 1$ en $(P^n)^i_j \leqq 1$.

Hieruit volgt $\forall_i^J |f^i| = 0$ dus $f = 0$, q.e.d.

Stelling 2

Als $\lambda$ een eigen waarde van de Markof-matrix $P$ is met modulus 1 en $g$ een vector, dan volgt voor iedere $n$ uit $(P-\lambda)^n g = 0$, dat $(P-\lambda) g = 0$ is [1]

Bewijs: Dat de stelling geldt voor $n=2$, volgt direct uit (1). Immers stel $(P-\lambda)^2 g = 0$ en noem $(P-\lambda) g = f$, dan is dus $(P-\lambda) f = 0$ en volgt uit (1), dat $f=0$ is, dus $(P-\lambda) g = 0$ is. Stel nu, dat de stelling reeds bewezen is voor $n-1$. Dus uit $(P-\lambda)^{n-1} f = 0$ volgt, dat $(P-\lambda) f = 0$ is. Stel $(P-\lambda) g = f$, dan is dus $(P-\lambda)^n g = (P-\lambda)^{n-1} f = 0$ dus $(P-\lambda) f = 0$, dus $(P-\lambda)^2 g = 0$, dus $(P-\lambda) g = 0$ volgens het voorgaande.

----------

1) Met $(P-\lambda)$ bedoelen wij de matrix $(P-\lambda I)$

In de appendix (8.4) en (8.5) wordt bewezen, dat elke (vierkante) matrix $M$ te schrijven is in de vorm:

(8.5.12) $\quad M = \sum^{\rho} (\lambda_\rho E_\rho + F_\rho)$,

waarbij de $\lambda_\rho$ 's de eigen waarden van $M$ voorstellen, de $E$ 's orthogonaal en idempotent zijn (zie (8.5.6)), gedefinieerd volgens (8.5.4), terwijl volgens (8.5.5) $\sum E_\rho = I$ is. Verder is (8.5.7) $F_\rho = (M - \lambda_\rho) E_\rho$, de $F_\rho$ 's zijn orthogonaal en nilpotent (zie (8.5.8) en (8.5.11)). Verder geldt (8.5.9) en (8.5.10). Er is dus een kleinste $m$, waarvoor $F_\rho^m = 0$ is. Dezen noemen wij $m_\rho$. Uit 8.5 blijkt, dat $m_\rho$ de multipliciteit voorstelt van de factor $\lambda - \lambda_\rho$ in de minimumveelterm $\psi(\lambda)$ (gedefinieerd als het kleinste polynoom in $\lambda$, waarvoor $\psi(M) = 0$ is, zie 8.4). Er geldt $0 < m_\rho \leqq n_\rho$ (8.4, stelling 2).

Nu geldt voor Markof-matrices de

<u>Stelling 3</u>

Als $P$ een Markof-matrix is en $\lambda_\rho$ een eigenwaarde van $P$ met modulus 1, dan is de bij $\lambda_\rho$ behorende nilpotente matrix $F_\rho$ identiek gelijk 0.

Bewijs:

$F_\rho^{m_\rho} = 0$, dus $(P - \lambda_\rho)^{m_\rho} E_\rho = 0$, dus $\forall_i \forall_h \left((P - \lambda_\rho)^{m_\rho}\right)^i_j \cdot E_{\rho}{}^j_h = 0$.

Nu is voor vaste $h$ $E_{\rho}{}^j_h$ een vector, dus volgt uit stelling 2:

$\forall_i \forall_h (P - \lambda_\rho)^i_j E_{\rho}{}^j_h = 0$, dus $\forall_i \forall_h F_{\rho}{}^i_h = 0$, dus $F_\rho = 0$, q.e.d.

Daar de rang van $E_\rho$ gelijk $n_\rho$ is (zie 8.6, $n_\rho$ is de multipliciteit van $\lambda_\rho$ in de karakteristieke veelterm van $P$), zijn er precies $n_\rho$ kolommen van $E_\rho$ l.o., dus zijn er precies $n_\rho$ vectoren $f_\rho$, waarvoor geldt $(P - \lambda_\rho) f_\rho = 0$.

Hiermee is de aan het begin van deze paragraaf genoemde stelling bewezen, die wij nog even herhalen:

<u>Stelling 4</u>

Bij een Markof-matrix is het aantal l.o. eigen vectoren, dat gevonden kan worden bij een eigen waarde <u>met modulus 1</u> precies gelijk aan de multipliciteit van deze eigen waarde in de karakteristieke veelterm.

Uit $P = \sum^{\rho} (\lambda_\rho E_\rho + F_\rho)$, stelling 3 en de betrekkingen (8.5.6) en (8.5.10) volgt direct: Als $|\lambda_\rho| = 1$ is, geldt $E_\rho P = P E_\rho = \lambda_\rho E_\rho$.

Dus $E_{\rho}{}^i_j P^j_k = \lambda_\rho E_{\rho}{}^i_k$ en

$P^i_j E_{\rho}{}^j_k = \lambda_\rho E_{\rho}{}^i_k$, dus:

Stelling 5

Als $|\lambda_\rho| = 1$ is, geven de rijen van $E_\rho$ de bij $\lambda_\rho$ behorende eigen covectoren; de kolommen van $E_\rho$ geven de bij $\lambda_\rho$ behorende eigen vectoren.

Als de eigen waarde 1 van P enkelvoudig is, dan volgt dus uit het bovenstaande, dat alle rijen van $E_o$ (de bij $\lambda_o = 1$ behorende E ) gelijk zijn. Elke rij geeft de bij 1 behorende eigen covector $g_{oj}$ , zo genormeerd dat $\sum_1^r g_{oj} = 1$ is. Dit laatste blijkt als volgt:

$$E_{o\,j}^{\,i} = 1^i g_{oj}\,,\quad E_o^2 = E_o\,,\ \text{dus}\ \ 1^i g_{oj}\, 1^j g_{ok} = 1^i g_{ok}\,,\ \text{dus}\ \ g_{oj}\, 1^j = 1 = \sum_1^r g_{oj}$$

Uit $F_\rho = 0$ als $|\lambda_\rho| = 1$ is volgt nu (zie 8.6), dat de normaalvorm van een Markof-matrix er bijvoorbeeld als volgt uit zal zien:

$$(2)\qquad P^* = \begin{pmatrix}
\boxed{1} & 0 & & & & & & & & \\
 & \boxed{1} & 0 & & & & & & & \\
 & & \boxed{1} & 0 & & & & & & \\
 & & & \boxed{\lambda_1} & 0 & & & & & \\
 & & & & \boxed{\lambda_1} & 0 & & & & \\
 & & & & & \lambda_2 & 1 & & 0 & \\
 & & & & & & \lambda_2 & 1 & & \\
 & & & & & 0 & & \lambda_2 & 1 & \\
 & & & & & & & & \lambda_2 & \\
 & & & & & & & & & \lambda_3
\end{pmatrix}$$

waarbij $|\lambda_1| = 1, |\lambda_2| \neq 1$ is, en de veelvuldigheid van 1 gelijk 3, van $\lambda_1$ 2 en van $\lambda_2$ gelijk 4 is.

Bovenstaande stellingen gelden alleen voor eigen waarden met modulus 1. Voor eigen waarden met modulus kleiner dan 1, zullen zij in het algemeen niet gelden, zoals blijkt uit het volgende voorbeeld:

Voorbeeld:

Gegeven is een getal $p$, $0 < p < 1$ en de volgende Markof-matrix:

$$P_j^i = \begin{pmatrix} p & 0 & 1-p & 0 \\ 0 & 0 & p & 1-p \\ 0 & 0 & p & 1-p \\ 0 & 0 & 0 & 1 \end{pmatrix}; \quad D(\lambda) = \begin{vmatrix} p-\lambda & 0 & 1-p & 0 \\ 0 & -\lambda & p & 1-p \\ 0 & 0 & p-\lambda & 1-p \\ 0 & 0 & 0 & 1-\lambda \end{vmatrix} = -\lambda(p-\lambda)^2(1-\lambda).$$

Uit $D(\lambda)=0$ vinden wij de eigen waarden $\lambda_0=1, \lambda_1=0, \lambda_2=\lambda_3=p<1$
De minimumveelterm $\psi(\lambda)$ moet de factoren $\lambda$, $p-\lambda$ en $1-\lambda$ bevatten (zie stelling 2 van 8.4). Uitrekenen geeft echter:

$$P(P-pI)(P-I)=\begin{pmatrix} 0 & 0 & -p(1-p)^2 & p(1-p)^2 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \neq 0.$$

Dus $\psi(\lambda)$ is op een constante na gelijk aan $D(\lambda)$.
Er is nu een $f$ en een $\lambda$, zodanig dat $(P-\lambda)^2 f=0, (P-\lambda)f\neq 0$ is.
Neem namelijk $\lambda=p$:

$$(P-pI)^2=\begin{pmatrix} 0 & 0 & 1-p & 0 \\ 0 & -p & p & 1-p \\ 0 & 0 & 0 & 1-p \\ 0 & 0 & 0 & 1-p \end{pmatrix}^2=\begin{pmatrix} 0 & 0 & 0 & (1-p)^2 \\ 0 & p^2 & -p^2 & (1-p)^2 \\ 0 & 0 & 0 & (1-p)^2 \\ 0 & 0 & 0 & (1-p)^2 \end{pmatrix}$$

$$(P-pI)^2 f=\begin{pmatrix} (1-p)^2 f^4 \\ p^2 f^2 - p^2 f^3 + (1-p)^2 f^4 \\ (1-p^2) f^4 \\ (1-p^2) f^4 \end{pmatrix} \quad \text{als } f^4=0 \text{ en } f^2=f^3 \text{ is.}$$

Hieraan voldoet bijvoorbeeld: $f=\begin{pmatrix}1\\1\\0\\0\end{pmatrix}$ ; $(P-pI)f=\begin{pmatrix}1-p\\0\\0\\0\end{pmatrix}\neq\begin{pmatrix}0\\0\\0\\0\end{pmatrix}$.

In dit eenvoudige geval willen wij ook de $E$ 's en de $F$ 's berekenen: (zie voor de methode 8.5)

$$\frac{1}{D(\lambda)}=\frac{1}{-\lambda(p-\lambda)^2(1-\lambda)}=\frac{a}{\lambda}+\frac{b}{\lambda-1}+\frac{c+d\lambda}{(\lambda-p)^2}$$

Door rechts en links met $\lambda(\lambda-1)$ te vermenigvuldigen en daarna $\lambda=0$, resp. $\lambda=1$ in te vullen vinden wij $a=-\frac{1}{p^2}$, $b=\frac{1}{(1-p)^2}$ ; de berekening van $c$ en $d$ geeft meer werk, doch het zal blijken, dat dit te vermijden is.
Als wij nu de $E$, behorende bij de eigen waarde $\lambda$, aangeven met $E^{(\lambda)}$, dan volgt uit (8.5.1), (8.5.2) en (8.5.4):

$$E^{(0)} = -\frac{1}{p^2}(P-I)(P-pI)^2$$

$$E^{(1)} = \frac{1}{(1-p)^2} P(P-pI)^2$$

$$E^{(p)} = (c+dP)\,P(P-I)$$

Door de vermenigvuldigingen uit te voeren vinden wij:

$$E^{(0)} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad E^{(1)} = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Uit $\sum E_\rho = I$ volgt nu:

$$E^{(p)} = \begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix},$$

zodat het inderdaad niet nodig was $c$ en $d$ te berekenen. Controleer, dat de $E$ 's idempotent en orthogonaal zijn!

Daar $F_\rho = (P - \lambda_\rho I)\,E_\rho$ is, vinden wij

$$F^{(0)} = 0, \quad F^{(1)} = 0, \quad F^{(p)} = \begin{pmatrix} 0 & 0 & 1-p & -1+p \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Controleer $F^{(p)2} = 0$ ($p-\lambda$ komt immers twee maal voor in $\psi(\lambda)$)

Uit (8.5.12) volgt nu:

$$P = 0.E^{(0)} + 1.E^{(1)} + p.E^{(p)} + F^{(p)} = \begin{pmatrix} p & 0 & 1-p & 0 \\ 0 & 0 & p & 1-p \\ 0 & 0 & p & 1-p \\ 0 & 0 & 0 & 1 \end{pmatrix}, \text{hetgeen klopt.}$$

## 2.7. Asymptotisch gedrag van Markof-matrices.

In het algemeen bestaat $\lim\limits_{n\to\infty} (P^n)^i_j$ niet. Dit blijkt direct als wij $P = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ nemen. In 2.3 bewezen wij, dat deze limiet wel bestaat als alle eigen waarden enkelvoudig zijn en behalve 1 geen eigen waarden met modulus 1 voorkomen. Wij merkten toen reeds op, dat de multipliciteit van de eigen waarden met modulus kleiner dan 1 niet ter zaken doet, zodat wij thans zullen bewijzen:

Stelling 1:

$\lim_{n\to\infty} (P^n)^i_j$ bestaat als de eigen waarde 1 enkelvoudig is en eigen waarden $\neq 1$ met modulus 1 niet voorkomen en is gelijk $g_{oj}$

Bewijs:

$$P = E_o + \sum_{|\lambda_\rho|<1}^{\rho} (\lambda_\rho E_\rho + F_\rho)$$

$$P^n = E_o + \sum_{|\lambda_\rho|<1}^{\rho} \sum_{o}^{m-1}{}^{k} \binom{n}{k} \lambda_\rho^{n-k} F_\rho^{\,k}, \qquad \text{waarbij } m = \max_{|\lambda_\rho|<1} m_\rho \text{ en}$$

$F_\rho^{\,m_\rho} = 0$ (zie 8.5.15).

Nu is $(F_\rho^{\,k})^i_j$ begrensd, dus $\forall_k \forall_i \forall_j \left|(F_\rho^{\,k})^i_j\right| \leq A$,

daar $F$ nilpotent is. Door $n$ voldoende groot te kiezen, kunnen wij er voor zorgen, dat als $0 \leq k \leq m-1$ is:

$\left|\binom{n}{k}\lambda_\rho^{n-k}\right| \leq \frac{\varepsilon}{m\tau A}$ ($\varepsilon$ willekeurig $> 0$ ) (immers $\lim_{n\to\infty} n^k a^n = 0$ als $|a|<1$ voor elke $k$ ).

Dus:

$$\left|(P^n)^i_j - E_{oj}^{\,i}\right| \leq \tau m \frac{\varepsilon}{m\tau A} A = \varepsilon, \quad \text{waaruit volgt:}$$

$\lim_{n\to\infty} (P^n)^i_j = E_{oj}^{\,i}$, dus ook $= g_{oj}$ , daar $E_{oj}^{\,i} = 1^i g_{oj}$ in dit geval (zie 2.6).

Ook in andere gevallen kan $\lim (P^n)^i_j$ bestaan, als $P$ een speciale vorm heeft. Neem bijvoorbeeld $P = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$, dan is $\lim P^n = P$. 1)

De limiet is nu niet gelijk aan $g_{oj}$ , dus nu geldt niet, dat de wh om na een groot aantal spelen in de toestand $j$ te verkeren onafhankelijk is van de begintoestand.

Altijd bestaat de Cesarolimiet: 2)

---

1) $\lim_{n\to\infty} P^n = A$ wil zeggen, dat $\forall_i^{\exists} \forall_j^{\exists} \lim_{n\to\infty} (P^n)^i_j = A^i_j$.

2) De Cesarolimiet van een rij getallen $a_1, a_2, \ldots$ wordt gedefinieerd als $\lim_{n\to\infty} \frac{a_1 + a_2 + \ldots + a_n}{n}$, wanneer deze limiet bestaat. Bestaat $\lim_{n\to\infty} a_n = a$, dan bestaat de Cesarolimiet ook en is gelijk aan de gewone limiet (dus $= a$). Dit bewijzen wij door aan te tonen, dat de vorm $\left|a - \frac{a_1 + a_2 + \ldots + a_n}{n}\right|$ voor voldoend grote $n$ willekeurig weinig van 0 verschilt. Kies hiertoe $\varepsilon > 0$ willekeurig en $N$ zodanig, dat voor $n > N$ geldt: $|a - a_n| < \frac{\varepsilon}{3}$. Wij kiezen $n$ zo groot, dat bovendien geldt $\frac{N|a|}{n} < \frac{\varepsilon}{3}$ en $\left|\frac{a_1 + a_2 + \ldots + a_N}{n}\right| < \frac{\varepsilon}{3}$.

Nu is $\left|a - \frac{a_1 + a_2 + \ldots + a_n}{n}\right| = \left|\frac{Na}{n} - \frac{a_1 + a_2 + \ldots + a_N}{n} + \frac{(a - a_{N+1}) + \ldots + (a - a_n)}{n}\right| <$

$< \frac{N|a|}{n} + \frac{|a_1 + a_2 + \ldots + a_N|}{n} + \frac{n-N}{n}\frac{\varepsilon}{3} < \varepsilon$, dus $\lim_{n\to\infty} \frac{a_1 + a_2 + \ldots + a_n}{n} = a$.

Stelling 2:

Als $P$ een (eindige) Markof-matrix is, bestaat $\lim_{n\to\infty} \frac{P+P^2+\ldots+P^n}{n} = Q$, waarbij $Q$ ook een Markof-matrix is. En wel geldt: $Q = E_0$ , de bij eigen waarde 1 behorende idempotente matrix. Bewijs:

$$Q^{(n)}{}^i_j \stackrel{def}{=} \frac{1}{n}(P+P^2+\ldots\ldots+P^n)^i_j\ , \quad 0 \leqq Q^{(n)}{}^i_j \leqq 1.$$

Wij nemen $i$ en $j$ vast en $n$ variabel. Er is nu een deelrij met indices $n_1, n_2 \ldots\ldots\ldots$ , zodanig, dat $\forall_i^3 \forall_j^3 \lim_{\nu\to\infty} Q^{(n_\nu)}{}^i_j$ bestaat. Deze limiet noemen wij $R^i_j$ .

Er is namelijk een deelrij, zodanig, dat $\lim Q^{(n)}{}^1_1$ bestaat. Van deze deelrij kunnen wij weer een deelrij nemen, waarvoor ook $\lim Q^{(n)}{}^1_2$ bestaat. Hiervan weer een deelrij, zodanig, dat ook $\lim Q^{(n)}{}^1_3$ bestaat. Na eindig veel stappen convergeert zo voor alle $i$ en $j$ de deelrij van $Q^{(n)}{}^i_j$ . De indices van de zo verkregen deelrij noemen wij $n_1, n_2, \ldots\ldots$

$$Q^{(n_\nu)} = \frac{P+P^2+\ldots\ldots+P^n}{n_\nu} \longrightarrow R \quad \text{als} \quad n_\nu \longrightarrow \infty.$$

Daar $PQ^{(n_\nu)} = Q^{(n_\nu)}P$ is ook $PR = RP$ , $P$ en $R$ zijn verwisselbaar.

$$Q^{(n_\nu)} - PQ^{(n_\nu)} = \frac{P}{n_\nu} - \frac{P^{n_\nu+1}}{n_\nu} \longrightarrow R - PR, \text{ terwijl}$$

$$\frac{P^i_j}{n_\nu} \longrightarrow 0 \quad \text{en} \quad \frac{(P^{n_\nu+1})^i_j}{n_\nu} \leqq \frac{1}{n_\nu} \longrightarrow 0.$$

dus $\quad PR = RP = R.$

Ook $\quad P^k R = RP^k = R \quad$ en als $\varphi(P)$ een willekeurig polynoom in $P$ voorstelt:

$$\varphi(P)R = R\varphi(P) = \varphi(1)R.$$

Als buiten de rij $n_1, n_2, \ldots\ldots$ nog oneindig veel natuurlijke getallen voorkomen, dan is daaruit weer een deelrij van getallen $m_1, m_2, \ldots\ldots\ldots$ te vormen zodanig, dat de rij

$$S^{(m_\nu)} = \frac{1}{m_\nu}(P+P^2+\ldots\ldots+P^{m_\nu})$$

convergent is. De limietmatrix noemen wij $S$ . $S^{(m_\nu)}$ is een polynoom in $P$, $S^{(m_\nu)}(1) = 1$ dus $S^{(m_\nu)}R = RS^{(m_\nu)} = R$, dus ook $SR = RS = R$. Omgekeerd geldt ook wegens de symmetrie $RS = SR = S$ zodat $R = S$.

Iedere convergente deelrij nadert tot dezelfde matric, die wij Q zullen noemen, zodat ook de gehele rij naar Q convergeert. Daar elke $Q^{(n)}$ een Markof-matrix is $(\Sigma\, Q^{(n)}{}^{i}_{j} = 1)$, is ook Q een Markof-matrix.
Wij moeten nu nog bewijzen $Q = E_0$:

$$P = E_0 + \sum_{\substack{|\lambda_\rho|=1 \\ \lambda_\rho \neq 1}}^{\rho} \lambda_\rho E_\rho + \sum_{|\lambda_\rho|<1}^{\rho} (\lambda_\rho E_\rho + F_\rho) \;;$$

$$P^n = E_0 + \sum_{\substack{|\lambda_\rho|=1 \\ \lambda_\rho \neq 1}}^{\rho} \lambda_\rho^n E_\rho + \sum_{|\lambda_\rho|<1}^{\rho} \sum_{0}^{m-1}{}^{k} \binom{n}{k} \lambda_\rho^{\,n-k} F_\rho^{\,k} \,,$$

waarbij $m = \max\limits_{|\lambda_\rho|<1} m_\rho$ en $F_\rho^{\,m_\rho} = 0$ is (zie 8.5.15).

$$Q^{(n)} = E_0 + \sum_{|\lambda_\rho|=1}^{\rho} \frac{\lambda_\rho + \dots + \lambda_\rho^n}{n} E_\rho + \frac{1}{n} \sum_{|\lambda_\rho|<1}^{\rho} \sum_{0}^{m-1}{}^{k} \sum_{1}^{n}{}^{h} \binom{h}{k} \lambda_\rho^{\,h-k} F_\rho^{\,k} \,.$$

Nu is $$\sum_{1}^{n}{}^{h} \binom{h}{k} \lambda_\rho^{\,h-k} = \sum_{1}^{\infty}{}^{h} \binom{h}{k} \lambda_\rho^{\,h-k} - \sum_{n+1}^{\infty}{}^{h} \binom{h}{k} \lambda_\rho^{\,h-k} = \frac{1}{(1-\lambda_\rho)^{k+1}} - \sum_{+1}^{\infty}{}^{h} \binom{h}{k} \lambda_\rho^{\,h-k}.$$

~~Door n voldoende~~ groot te kiezen, kunnen wij voor elke $k$
$\sum_{n+1}^{\infty}{}^{h} \binom{h}{k} \lambda_\rho^{\,h-k}$ zo weinig van 0 doen verschillen als wij willen, zodat voor de derde term van het rechterlid van (1) te schrijven is $\frac{1}{n} \sum\limits_{|\lambda_\rho|<1} \sum\limits_{0}^{m-1}{}^{k} F_\rho^{\,k} \frac{1}{(1-\lambda_\rho)^{k+1}} + \varepsilon$
voor $n > n(\varepsilon)$

Dit nadert tot 0 als $n \to \infty$ ($m$ constant, $\varepsilon$ willekeurig $> 0$ )
De tweede term van het rechterlid van (1) bestaat uit een eindig aantal termen, die ieder nul zijn als $n \equiv 0\ (\ell_\rho)$ [1] is, waarbij $\ell_\rho$ de kleinste waarde $\neq 0$ is, zodanig, dat $\lambda_\rho^{\ell_\rho} = 1$ is. Is $n \not\equiv 0 (\ell_\rho)$, maar bijvoorbeeld $a\ell_\rho < n < (a+1)\ell_\rho$ , $a$ geheel, dan blijft een rest over van de vorm $\frac{1}{n}(\lambda_\rho^{\ell_\rho+1} + \dots + \lambda_\rho^{n})$ en dit nadert tot 0 als $n \to \infty$, daar het aantal termen tussen de haken nooit meer dan $\ell_\rho - 1$ zal bedragen. Dit geldt voor elke $\lambda_\rho$ met $|\lambda_\rho| = 1$ zodat de hele tweede term van (1) tot 0 nadert. Dus

$$\lim_{n \to \infty} Q^{(n)} = E_0 \,, \quad \text{q.e.d.}$$

------------

1) $a \equiv 0 (b)$ wil zeggen: $a$ is deel aar door b.
$a \equiv 1 (b)$ " " $a$ is niet deel aar door b.

Wij bewezen dus:

(2) $$\lim_{n\to\infty} \frac{P+P^2+\ldots\ldots+P^n}{n} = E_0$$

Gevolg: Daar bij convergente rijen de Cesarolimiet gelijk is aan de gewone limiet, zal als $\lim (P^n)^i_j$ bestaat ook gelden

$$\lim_{n\to\infty} (P^n)^i_j = (E_0)^i_j .$$

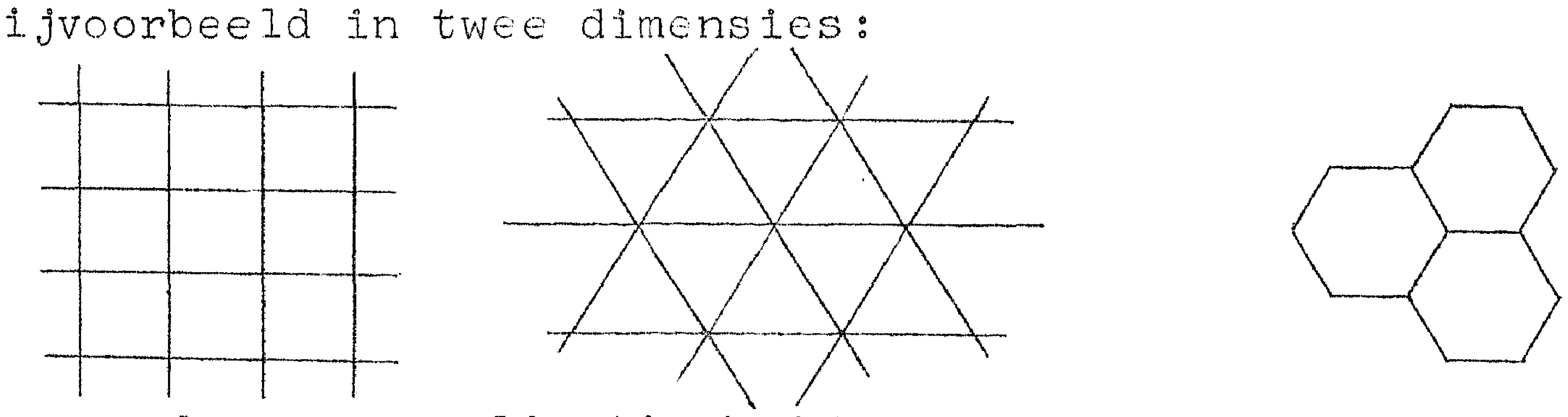
## 2.8. De stochastische wandeling

Als toepassing van het tot nu toe behandelde beschouwen wij de "stochastische wandeling" (Engels: random walk).
Gegeven is een discrete puntrij, bijvoorbeeld de punten op de reële getallenrechte, overeenkomende met de gehele getallen.
De stochastische wandeling bestaat nu hieruit, dat een of ander object, het wandelende, bewegende of springende punt, steeds met dezelfde whn van het punt, waarin het zich bevindt naar het erop volgende of eraan voorafgaande springt.

Noemen wij $P^i_j$ de wh, dat het bewegende punt in een sprong van punt $i$ naar punt $j$ springt (geformaliseerd, dat het systeem van toestand $i$ naar toestand $j$ overgaat), dan moeten voor de $P^i_j$ de volgende eigenschappen gelden:

(1)
1. $P^i_j$ hangt alleen af van het verschil $i-j$
2. $P^i_j = 0$ als $i-j \neq \pm 1$.

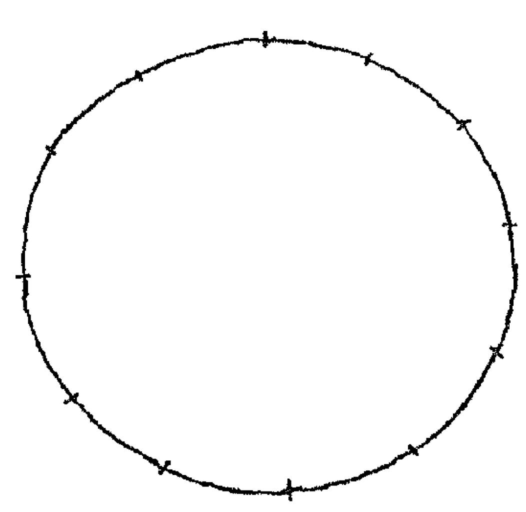
Verschillende generalisaties zijn hiervan mogelijk:
In twee en meer dimensies kunnen wij verschillende soorten roosters beschouwen, waarbij wij steed twee punten met elkaar verbinden, wanneer het bewegende punt in één sprong van het ene naar het andere punt kan komen.
Bijvoorbeeld in twee dimensies:



Een verdere generalisatie krijgen wij door toe te laten, dat het wandelende punt ook grotere afstanden dan 1 per keer aflegt. Wij handhaven dan alleen eis 1 van (1) en krijgen een sprongproces of "stochastische dans".
Voor al deze weinig gecompliceerde processen is de tot dusverre ontwikkelde theorie al niet meer voldoende. De vz van mogelijke toestanden is namelijk oneindig. Wij kunnen in het ééndimensionale geval hieraan op verschillende manieren tegemoet komen.

1) Het aantal punten nemen wij eindig. Voor de uiteinden kan dan gelden:
   a) De uiteinden zijn fuiken. Het bewegende punt, eenmaal in een der uiteinden gekomen, kan niet meer terug, doch blijft steeds in hetzelfde punt. Wij spreken van absorberende grenzen. Op de hiermee samenhangende absorptieproblemen komen wij in hoofdstuk 4 terug.
   b) In een uiteinde aangekomen moet het punt terug. Is $a$ dus het rechteruiteinde, dan is $P^{a}_{a-1} = 1$ . Wij spreken van elastische of reflecterende wanden.

   In beide gevallen gaat de symmetrie verloren.
2) Deze symmetrie kunnen wij handhaven door $\tau$ aequidistante punten op een cirkel te beschouwen, dus de hoekpunten van een regelmatige $\tau$ -hoek.



Cyclische stochastische wandeling.

*fig.* 1.

Bij de gewone stochastische wandeling geldt voor de elementen van de matrix van overgangswhn:

$$(2)\qquad P^i_j = p\,I^{i+1}_j + q\,I^{i-1}_j\,, \quad (p+q=1\,,\ pq \geqq 0)\,,$$

zodat $P$ oneindig is, geheel uit nullen bestaat, behalve elementen $q$ direct links en elementen $p$ direct rechts van de hoofddiagonaal. Bij een cyclische stochastische wandeling nemen wij $i+\tau=i$, dus beschouwen wij de indices modulo $\tau$ . Indices, die een $\tau$ -voud verschillen, beschrijven dezelfde toestand.
De matrix der overgangswhn en de karakteristieke vergelijking worden nu:

$$(3)\ P = \begin{array}{c} \\ 1 \\ 2 \\ \vdots \\ \\ \\ \tau \end{array} \begin{array}{c} 1 \quad 2 \cdots\cdots \tau \\ \begin{pmatrix} 0 & p & \cdots & & q \\ q & 0 & p & & \\ & q & 0 & p & \\ & & \ddots & \ddots & \ddots \\ & & q & 0 & p \\ p & \cdots & & q & 0 \end{pmatrix} \end{array} \ ; \qquad D(\lambda) = \begin{vmatrix} -\lambda & p & \cdots & & q \\ q & -\lambda & p & & \\ & q & -\lambda & p & \\ & & \ddots & \ddots & \ddots \\ & & q & -\lambda & p \\ p & \cdots & & q & -\lambda \end{vmatrix} .$$

De eigen waarden kunnen wij vinden door $D(\lambda)=0$ op te lossen, hetgeen veel rekenwerk geeft. Eenvoudiger is de volgende methode:

$$Pf=\lambda f,\ \text{dus} P^i_j f^j=\lambda f^i;\ \text{dus} \lambda f^i = p f^{i+1}+q f^{i-1},\quad i\in\mathfrak{I},\ f^{i+\tau}=f^i.$$

Dit is een differentievergelijking, die slechts voor bepaalde waarden van $\lambda$ oplosbaar is. Ter oplossing stellen wij:

$f^i=\omega^i$ ($\omega$ tot de macht $i$ ; links van het gelijkteken is $i$ dus een index, rechts een exponent)

$\lambda\omega^i = p\omega^{i+1}+q\omega^{i-1}$, dus na deling door $\omega^i$:

$\lambda = p\omega + q\omega^{-1}$

Verder moet gelden $\omega^\tau=1$ , $\omega$ is een $\tau$-de eenheidswortel, zodat wij voor $\omega$ moeten substitueren:

$$\omega_\rho = e^{\frac{2\pi i\rho}{\tau}},\quad \rho\in\{0,1,\ldots\ldots,\tau-1\}.$$

Bij $\omega_\rho$ behoort $\lambda_\rho = p\omega_\rho + q\omega_\rho^{-1}$ , zodat wij de $\tau$ eigen waarden gevonden hebben.
Wanneer zullen twee eigen waarden aan elkaar gelijk zijn?
Stel hiertoe $\lambda_\rho=\lambda_\sigma$, $\sigma\neq\rho$ ; $\sigma,\rho\in\{0,1,..,\tau-1\}$, dus

$$p\omega_\rho + q\omega_\rho^{-1} = p\omega_\sigma + q\omega_\sigma^{-1},\ \text{dus}\ p(\omega_\rho-\omega_\sigma)=q(\omega_\sigma^{-1}-\omega_\rho^{-1}) = q\,\frac{\omega_\rho-\omega_\sigma}{\omega_\rho\omega_\sigma},$$

of na deling door $\omega_\rho-\omega_\sigma$ (mogelijk daar altijd $\omega_\rho\neq\omega_\sigma$ ):

$$p=\frac{q}{\omega_\rho\omega_\sigma},\quad \frac{q}{p}=\omega_\rho\omega_\sigma.$$

Dit kan alleen als $\frac{q}{p}$ eenheidswortel is, waaruit volgt $q=p=\frac{1}{2}$. Dit is dus een <u>noodzakelijke</u> voorwaarde voor het optreden van gelijke eigen waarden. Stel $p=q$ , dan $\lambda_\rho=\cos\frac{2\pi\rho}{\tau}$ , zodat bijvoorbeeld $\lambda_1$ en $\lambda_{\tau-1}$ gelijk zijn, evenzo $\lambda_2$ en $\lambda_{\tau-2}$ . Enkelvoudig zijn nu dus $\lambda_0=1$ en als $\tau$ even is, $\lambda_{\frac{1}{2}\tau}=-1$ . Alle overige eigen waarden hebben de multipliciteit 2. $\lambda=1$ is dus altijd enkelvoudig, ook als $p=q$ is. Het in 2.4 behandelde geval kan zich hier dus nimmer voordoen. Als $p\neq q$ is, zijn dus alle $\lambda$'s verschillend en geldt: $D(\lambda)=\prod_0^{\tau-1}(\lambda-\lambda_\rho)$.

<u>Het asymptotische gedrag</u>

In het algemeen geldt volgens (8.5.13) en (8.5.15):

$$P^n = \sum^\rho (\lambda_\rho E_\rho + F_\rho)^n = \sum^\rho \sum_0^{m_\rho-1} {}^k \binom{n}{k}\lambda_\rho^{n-k}F_\rho^{\,k}.$$

Als $n\to\infty$ zal ook $n-k\to\infty$ ($k$ begrensd), dus voor grote $n$ gaan de termen met $|\lambda_\rho|=1$ steeds meer overheersen, daar voor Markof-matrices reeds bewezen is $|\lambda|\leqq 1$. Dus:

$$P^n \approx \sum_{|\lambda_\rho|=1}\rho \sum_{0}^{m_\rho-1} k \binom{n}{k} \lambda_\rho^{n-k} F_\rho^k \qquad \text{voor grote } n.$$

Stellen wij nu $p=\frac{1}{2}(1+u)$, $q=\frac{1}{2}(1-u)$, dan $0\leqq|u|\leqq 1$ en

$$\lambda_\rho = pe^{\frac{2\pi i\rho}{\tau}} + qe^{-\frac{2\pi i\rho}{\tau}} = \cos\frac{2\pi\rho}{\tau} + iu\sin\frac{2\pi\rho}{\tau}$$

$$|\lambda_\rho|^2 = \cos^2\frac{2\pi\rho}{\tau} + u^2\sin^2\frac{2\pi\rho}{\tau} = 1-(1-u^2)\sin^2\frac{2\pi\rho}{\tau}.$$

Dus zal $|\lambda_\rho|=1$, $\lambda_\rho\neq 1$ alleen voorkomen als $\tau$ even en $\rho=\frac{1}{2}\tau$ of als $u^2=1$ is.

Verschillende gevallen:

1) $\left.\begin{array}{l} 0<|u|<1 \\ \tau\equiv 1\,(2)\end{array}\right\}$

Wij hebben nu nergens last mee: $\lambda_0=1$, $|\lambda_\rho|\neq 1$ als $\rho\neq 0$ is. Dus $\lim\limits_{n\to\infty}(P^n)^i_j = g_{0j}$ bestaat (zie stelling 1 van 2.7), waarbij voor de $g_0$ geldt:

$$g_0 P = g_0 \,, \quad g_{0i} P^i_j = g_{0j} \,, \quad \text{dus} \quad g_{0j} = p\, g_{0\,j-1} + q\, g_{0\,j+1}.$$

Een niet triviale oplossing hiervan is: $\forall_j^{\mathcal{J}}\; g_{0j}=\frac{1}{\tau}$. Daar de oplossing ondubbelzinnig bepaald is, is dit ook de enige. Uit

$\lim\limits_{n\to\infty}(P^n)^i_j=\frac{1}{\tau}$ volgt nu:

Wandelen veel punten stochastisch rond de cirkel, dan zullen zij op den duur gelijkmatig over de cirkel verdeeld zijn, ook, wanneer ze dat oorspronkelijk niet waren.

2) $\left.\begin{array}{l} 0<|u|<1 \\ \tau\equiv 0\,(2)\end{array}\right\}$

Nu is $\lambda_{\frac{1}{2}\tau}=-1$ een tweede eigen waarde met modulus 1, waarbij dus een kringfuik hoort van de orde 2 ($\lambda^2_{\frac{1}{2}\tau}=1$). Wij hebben dus twee deelsystemen $S$ en $T$ met altijd overgang $S\to T\to S\to\ldots\ldots$ Deze deelsystemen zijn natuurlijk de vzn van alle even en die van alle oneven genummerde toestanden van $\mathcal{J}$. Nu kan $(P^n)^i_j$ geen limiet hebben.

3) $|u|=1$. Het punt loopt nu steeds dezelfde kant uit. Wij hebben een gedetermineerd proces. Als stochastisch proces gezien is er een enkele kringfuik van de orde $\tau$.

3. Klassificatie der mogelijke toestanden (zie ook [30], hoofdstuk 15).

3.1. Definities

In hoofdstuk 2 beperkten wij ons uitsluitend tot eindige (stationaire) Markof-ketens. Bij de in dit hoofdstuk te geven definities en stellingen is deze beperking niet noodzakelijk. Derhalve zullen wij thans aannemen, dat de vz van mogelijke toestanden aftelbaar oneindig is. Deze generalisatie heeft voor de in hoofstuk 2 behandelde theorie der Markof-matrices vèrgaande gevolgen, waar in hoofstuk 5 op teruggekomen zal worden. Komen echter in dit hoofdstuk oneindige sommen voor, dan zij steeds stilzwijgend ondersteld, dat deze convergeren. Bij eindige ketens zijn de sommaties ook eindig en is dus geen convergentievoorwaarde nodig. De stochastische grootheid $\underline{i}_n$ zal weer de mogelijke toestanden aangeven na $n$ overgangen (zie 1.6).
Wij voeren nu in:

(1) $$w(i) \overset{def}{=} \mathcal{P}\left\{ \exists_n^{N'} \underline{i}_n = i \mid \underline{i}_0 = i \right\}.$$

Dus is $w(i)$ de wh, dat het systeem, uitgaande van de toestand $i$, ooit weer in de toestand $i$ zal terugkeren. Daar de Markofketen stationair is, geldt ook:

(2) $$w(i) = \mathcal{P}\left\{ \exists_n^{N'} \underline{i}_{n+k} = i \mid \underline{i}_k = i \right\}.$$

In plaats van $w(i)$ zullen wij ook wel alleen $w$ schrijven.

Definities:

a) Wij noemen $i$ een doorgangstoestand (Engels: transient state) als $w(i) < 1$, dus $1 - w(i) > 0$. Er is dus een positieve ontsnappingskans uit de toestand $i$, een positieve kans, dat het systeem nooit meer in de toestand $i$ terugkomt.

b) Wij noemen $i$ een periodieke toestand met periode $m$ als:
   1) $(P^n)_i^i = 0$, tenzij $n \equiv 0 \ (m)$
   2) $m > 1$
   3) $m$ maximaal.

   Dus de kans, dat het systeem na $n$ stappen terugkeert, is 0, tenzij $n$ een $m$-voud is.

c) Wij noemen $i$ een terugkeertoestand (recurrent) als $w(i)=1$ is. Het systeem komt zeker in de toestand $i$ terug. Deze terugkeer zal dan ook oneindig vaak geschieden. De kans, dat het systeem slechts een eindig aantal malen in de toestand $i$ terug zal keren, is dus gelijk 0.
Wanneer wij met de stochastische grootheid $\underline{t}$ het aantal overgangen aangeven, waarna het systeem voor het eerst weer in de toestand $i$ teruggekeerd is, en de verwachting van $\underline{t}$ -dat is de gemiddelde terugkeertijd - aangeven met $\mu$, $\mu=\mathcal{E}\underline{t}$, dan kunnen wij nog twee gevallen onderscheiden als $i$ recurrent is:
 1) $\mu<\infty$. Wij noemen de toestand $i$ ergodisch.
 2) $\mu=\infty$. Wij noemen de toestand $i$ een nultoestand.

De gevallen a) en c) sluiten elkaar uit. Dit is niet zo met de gevallen b) en c): een periodieke toestand kan bijvoorbeeld ergodisch zijn. Ook de gevallen a) en b) hoeven elkaar niet uit te sluiten.

3.2. Doorgangstoestanden

Wij definiëren verder:

$$(1)\qquad \begin{cases} u_n \overset{\text{def}}{=} (\mathcal{P}^n)_i^i = \mathcal{P}\{\underline{i}_n=i \mid \underline{i}_0=i\} \\ u_0 \overset{\text{def}}{=} 1 \end{cases}$$

Dus $u_n$ is de kans, dat het systeem na $n$ overgangen zich weer in de toestand $i$ bevindt. Dit hoeft dan niet voor het eerst te zijn. Het systeem kan reeds een of meer keren eerder in de toestand $i$ terug geweest zijn. De $u_n$'s hangen uiteraard ook van de begintoestand $i$ af, zodat wij zouden moeten schrijven $u_n(i)$. Dit doen wij terwille van de eenvoud in de notatie echter niet. De $u_n$'s zijn whn, dus:

$$(2)\qquad \forall_n^N \; 0 \leq u_n \leq 1$$

De kans, dat het systeem na $n$ overgangen voor het eerst weer in $i$ terugkeert, noemen wij $w_n$, dus:

$$(3)\qquad w_n \overset{\text{def}}{=} \mathcal{P}\{\underline{t}=n \mid i\} = \mathcal{P}\{\underline{i}_n=i,\ \underline{i}_1\neq i,\ldots\ldots,\underline{i}_{n-1}\neq i \mid \underline{i}_0=i\}.$$

Uit de definities volgt:

$$(4)\qquad u_n = w_1u_{n-1}+w_2u_{n-2}+\ldots+w_{n-1}u_1+w_nu_0=\sum_1^n{}_k\, w_k u_{n-k}=\sum_0^n{}_k\, w_k u_{n-k} \quad (n\geq 1)$$

Als (5) $w_0 \overset{def}{=} 0$.

Gegeven de $w$'s zijn de $u$'s hieruit door recursie te berekenen. Dit geldt ook omgekeerd, daar de coëfficiënt van $w_n$ gelijk 1 is. Nu geldt:

$$\sum_0^\infty{}^n u_n = u_0 + \sum_1^\infty{}^n u_n = 1 + \sum_1^\infty{}^n \sum_1^n{}^k w_k u_{n-k} = 1 + \sum_1^\infty{}^k w_k \sum_k^\infty{}^n u_{n-k} =$$
$$= 1 + \sum_1^\infty{}^k w_k \sum_0^\infty{}^m u_m = 1 + w \sum_0^\infty{}^m u_m,$$

daar

(6) $$\sum_1^\infty{}^k w_k = w,$$

waarbij $w$ gedefinieerd is volgens (3.1.1).
Dus:

(7) $$\left(\sum_0^\infty{}^n u_n\right)(1-w) = 1.$$

Uit (7) volgt:
a) $\sum_0^\infty{}^n u_n < \infty \Longleftrightarrow w < 1$, dus $i$ is doorgangstoestand. De toestand $i$ is dan en slechts dan transient als $\sum_0^\infty u_n < \infty$.
Convergeert $\sum_0^\infty u_n$, dan moet $\lim_{n\to\infty} u_n = 0$ zijn, zodat als $i$ transient is, $\lim u_n = 0$ is. Dit geldt natuurlijk niet omgekeerd.
b) Divergentie van $\sum_0^\infty u_n$ is nodig en voldoende voor recurrentie van de toestand $i$.
Afzonderlijk beschouwen wij het periodieke geval, waarbij dus $u_n = 0$, tenzij $n \equiv 0\,(m)$, $m > 1$ en bewijzen de stelling:
$u_n = 0$, tenzij $n \equiv 0\,(m) \Longleftrightarrow w_n = 0$ tenzij $n \equiv 0\,(m)$.
Bewijs: Wij voeren in:

(8) $$\begin{cases} U(z) \overset{def}{=} \sum_0^\infty{}^n u_n z^n. \\ W(z) \overset{def}{=} \sum_0^\infty{}^n w_n z^n. \end{cases}$$

Hierbij is $z$ een hulpvariabele. De reeksen convergeren altijd als $|z| < 1$.

$$U(z) = \sum_0^\infty{}^n u_n z^n = 1 + \sum_1^\infty{}^n \sum_1^n{}^k w_k u_{n-k} z^n =$$
$$= 1 + \sum_1^\infty{}^k w_k z^k \sum_k^\infty{}^n u_{n-k} z^{n-k} = 1 + W(z)\,U(z),$$

dus :

(9) $$U(z) = \frac{1}{1 - W(z)}.$$

Is dus $U(z)$ een machtreeks in $z^m$ (is dus $u_n = 0$ tenzij $n \equiv 0\,(m)$ ), dan zal ook $W(z)$ een machtreeks in $z^m$ zijn, waaruit volgt $w_n = 0$ tenzij $n \equiv 0\,(m)$ . Dit geldt ook omgekeerd, q.e.d.
Wij kunnen nu schrijven:

$$(10) \qquad u_{nm} = \sum_{1}^{n} {}_k\, w_{km}\, u_{(n-k)m} \,,$$

welke betrekking van precies dezelfde vorm is als (4).
In het algemeen bestaat $\lim\limits_{n\to\infty} u_n$ nu niet.
Noemen wij namelijk een deelrij met indices $n_\nu \not\equiv 0\,(m)$, dan zal $\forall_\nu\, u_{n_\nu} = 0$ dus $\lim u_{n_\nu} = 0$.
Nemen wij daarentegen een deelrij met indices $\nu_m$ , dan zal deze deelrij bijvoorbeeld de limiet 0 hebben, als $i$ een doorgangstoestand voor het $m$-voudig herhaalde proces is Wij verkeren immers nu in hetzelfde geval met betrekking tot $P^m$ als oorspronkelijk met betrekking tot $P$ . Dus als $\lim\limits_{n\to\infty} P^{mn} = 0$ is, dan is $\lim\limits_{n\to\infty} u_{nm} = 0$ en bestaat $\lim u_n$. In het algemeen is dit dus niet zo.

### 3.3. Het recurrente geval

Wij sluiten nu het periodieke geval met periode $> 1$ uit. Hieruit volgt dus, dat de indices van de $w$'s $\neq 0$ de g.g.d. 1 hebben. Reeds eindig veel indices, $a_1, \ldots\ldots, a_k$ zullen dan deze eigenschap hebben: $(a_1, \ldots\ldots, a_k) = 1$ , $w_{a_1} \neq 0, \ldots\ldots\ldots, w_{a_k} \neq 0$.
Nu geldt de <u>stelling</u>:
Als $i$ een recurrente, niet periodieke toestand is, dan geldt

$$\lim_{n\to\infty} u_n = \frac{1}{\mu}.$$

Hierbij is $\mu = \mathcal{E}\underline{t}$ de verwachte terugkeertijd (zie 3.1). De stelling geldt zowel als $\mu$ eindig als ook wanneer $\mu$ oneindig is.
Bewijs: Wij zullen het bewijs van Feller op de voet volgen (zie [30] , p. 259-261)
Stel:

$$(1) \qquad r_n = w_{n+1} + w_{n+2} + \ldots\ldots = \sum_{n+1}^{\infty} {}_k\, w_k .$$

---

1) $(a_1, \ldots\ldots\ldots, a_k) \overset{def}{=}$ Grootste gemene deler van $a_1, \ldots\ldots, a_k$.
$[a_1, \ldots\ldots, a_k] \overset{def}{=}$ Kleinste gemene veelvoud van $a_1, \ldots, a_k$.

Daar $\sum_{1}^{\infty k} w_k = 1$ zal

(2) $\lim_{n\to\infty} \tau_n = 0$, $\tau_0 = 1$.

(3) $\mu = w_1 + 2w_2 + 3w_3 + \ldots\ldots = \tau_0 + \tau_1 + \tau_2 + \ldots\ldots = \sum_{0}^{\infty n} \tau_n$.

Ook geldt:

(4) $w_n = \tau_{n-1} - \tau_n$.

Wij substitueren (4) in (3.2.4):

$$u_n + \sum_{1}^{n k} \tau_k u_{n-k} = \sum_{1}^{n k} \tau_{k-1} u_{n-k}$$

dus daar $\tau_0 = 1$ is:

$$\sum_{0}^{n k} \tau_k u_{n-k} = \sum_{1}^{n k} \tau_{k-1} u_{n-k} = \sum_{0}^{n-1\, k} \tau_k u_{n-1-k}.$$

Deze uitdrukking heeft dus voor alle $n$ dezelfde waarde, namelijk $\tau_0 u_0$ (neem $n=1$):

(5) $\forall_n \sum_{0}^{n k} \tau_k u_{n-k} = 1$ (want $\tau_0 = u_0 = 1$).

Daar $0 \leqq u_n \leqq 1$ bestaat altijd $\limsup_{n\to\infty} u_n = \lambda$, waarbij $0 \leqq \lambda \leqq 1$ [1]
Er moet dus minstens een deelrij $n_1, n_2, \ldots\ldots$ van de natuurlijke getallen bestaan, zodanig dat $\lim_{\nu\to\infty} u_{n_\nu} = \lambda$ is, terwijl verder bij elke $\varepsilon > 0$ een $m(\varepsilon)$ te vinden is, zodanig, dat voor $n > m(\varepsilon)$ geldt $u_n \leqq \lambda + \varepsilon$. Het aantal $u_n$'s met $u_n \geqq \lambda + \varepsilon$ is namelijk eindig, daar een oneindig aantal $u_n$ 's met $(\lambda+\varepsilon) \leqq u_n \leqq 1$ een ophopingspunt $> \lambda$ zou hebben, waarnaar een deelrij zou convergeren, zodat $\lambda$ niet de limes superior zou zijn.
Is $u_{n_1}, u_{n_2}, \ldots$ nu een naar $\lambda$ convergerende deelrij en is $j$ een natuurlijk getal, waarvoor $w_j > 0$ is, dan zal gelden:

(6) $\lim_{\nu\to\infty} u_{n_\nu - j} = \lambda$.

Om dit te bewijzen, kiezen wij een $j$ met $w_j > 0$ en nemen wij aan, dat het gestelde niet juist is. Dan kan er dus een deelrij van de getallen $n_\nu$ gevonden worden $n_{m_1}, n_{m_2}, \ldots$ zodanig, dat $\lim_{\nu\to\infty} u_{n_{m_\nu} - j} < \lambda$ is. Deze deelrij is zo te kiezen, dat voor alle termen geldt $u_{n_{m_\nu} - j} \leqq \lambda' < \lambda$.
Kies $\varepsilon > 0$. Door eventueel eindig veel termen weg te laten, zorgen wij, dat geldt:

(7) $\forall_\nu\ \lambda - \varepsilon \leqq u_{n_{m_\nu}}$ en $\forall_n\ u_n \leqq \lambda + \varepsilon$.

----------

1) De limes superior van een rij getallen is het grootste getal, dat limiet is van een deelrij van die getallen.

Verder kiezen wij $N$ zo, dat $\tau_n \leqq \varepsilon$ als $n \geqq N$ en $N \geqq j$ is, en nemen $n \geqq N$ .

$$u_n = w_1 u_{n-1} + \dots + w_N u_{n-N} + w_{N+1} u_{n-N-1} + \dots + w_n u_0 \leqq$$

$$\leqq w_1 u_{n-1} + \dots + w_N u_{n-N} + w_{N+1} + w_{N+2} + \dots + w_n \leqq w_1 u_{n-1} + \dots + w_N u_{n-N} + \varepsilon .$$

Dit passen wij toe voor $n = n_{m_\nu}$ , waarbij wij de termen $w_j u_{n_{m_\nu}-j}$ apart beschouwen: $w_j u_{n_{m_\nu}-j} \leqq \lambda' w_j$ .
Verder is $w_k u_{n_{m_\nu}-k} \leqq (\lambda + \varepsilon) w_k$, zodat met behulp van (7) geldt:

$$\lambda - \varepsilon \leqq u_{n_{m_\nu}} \leqq \lambda' w_j + (\lambda + \varepsilon)(w_1 + \dots + w_{j-1} + w_{j+1} + \dots + w_N) + \varepsilon \leqq$$
$$\leqq \lambda' w_j + (\lambda + \varepsilon)(1 - w_j) + \varepsilon \leqq \lambda + 2\varepsilon - (\lambda - \lambda') w_j .$$

Voor iedere $\varepsilon > 0$ zou dus gelden $(\lambda - \lambda') w_j \leqq 3\varepsilon$ en $w_j > 0$ . Dit is onmogelijk, waarmee (6) is aangetoond.
Neem dus een $j$ , waarvoor $w_j > 0$ , dan zal dus:

$$u_{n_\nu - j} \to \lambda ; \; u_{n_\nu - 2j} \to \lambda ; \; u_{n_\nu - 3j} \to \lambda, \text{ enz.}$$

Wij onderscheiden verder twee gevallen:
1. $w_1 > 0$
Dan nemen wij $j = 1$ : $u_{n_\nu - 1} \to \lambda$ ; $u_{n_\nu - 2} \to \lambda$ .
Wij kiezen een bepaalde N en verder $n_\nu \geqq N$ .
Uit (5) volgt:

(8) $$1 \geqq \tau_0 u_{n_\nu} + \tau_1 u_{n_\nu - 1} + \dots + \tau_N u_{n_\nu - N} .$$

dus daar $u_{n_\nu - k} \to \lambda$ als $k \leqq N$ : $1 \geqq (\tau_0 + \dots + \tau_N)\lambda$ .
Dit geldt voor iedere $N$ , dus ook $1 \geqq (\tau_0 + \tau_1 + \dots)\lambda = \mu\lambda$, dus

(9) $$\limsup_{n\to\infty} u_n \leqq \frac{1}{\mu} .$$

Vervolgens definiëren wij $\liminf_{n\to\infty} u_n = \gamma$ . Er is weer een deelrij $u_{n_\nu}$ , die naar $\gamma$ convergeert, terwijl op dezelfde manier te bewijzen is: $u_{n_\nu - j} \to \gamma$ voor elke vaste $j$ , waarvoor $w_j > 0$ is. Kies weer $\varepsilon > 0$ en $N$ zodanig, dat $\tau_N \leqq \varepsilon$ als $n \geqq N$ is, dan is uit (5) af te leiden:

$$1 \leqq (\tau_0 + \dots + \tau_N)\gamma + \varepsilon \text{ dus } \gamma \geq \frac{1 - \varepsilon}{\tau_0 + \dots + \tau_N} , \qquad \text{dus ook}$$

(10) $$\liminf_{n\to\infty} u_n \geqq \frac{1}{\mu} .$$

Uit (9) en (10) volgt nu: $\lim_{n\to\infty} u_n = \frac{1}{\mu}$ .

2.
Wij maken gebruik van de volgende stelling uit de elementaire getallentheorie:
Als $a_1, \ldots\ldots, a_k$ natuurlijke getallen zijn en $(a_1, \ldots\ldots, a_k) = 1$ is terwijl hoogstens één der $a$'s $= 1$ is, dan kan ieder natuurlijk getal $M$, dat groter is dan het product $a_1 a_2 \ldots a_k$ geschreven worden in de vorm:

$M = m_1 a_1 + m_2 a_2 + \ldots\ldots + m_k a_k$, waarbij de $m$'s geheel en positief zijn.[1)]

Volgens het gegeven bestaan er getallen $a_1, \ldots\ldots, a_k$ zodanig, dat $(a_1, \ldots\ldots, a_k) = 1$ en $w_{a_1} \neq 0, \ldots\ldots, w_{a_k} \neq 0$ is. Als $M$ een natuurlijk getal is en $M \geqq a_1 a_2 \ldots a_k$, dan zal op grond van het bovenstaande nu ook gelden $w_M > 0$.
Dus als $\lim\limits_{\nu \to \infty} u_{n_\nu} = \lambda$ is, dan zal ook $\lim\limits_{\nu \to \infty} u_{n_\nu - M} = \lambda$ zijn.
Om de ongelijkheid (8) te krijgen, moeten wij nu (5) toepassen met $n = n_\nu - a_1 a_2 \ldots\ldots a_k$. Het verdere bewijs blijft hetzelfde.

-----------

1) Bewijs: Wanneer een der $a$'s $= 1$, noemen wij die $a_1$. Het is eenvoudig te bewijzen, dat, wanneer de stelling geldt voor $k-1$ $a$'s en $k > 2$, de stelling ook geldt voor $k$ $a$'s. Immers kies $k > 2$ en een $M \geqq a_1 \ldots\ldots a_k + 1$. Dan is $M - a_k \geqq (a_1 \ldots a_{k-1} - 1)(a_k - 1) + a_1 \ldots a_{k-1} + a_1 \ldots a_{k-1} > a_1 \ldots a_{k-1}$ daar $a_k \neq 1$ als $k > 1$ is. Geldt de stelling dus voor $k-1$, dan is $M - a_k = \sum\limits_1^{k-1} m_i a_i$ dus $M = \sum\limits_1^{k} m_i$ met $m_k = 1$. Het is dus voldoende de stelling te bewijzen voor $k = 2$:
Wij noemen $a_1 = a$ en $a_2 = b$ en kiezen een $M > ab$.
Daar $(a, b) = 1$ is 1, dus ook $M$ te schrijven als een lineaire combinatie van $a$ en $b$ met gehele, niet noodzakelijk positieve coëfficiënten, dus $M = pa + qb$, $p \in I$, $q \in I$. Als hierbij $p > 0$ en $q > 0$ is zijn wij klaar. Het geval $p \leqq 0$ en $q \leqq 0$ kan niet voorkomen, zodat wij ons kunnen beperken tot het geval $p > 0 \geqq q$.
Dan geldt $ab < pa + qb \leqq pa$, dus $p > b$, zodat wij kunnen schrijven $p = kb + r$ met $k > 0$, $0 < r \leqq b$, $k$ en $r$ geheel, dus:

$$M = pa + qb = ra + (q + ka)b, \; r \leqq b, \text{ dus}$$

$$ab < M \leqq ab + (q + ka)b, \text{ waaruit volgt } q + ka < 0$$

. Hiermee is de stelling bewezen.

Bewezen is dus:

(11) $\lim\limits_{n\to\infty} u_n = \frac{1}{\mu}$ als $i$ recurrent en aperiodiek is.

Recapitulerende volgt uit 3.2 en 3.3

$$(12)\quad \begin{cases} \left.\begin{array}{l}\sum u_n = \infty \\ \lim u_n \neq 0\end{array}\right\} \leftrightarrows \text{toestand } i \text{ is ergodisch, aperiodiek.} \\ \left.\begin{array}{l}\sum u_n = \infty \\ \lim u_n = 0\end{array}\right\} \leftrightarrows i \text{ is een nultoestand.} \\ \left.\begin{array}{l}\sum u_n < \infty \\ [\lim u_n = 0]\end{array}\right\} \leftrightarrows i \text{ is een doorgangstoestand.} \end{cases}$$

3.4. Bereikbaarheid.

Definitie: Wij noemen de toestand $j$ vanuit de toestand $i$ bereikbaar, wanneer de wh, dat het systeem vanuit $i$ na een eindig aantal overgangen in $j$ komt, positief is, dus wanneer $\exists_n (P^n)^i_j > 0$.

Stelling

Alle vanuit een recurrente toestand bereikbare toestanden zijn eveneens recurrent met dezelfde periode (die natuurlijk gelijk aan 1 kan zijn). Vanuit een nultoestand kunnen slechts nultoestanden bereikt worden.

Is de Markof-keten eindig, dan kunnen nultoestanden in het geheel niet voorkomen, terwijl niet alle toestanden transiënt kunnen zijn.

Bewijs: Laat $i$ een recurrente toestand zijn en laat $j$ vanuit $i$ bereikbaar zijn. De kleinste $n$, waarvoor geldt $(P^n)^i_j > 0$ noemen wij $n_1$:

(1) $\alpha \overset{\text{def}}{=} (P^{n_1})^i_j > 0$.

Er is nu ook een $n$ te vinden, waarvoor $(P^n)^j_i > 0$ is. Immers als $\forall_n (P^n)^j_i = 0$ is, dan zou er een positieve ontsnappingskans $\geqq \alpha$ uit de toestand $i$ zijn, dus zou $i$ niet recurrent zijn, hetgeen wij onderstelden. Dus $\exists_n (P^n)^j_i > 0$.

De kleinste $n$, waarvoor dit geldt, geven wij aan met $n_2$:

(2) $\beta \overset{\text{def}}{=} (P^{n_2})^j_i > 0$.

Nu geldt voor alle $n$:

$$(3)\quad (P^{n_1+n+n_2})^i_i = \textstyle\sum_{h,k} (P^{n_1})^i_h (P^n)^h_k (P^{n_2})^k_i \geqq (P^{n_1})^i_j (P^n)^j_j (P^{n_2})^j_i = \alpha\beta (P^n)^j_j .$$

$$(4) \qquad (P^{n_2+n+n_1})^j_j = \sum^{h,k} (P^{n_2})^j_h (P^n)^h_k (P^{n_1})^k_j \geqq (P^{n_2})^j_i (P^n)^i_i (P^{n_1})^i_j = \alpha\beta (P^n)^i_i .$$

Hieruit volgt, dat $(P^n)^i_i$ en $(P^n)^j_j$ asymptotisch hetzelfde gedrag hebben. Daar $\sum^n (P^n)^i_i = \infty$, is ook $\sum^n (P^n)^j_j = \infty$, dus ook $j$ recurrent.
Is $i$ een recurrente nultoestand, dan volgt uit (3) en (4) tevens, dat ook $j$ een nultoestand moet zijn. (Immers uit $\lim\limits_{n\to\infty} (P^n)^i_i = 0$ volgt $\lim\limits_{n\to\infty} (P^n)^j_j = 0$.)
Laat nu $i$ periodiek zijn met de periode $m$, dus $(P^n)^i_i = 0$, tenzij $n \equiv 0\,(m)$ is. Er is een positieve kans $\geqq \alpha\beta$, dat het systeem na $n_1+n_2$ overgangen weer in de toestand $i$ is teruggekeerd, dus dat $(P^{n_1+n_2})^i_i > 0$ is, waaruit wij vinden $n_1+n_2 \equiv 0\,(m)$.
Nu is $(P^{n_1+n+n_2})^i_i = 0$, tenzij $n_1+n+n_2 \equiv 0\,(m)$ is, dus tenzij $n \equiv 0\,(m)$ is, zodat uit (3) volgt $(P^n)^j_j = 0$, tenzij $n \equiv 0\,(m)$ is, zodat de toestand $j$ een periode $\geqq m$ heeft. Uit (4) volgt, dat de periode van $i$ minstens die van $j$ is, dus de toestanden $i$ en $j$ hebben dezelfde periode $m$.
Wij definiëren

$$(5) \qquad w^{(n)}{}^i_j \overset{def}{=} \mathcal{P}\left\{ \underline{i}_1 \neq j, \ldots\ldots, \underline{i}_{n-1} \neq j,\ \underline{i}_n = j \;\middle|\; \underline{i}_0 = i \right\},$$

dus $w^{(n)}{}^i_j$ is de wh, dat $n$ het kleinste aantal overgangen is, waarin het systeem vanuit $i$ in $j$ zal komen (als $i=j$ zal $w^{(n)}{}^i_i = w_n$, waarbij $w_n$ gedefinieerd is volgens (3.2.3).
Duidelijk is, dat geldt:

$$(6) \qquad (P^n)^i_j = \sum_1^n{}^k\, w^{(k)}{}^i_j\, (P^{n-k})^j_j$$

en

$$(7) \qquad \sum_1^\infty{}^n\, w^{(n)}{}^i_j \leqq 1,$$

want $\sum_1^\infty{}^n w^{(n)}{}^i_j$ is de totale wh, dat het systeem van $i$ naar $j$ overgaat. Laat $j$ nu een doorgangstoestand of een nultoestand zijn, dan weten wij reeds $\lim\limits_{n\to\infty} (P^n)^j_j = 0$. Wij kiezen nu $\varepsilon > 0$ en verder $n_0(\varepsilon)$ en $n_1(\varepsilon)$ zo groot, dat voor alle $n \geqq n_0(\varepsilon)$ geldt $(P^n)^j_j \leqq \varepsilon$ en dat verder $\sum_{n_1+1}^\infty{}^n w^{(n)}{}^i_j \leqq \varepsilon$ is. Verder nemen wij $n \geqq n_0 + n_1$, dan is dus ook voor alle $k \leqq n_1$ $n-k \geqq n_0$ en geldt onder gebruikmaking van (6) :

$$(P^n)^i_j = \sum_1^{n_1}{}^k w^{(k)}{}^i_j (P^{n-k})^j_j + \sum_{n_1+1}^{n}{}^k w^{(k)}{}^i_j (P^{n-k})^j_j \leqq$$
$$\leqq \sum_1^{n_1}{}^k w^{(k)}{}^i_j \,.\, \varepsilon + \sum_{n_1+1}^{n}{}^k w^{(k)}{}^i_j \,.\, 1 \leqq \varepsilon + \varepsilon = 2\varepsilon,$$

zodat wij vinden:

(8) $\quad \forall_i^{\mathcal{J}} \lim_{n\to\infty} (P^n)_j^i = 0.$

Is $\mathcal{J}$ nu eindig, dan kunnen niet alle toestanden nultoestand of doorgangstoestand zijn. Anders zou immers bij een vaste $i$ voor alle $j \in \mathcal{J}$ gelden $\lim_{n\to\infty} (P^n)_j^i = 0$, hetgeen als $\mathcal{J}$ eindig is, in strijd is met $\sum (P^n)_j^i = 1$.
Is $i$ een nultoestand, dan zijn volgens het voorgaande alle vanuit $i$ bereikbare toestanden eveneens nultoestanden, zodat $\lim_{n\to\infty} (P^n)_j^i = 0$ is voor alle $j$, uit $i$ bereikbaar.
Is $j$ daarentegen niet bereikbaar uit $i$, dan is $\forall_n (P^n)_j^i = 0$, dus ook $\lim_{n\to\infty} (P^n)_j^i = 0$, dus $\forall_j^{\mathcal{J}} \lim_{n\to\infty} (P^n)_j^i = 0$, hetgeen - als $\mathcal{J}$ eindig is - in strijd is met de eigenschap $\sum (P^n)_j^i = 1$, zodat $i$ geen nultoestand kan zijn.
Hiermee is de stelling geheel bewezen.

3.5. Klassificatie der toestanden bij Markof-ketens.

In 2.4. definieerden wij een fuik als een vz $S \subset \mathcal{J}$ met de eigenschap, dat als $i \in S$ en $j \in \mathcal{J}-S$ zijn, geldt $P_j^i = 0$. Dan is tevens $\forall_n (P^n)_j^i = 0$, hetgeen intuïtief duidelijk, doch ook door inductie te bewijzen is:

$$(P^{n+1})_j^i = (P^n)_j^i P_j^k = \sum_{k\in S} (P^n)_k^i P_j^k + \sum_{k\in \mathcal{J}-S} (P^n)_k^i P_j^k = 0,$$

daar wij de inductie-veronderstelling $(P^n)_k^i = 0$ voor $k \in \mathcal{J}-S$ maken, terwijl voor $k \in S$ geldt $P_j^k = 0$. Verder geldt het gestelde voor $n = 1$.

Stelling 1

Bij een irreducibele eindige Markof-keten zijn alle toestanden ergodisch met dezelfde periode. Doorgangstoestanden kunnen niet voorkomen.

Bewijs: Een irreducibele keten bevat geen niet triviale fuiken. Dus is elke toestand vanuit elke andere toestand bereikbaar, want stel, dat er een $i$ en een $j$ gevonden konden worden, zodanig, dat $\forall_n (P^n)_j^i = 0$ is, dan zou $j$ niet behoren tot de vz $S$ van toestanden, die vanuit $i$ wel bereikbaar zijn. Deze vz $S$ is nu stochastisch afgesloten, want als $k \in S$ en $\ell \in \mathcal{J}-S$ is, dan geldt $\forall_n (P^n)_\ell^k = 0$. Ware dit niet zo, dan zou $\ell$ immers vanuit $k$ en dus vanuit $i$ bereikbaar zijn, dus zou $\ell \in S$ zijn.

Daar de vz $S$ niet leeg is (immers $\sum_j P_j^i = 1$ , zodat minstens één toestand $j$ vanuit $i$ bereikbaar moet zijn) en $\dot{P}$ irreducibel is, volgt hieruit $S = \mathfrak{J}$ , dus alle toestanden zijn vanuit elke andere toestand bereikbaar.
Uit 3.4 volgt nu, dat, als een toestand ergodisch is, alle vanuit die toestand bereikbare toestanden ergodisch zijn met dezelfde periode en dat nultoestanden niet voor kunnen komen. Daar niet alle toestanden transiënt kunnen zijn, dus minstens één ergodisch volgt uit de bereikbaarheid van alle toestanden, dat bij een irreducibele eindige keten alle toestanden ergodisch zijn met dezelfde periode, q.e.d.

<u>Stelling 2</u>

Bij een algemene Markof-keten is de vz van alle recurrente toestanden te verdelen in een aantal onderling disjuncte stochastisch afgesloten deelvzn, waarbij in elke deelvz een irredubel Markof-proces plaats vindt en toestanden, gelegen in dezelfde deelvz, dezelfde periode hebben, welke periode voor verschillende deelvzn niet gelijk behoeft te zijn.
Bewijs: $i \in \mathfrak{J}$. $i$ recurrent; $S_i \overset{def}{=} \mathrm{Ens}\{k \mid k$ bereikbaar uit $i\}$,
dan zijn alle $k \in S_i$ recurrent met dezelfde periode (zie 3.4).
$S_i$ is afgesloten. Als ook $j \in \mathfrak{J}$ en $j$ recurrent is, dan is òf $S_i = S_j$ , òf $S_i \cap S_j = 0$, want stel $k \in S_i \cap S_j$, dan is dus $i$ bereikbaar vanuit $k$ (want $i$ recurrent) en ook $k$ bereikbaar vanuit $j$ , dus $j$ bereikbaar vanuit $i$, waaruit volgt $j \in S_i$ , dus $S_j \subset S_i$ . Evenzo $S_i \subset S_j$, dus $S_i = S_j$ of $S_i \cap S_j = 0$.
Voor $\mathfrak{J}$ kunnen wij dus schrijven:
$\mathfrak{J} = S_1 + S_2 + \ldots + S_m + R$ , waarbij $R$ de vz van alle doorgangstoestanden is. Uit $R$ kan minstens een der vzn $S_1, \ldots, S_m$ bereikt worden. Anders zou $R$ namelijk nog afgesloten vzn van recurrente toestanden bevatten.
Daar elke $S_i$ afgesloten is en vanuit elke toestand uit $S_i$ elke andere van $S_i$ bereikbaar is, vindt in elke $S_i$ dus een irreducibel Markof-proces plaats.

# 4. Absorptieproblemen.

## 4.1. Inleiding.

In dit hoofdstuk zullen wij een der klassieke problemen van de whr behandelen, namelijk het probleem van de ruïnering der spelers.

Zoals bekend wordt de correspondentie tussen Blaise Pascal en Pierre de Fermat in 1654 als het begin van de whr beschouwd. In 1657 verschijnt dan een boek van Christiaan Huygens: De ratiociniis in ludo aleae, Van Rekeningh in Spelen van Geluck. Hierin stelt Huygens vijf problemen, waarvan het derde als volgt luidt: A en B hebben elk 12 penningen en spelen met 3 dobbelstenen onder de volgende condities: Als 11 ogen geworpen worden moet A één penning aan B geven, worden daarentegen 14 ogen geworpen, dan moet B één penning aan A geven. Hij wint het spel, die voor het eerst alle penningen heeft. Hoe verhouden zich de winstkansen van A en B? Huygens geeft voor de verhouding der kansen van A en B: 244.140.625: 282.429.536.481, doch beredeneert zijn oplossing niet. Dit wordt gedaan door J Bernoulli in zijn boek Ars Conjectandi, 1713, posthuum.

Het vraagstuk is verder behandeld door P. de Montmort (Essai d'analyse sur les jeux de hasard, Paris (1708) en A de Moivre (De Mensura Sortis, 1708; Doctrine of Chances, 1718 (2e druk 1738)

De Moivre beschouwt ook de wh, dat een der spelers vóór of tijdens het $n$ -de spel geruïneerd zal zijn. Dit heet "Duration of Play". Bovendien neemt hij de beginkapitalen niet gelijk. Hij geeft twee regels, die juist zijn, doch die hij niet bewijst. Dit geschiedt door P.S. de Laplace omstreeks 1770 met behulp van de methode van de voortbrengende functies.

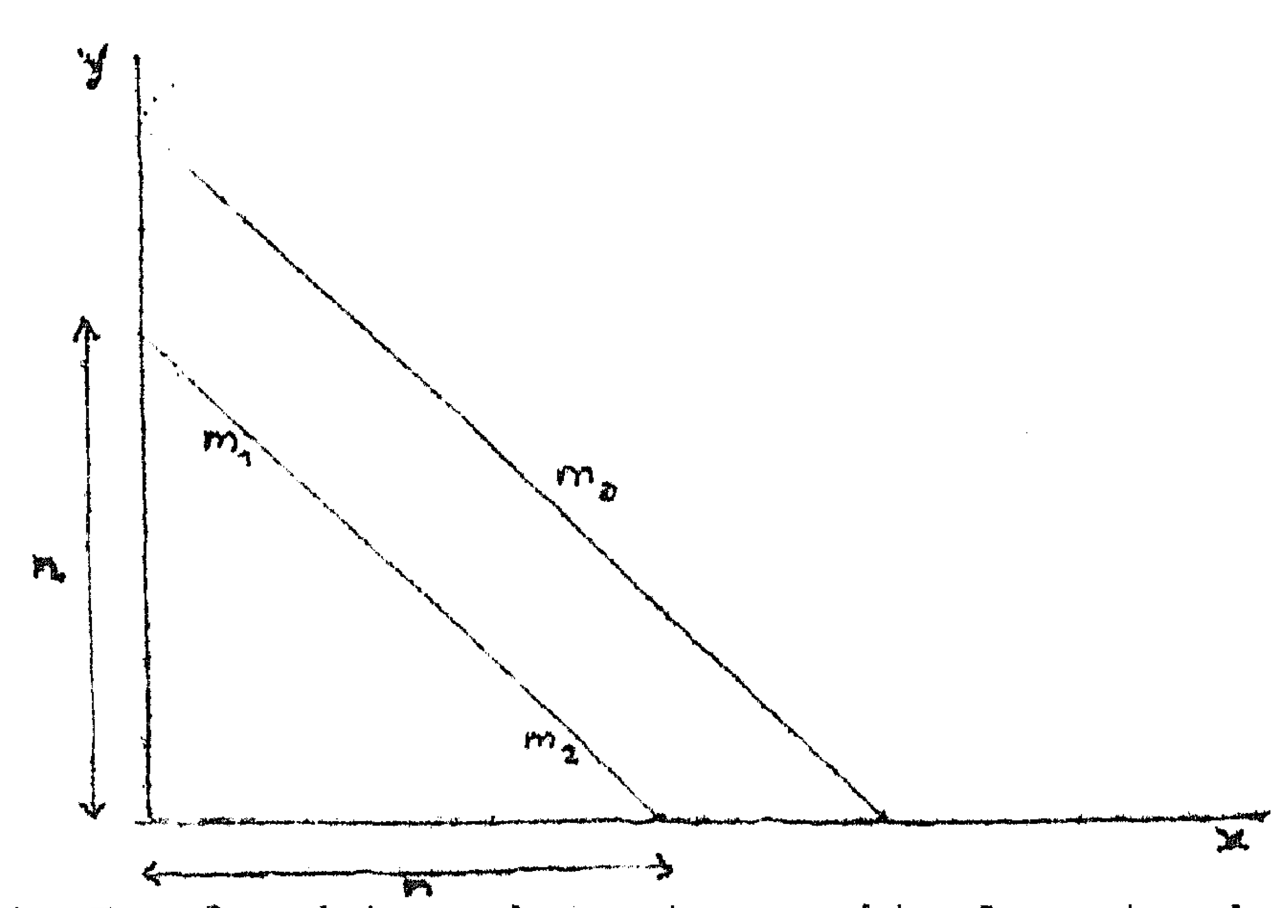
Beschouwt worden 2 spelers met beginbedragen $a$ en $b$ , $a+b=s$. Gedurende het hele spel blijft s constant. Bij elke partij heeft de eerste speler een kans $p$ om één munteenheid te winnen en een kans $q=1-p$ om er één te verliezen, Na verloop van enige partijen zal de eerste speler een bedrag $x$ bezitten, de tweede $s-x$ . De overgangswhn zijn dus $p$ voor de overgang $x \to x+1$ en $q$ voor $x \to x-1$ zolang $1 \leqq x \leqq s-1$ . De randvoorwaarden zijn

$0 \to 0$ met wh 1

$s \to s$ met wh 1

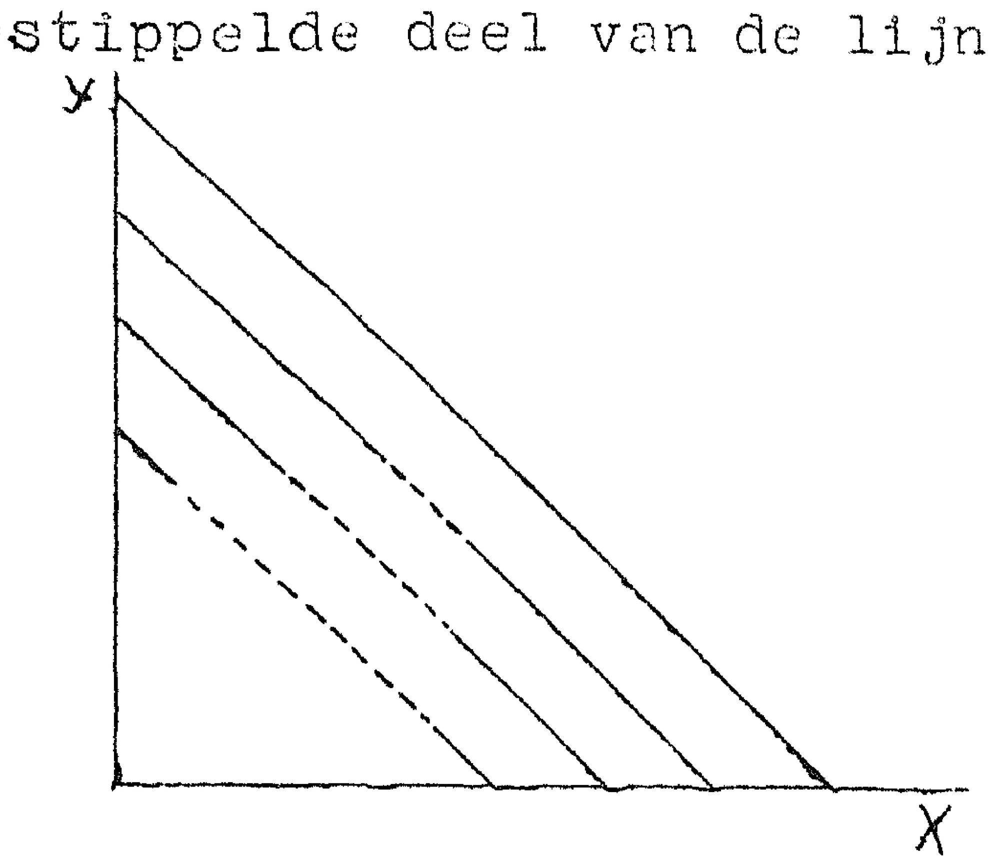
Gevraagd wordt de verwachte duur van het spel of liever de wh-verdeling van de duur van het spel.

Het duurt vervolgens tot 1945 voor het probleem weer ter sprake komt onder de algemene naam absorptieprobleem. De bedragen, welke de spelers bij winst en verlies aan elkaar moeten betalen, worden niet meer gelijk genomen, doch ondersteld wordt, dat het bedrag, dat de ene speler kan winnen (en de andere dus verliezen) een veelvoud is van het bedrag, dat de andere speler kan winnen.
Dit probleem is namelijk van belang voor de door A. Wald in Amerika en G.A. Barnard in Engeland gedurende de tweede wereldoorlog ontwikkelde sequente analyse.
De sequente analyse wordt gebruikt bij de keuring van massaproducten. Vroeger nam men voor dit doel één steekproef en verwierp de partij, wanneer de fractie ondeugdelijke exemplaren een bepaald percentage overtrof. Deze methode is vaak tijdrovend, omdat teveel objecten nodig zijn. Een verbetering is het dubbelsteekproefschema, opgesteld door Dodge en Romig. Ter illustratie hiervoor dient de volgende figuur:

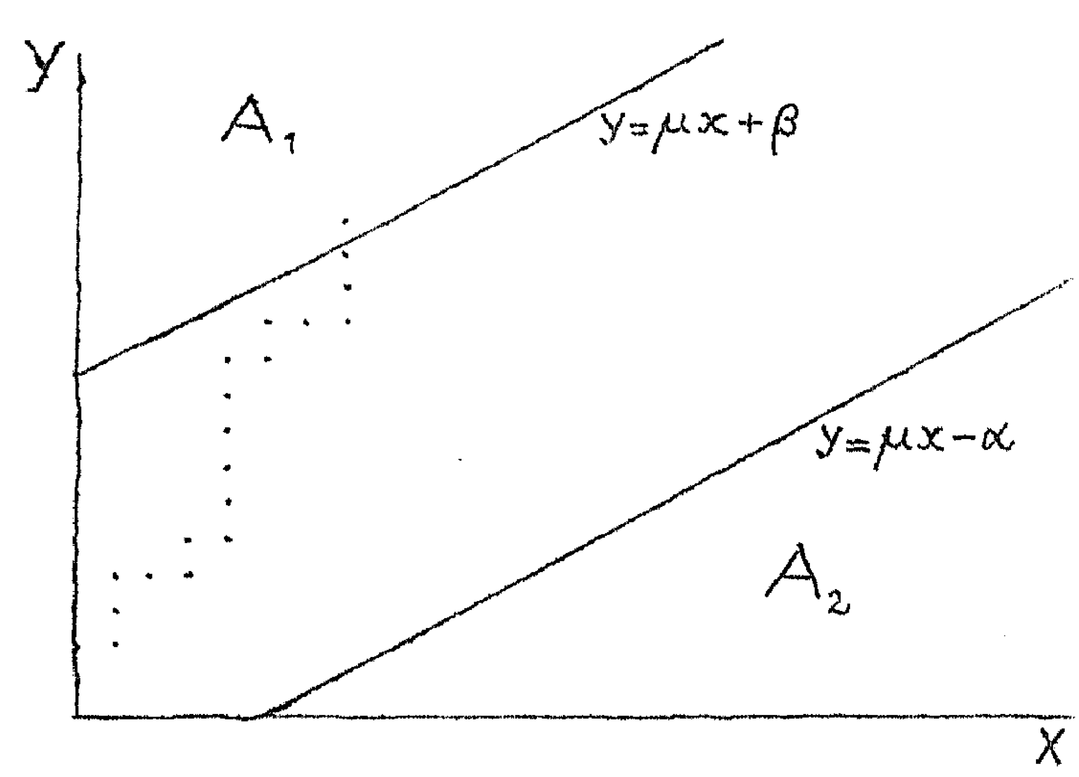



Het aantal slechte objecten uit de steekproef is $x$, het aantal goede objecten $y$; $x+y=n$, de grootte van de steekproef. De waarden van $x$ en $y$, die voor kunnen komen, liggen dus op de rechte $x+y=n$.
Vooraf worden nu twee punten op de lijn bepaald met absces $m_1$ en $m_2$, zodanig, dat als $x \leqq m_1$ wordt gevonden, de partij wordt goedgekeurd en dat als $x \geqq m_2$ is, de partij wordt afgekeurd. Wordt gevonden $m_1 < x < m_2$, dan wordt geen decisie genomen, doch de steekproef uitgebreid door er nieuwe objecten aselect bij te kiezen. Men krijgt dan een steekproef van grootte $n'$ en keurt de partij af, wanneer het totaal aantal slechte objecten een <u>vóór het experiment</u> bepaald aantal $m_0$ overschrijdt. Over zeer goede en zeer slechte partijen kan dus reeds na een kleine steekproef worden beslist.

Men kan de methode van Dodge en Romig verder uitbreiden door een groter aantal steekproeven te beschouwen, dus door de steekproefgrootte geleidelijk te vergroten. Zie onderstaande figuur, waarbij steeds de met een steekproef overeenkomende lijn getekend is. Alleen dan wordt geen beslissing genomen en de steekproefgrootte uitgebreid, wanneer het waargenomen aantal goede en slechte objecten $y$ en $x$ voorgesteld wordt door een punt op het gestippelde deel van de lijn.




Tenslotte kan men aan de steekproef telkens slechts één object toevoegen. Door succesieve onderzoekingen van telkens één object tracht men dan dus tot een beslissing te komen. Dit nu is de methode van de sequente analyse. Deze is op de volgende manier grafisch te illustreren.




$y$ is weer het aantal goede objecten, $x$ het aantal slechte. Op grond van theoretische overwegingen, welke hier niet nader uiteengezet zullen worden, kiest men nu twee rechte lijnen $y=\mu x+\beta$ en $y=\mu x-\alpha$ en keurt af, zodra $y \leqq \mu x-\alpha$, keurt goed, zodra $y \geqq \mu x+\beta$ geworden is ($\mu$ wordt altijd geheel gekozen). Men onderzoekt dus eerst één object, kiest er daarna één bij, daarna weer één enz. Grafisch krijgt men dan een bepaald pad en zodra dit uit de voorafbepaalde strook komt, wordt de decisie genomen. In de figuur wordt dus goedgekeurd, nadat 21 objecten onderzocht zijn. Deze methode geeft ten opzichte van de oude methode een belangrijke kostenbesparing.

Neemt men aan, dat in de hele partij een fractie $p$ der objecten slecht is, dan is deze methode op te vatten als een stochastisch proces met overgangswhn, bepaald door:

$$(x,y) \to (x+1,y) \text{ met wh } p$$

$$(x,y) \to (x,y+1) \text{ met wh } q = 1-p.$$

De vraag is, wat er bij het proces gebeurt met het getal $y-\mu x$ Hiervoor geldt:

$$y-\mu x \to (y-\mu x)-\mu \text{ met wh } p$$

$$y-\mu x \to (y-\mu x)+1 \text{ met wh } q.$$

Interpreteren wij $y-\mu x$ als het kapitaal van een speler, dan heeft deze speler een kans $p$ op verlies en een kans $q$ op winst bij elke partij. De bedragen bij winst en bij verlies zijn, als $\mu \neq 1$ is, echter niet gelijk. Dit houdt dus een verdere generalisatie in van het klassieke probleem, welke wij slechts terloops zullen behandelen. In de volgende paragraaf beginnen wij met aan te nemen, dat voor beide spelers het bij een partij te winnen bedrag gelijk is aan het te verliezen. Wij zullen dus het reeds eerder in deze paragraaf beschreven probleem behandelen. Het probleem der sequente analyse is evenals het hier te behandelen probleem een speciaal geval van de in hoofdstuk 6 te behandelen algemene absorptietheorie.

## 4.2. Ruïneringskansen

Wij beschouwen een bepaalde speler. Laat $p_z$ de wh zijn, dat deze speler tenslotte zijn tegenstander zal ruïneren, in de onderstelling, dat hij reeds een bedrag z van de totale som $s$ bezit. Laat $q_z$ de wh zijn, dat hij geruïneerd wordt onder dezelfde onderstelling. Blijken zal, dat $p_z + q_z = 1$ is. Dit spreekt niet vanzelf, daar er wellicht een positieve kans zou kunnen zijn op een nimmer eindigend spel. Nu geldt dus:

(1) $$q_z = p\,q_{z+1} + q\,q_{z-1} \quad \text{als} \quad 1 \leqq z \leqq s-1$$

(2) $$q_0 = 1, \quad q_s = 0.$$

Dit is een differentie vergelijking met randvoorwaarden. Wij kunnen hem immers schrijven in de vorm $p(q_{z+1}-q_z)=q(q_z-q_{z-1})$, dat is een betrekking tussen de differenties. De theorie van de differentievergelijkingen vertoont grote overeenkomst met die van de differentiaalvergelijkingen. In dit speciale geval schrijven wij om dit in te zien (1) in de vorm:

$$p(q_{z+1}-2q_z+q_{z-1})+(2p-1)(q_z-q_{z-1})=0,$$

dat is een betrekking tussen de tweede en eerste differentie van $q_z$, die sterk doet denken aan een homogene lineaire differentievergelijking van de tweede orde met constante coëfficiënten $a\frac{d^2q}{dz^2}+b\frac{dq}{dz}=0$.

Van een dergelijke differentiaalvergelijking weten wij, dat de algemene vorm geschreven kan worden als een lineaire combinatie van twee lineair onafhankelijke particuliere oplossingen, die als $ab\neq 0$ is, van de vorm $e^{\lambda z}=\omega^z$ $(\omega=e^\lambda)$ zijn, waarbij de twee waarden van $\lambda$ bepaald worden door de vergelijking $a\lambda^2+b\lambda=0$ Evengoed kunnen wij echter de twee waarden van $\omega$ rechtstreeks vinden door $\omega^z$ te substitueren.

Deze zelfde methode kunnen wij nu toepassen bij onze differentievergelijking. Hebben wij twee l.o. oplossingen van $q_z=p\,q_{z+1}+q\,q_{z-1}$, dan voldoet ook elke lineaire combinatie, zoals direct uit (1) blijkt. Zijn $q_0$ en $q_1$ bekend, dan volgt iedere volgende $q$ daaruit, dus er kunnen ook niet meer dan twee l.o. oplossingen zijn. Wij zoeken dus oplossingen van de vorm $q_z=\omega^z$, dus

$$\omega^z=p\omega^{z+1}+q\omega^{z-1},\quad 1=p\omega+q\omega^{-1},\quad p\omega^2-\omega+q=0,$$

waaruit volgt, als $p\neq q$ is:

(3) $\omega_1=1$, $\omega_2=\frac{q}{p}$, zodat

(4) $q_z=A+B\left(\frac{q}{p}\right)^z$

De randvoorwaarden $q_0=1$ en $q_s=0$ geven $A+B=1$ en $A+B\left(\frac{q}{p}\right)^s=0$, dus

$$A=\frac{-q^s}{p^s-q^s},\quad B=\frac{p^s}{p^s-q^s}$$

en

(5) $$q_z=\frac{q_z(p^{s-z}-q^{s-z})}{p^s-q^s}$$ als $p\neq q$ is.

Door verwisseling van $p$ en $q$ en van $z$ en $s-z$ vinden wij $p_z$ :

(6) $$p_z = \frac{p^{s-z}(q^z - p^z)}{q^s - p^s} = \frac{p^{s-z}(p^z - q^z)}{p^s - q^s} \quad \text{als } p \neq q \text{ is}$$

Optelling geeft

(7) $$p_z + q_z = 1 .$$

zodat inderdaad met zekerheid een der spelers geruïneerd zal worden. Als $p = q$ is, vinden wij maar één $\omega$, namelijk $\omega = 1$. Een andere oplossing van (1) is nu echter $q_z = z$, zodat de algemene oplossing luidt $q_z = A + Bz$ en uit de randvoorwaarden volgt $A = 1$ en $B = -\frac{1}{s}$, zodat:

(8) $$q_z = 1 - \frac{z}{s} \quad \text{als } p = q \text{ is,}$$

(9) $$p_z = 1 - \frac{s-z}{s} = \frac{z}{s} \quad \text{dus weer } p_z + q_z = 1 .$$

Dit resultaat had ook met behulp van de regel van de l'Hôpital uit (5) gevonden kunnen worden:

$$\lim_{\frac{q}{p}\to 1} q_z = \lim_{\frac{q}{p}\to 1} \frac{q^z(p^{s-z} - q^{s-z})}{p^s - q^s} = \lim_{\frac{q}{p}\to 1} \frac{\left(\frac{q}{p}\right)^s - \left(\frac{q}{p}\right)^z}{\left(\frac{q}{p}\right)^s - 1} = \lim_{\frac{q}{p}\to 1} \frac{s\left(\frac{q}{p}\right)^{s-1} - z\left(\frac{q}{p}\right)^{z-1}}{s\left(\frac{q}{p}\right)^{s-1}} = \frac{s-z}{s} .$$

De sequente analyse bestudeert de differentievergelijking:

(10) $$\begin{cases} q_z = p\,q_{z-\mu} + q\,q_{z+1} & (\mu \text{ geheel, } 1 \leqq z \leqq s-1) \\ q_0 = q_{-1} = \dots\dots = q_{-\mu+1} = 0 ,\ q_s = 1 . \end{cases}$$

Deze differentievergelijking is op dezelfde manier op te lossen als (4.2.1), namelijk door $q_z = \sum_1^{\mu+1} c_i\,\omega_i^z$ te substitueren. De $\omega$'s zijn dan de wortels van de vergelijking $\omega^\mu = q\,\omega^{\mu+1} + p$ en onderstellen wij alle verschillend. Is dit laatste niet het geval, doch komt de wortel $\omega_\lambda$ $k$ maal voor, dan zullen $z\omega_\lambda^z, z^2\omega_\lambda^z, \dots, z^{k-1}\omega_\lambda^z$ ook aan de differentievergelijking voldoen. De algemene oplossing is dus: $q_z = \sum_1^{\ell} \sum_0^{k_i - 1} c_{ij}\, z^j \omega_i^z$ als er $\ell$ verschillende wortels $\omega$ zijn en de $i$-de $\omega$ $k_i$ maal voorkomt $\left(\sum_1^{\ell} k_i = \mu + 1\right)$. De c's zijn uit de randvoorwaarden op te lossen. Burman (zie [9]) verkreeg de oplossing in elegante vorm.

Het ligt voor de hand ook de differentievergelijking:

(11) $$\begin{cases} q_z = p\,q_{z-a} + q\,q_{z+b} & (0 < z < s) \\ q_z = 0 \text{ als } -a < z \leqq 0,\ q_z = 1 \text{ als } s \leqq z < s+b \text{ is} \end{cases}$$

te bestuderen.

Zolang $a$ en $b$ rationaal zijn, is er geen principiële moeilijkheid en is de boven beschreven methode toe te passen. Door een verandering van de maateenheid van z kan altijd bereikt worden, dat de bedragen bij winst en verlies geheel worden. De karakteristieke vergelijking wordt dan: $\omega^a = q\omega^{a+b} + p$, waarin $a$ en $b$ geheel zijn. Dit is dus een vergelijking van de graad $a+b$ in $\omega$ De veranderingen, die z kan ondergaan, bedragen steeds een geheel aantal malen de grootste gemene deler van $a$ en $b$. De mogelijke waarden van z zijn dus beperkt tot een eindig aantal (waarbij wij uitgaan van een bepaalde begintoestand).
Wanneer of $a$, of $b$ echter irrationaal is, geldt dit niet meer. Nu kan z willekeurig dicht bij elke waarde in het interval o-s komen, daar uit de theorie van de diophantische approximaties volgt, dat de getallen $ma+nb$ ( $m$ en $n$ geheel) overal dicht liggen in de vz van de reële getallen. Het aantal mogelijke waarden van z is aftelbaar oneindig, de differentievergelijking is van oneindige orde. Het is nog niet gelukt een analytische uitdrukking voor $q_z$ te vinden, behalve in het geval $s < a+b$.

4.3. Verwachting van de duur van het spel.

Wij beschouwen weer het klassieke spel met gelijke bedragen bij winst en verlies (gelijk aan de eenheid) en zullen nu de verwachte duur van het spel gaan berekenen.
Wij nemen aan, dat deze verwachting eindig is en schrijven er $D_z$ voor als de speler een bedrag z bezit.
Er geldt:

(1) $$D_z = p\,D_{z+1} + q\,D_{z-1} + 1 \quad \text{als} \quad 0 < z < s$$

(2) $$D_0 = 0\,, \quad D_s = 0\,.$$

Wij moeten er namelijk aan denken, dat nadat een partij gespeeld is, de verwachting van de spelduur ook één kleiner geworden is, vandaar de 1.
Wij hebben hier een lineaire, niet homogene differentievergelijking met constante coëfficiënten. Het verschil van twee oplossingen voldoet aan de homogene vergelijking, zodat wij de algemene oplossing van (1) vinden door bij de algemene oplossing van de homogene vergelijking een particuliere van de inhomogene te tellen.

Nu is de algemeneoplossing van de homogene vergelijking $A+B\left(\frac{q}{p}\right)^z$ als $q \neq p$ is. Een particuliere oplossing van (1) zoeken wij door $D_z = cz$ te substitueren:

$cz = pc(z+1) + qc(z-1) + 1$, dus $0 = pc - qc + 1$, dus $c = \frac{1}{q-p}$,

waaruit volgt:

(2) $$D_z = A + B\left(\frac{q}{p}\right)^z + \frac{z}{q-p}.$$

De randvoorwaarden $D_0 = 0$ en $D_s = 0$ geven $A+B=0$ en $A+B\left(\frac{q}{p}\right)^s = -\frac{s}{q-p}$, zodat:

(3) $$D_z = \frac{z}{q-p} - \frac{s}{q-p}\cdot\frac{1-\left(\frac{q}{p}\right)^z}{1-\left(\frac{q}{p}\right)^s} \qquad \text{als } p \neq q \text{ is.}$$

Als $p = q$ is, is $D_z = -z^2$ een particuliere oplossing van (1), zodat $D_z = -z^2 + A + Bz$, $A = 0$, $-s^2 + Bs = 0$, dus $B = s$, zodat:

(4) $$D_z = z(s-z) \qquad \text{als } p = q \text{ is.}$$

(ook door limietovergang uit (3) te vinden)

Uit (4) volgt, dat als $p = q$ is, de verwachte spelduur het product is van de beginkapitalen der beide spelers, dus sterk toeneemt als $\frac{z}{s-z}$ tot 1 nadert.

Hebben beide spelers ƒ 500,-- en gooien zij met een zuivere munt steeds om ƒ 1,--, dan is de verwachte duur van het spel 250.000 partijen.

Zelfs wanneer de ene speler ƒ 1,-- en de andere ƒ 999,-- te verliezen heeft, is de verwachte spelduur nog 999. De winstkansen zijn dan volgens (4.2.8) 0,001 en 0,999.

4.4. <u>Verdeling van de duur van het spel.</u>

Wij willen nu naast de verwachting ook de verdeling te weten komen van de duur $\underline{n}$ van het spel.

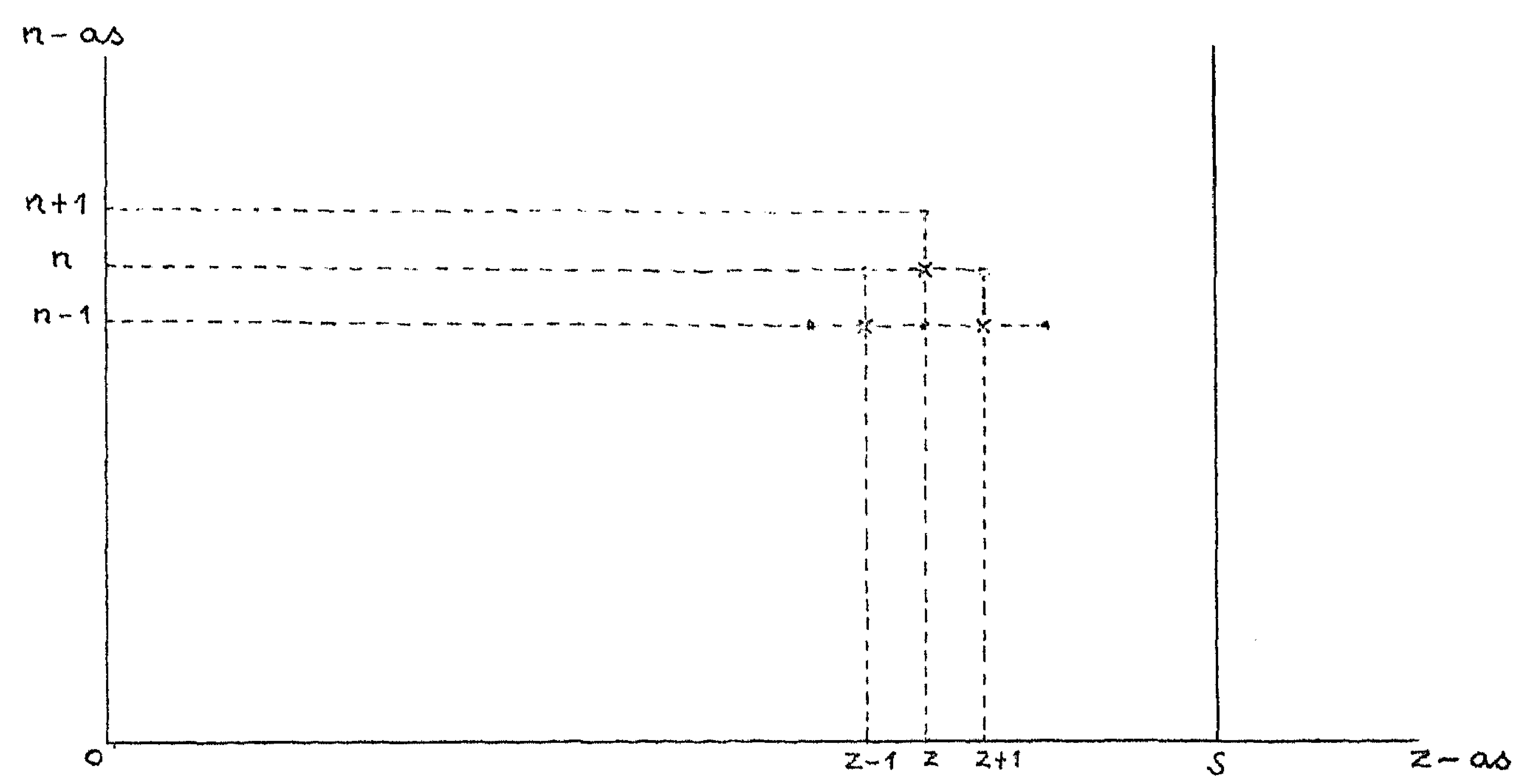
Wij definiëren hiertoe $u_{z,n}$ als de wh, dat precies na het spel de bewuste speler met beginkapitaal z geruïneerd zal zijn.

Nu is dus:

(1) $$u_{z,n+1} = p\,u_{z+1,n} + q\,u_{z-1,n}.$$

Hier variëren dus twee indices, zodat overeenkomst verwacht mag worden met de partiële differentiaalvergelijkingen.

Om de oplosbaarheid te onderzoeken maken wij een grafische voorstelling:




$u_{z,n+1}$ is een soort gewogen gemiddelde tussen $u_{z+1,n}$ en $u_{z-1,n}$
Uit de figuur blijkt, dat iedere rij behoudens de uiteinden uit de er direct onder liggende rij te berekenen is. Wil er een eenduidige oplossing zijn, dan moeten dus de eerste rij $(n=0)$ en de eerste en laatste kolom ( $z=0$ en $z=S$ ) gegeven zijn.
Dit zijn precies de randvoorwaarden

(2) $\quad u_{0,0}=1\ ,\ u_{z,0}=0\ \ (z>0),\ u_{0,n}=0\ (n>0)\ ,\ u_{S,n}=0\ \ (n\geqq 0).$

Wij weten nu dus, dat (1) en (2) eenduidig oplosbaar zijn en zouden de oplossing eventueel numeriek kunnen berekenen. Wij willen de oplossing echter graag analytisch hebben en zullen hiertoe gebruik maken van de methode der voortbrengende functies, zoals ook Laplace heeft gedaan.
Bij constante z en variabele n definiëren wij

(3) $\quad U_z(A) = \sum_0^\infty {}^n\, u_{z,n} A^n\ ,$

waarbij A een hulpvariabele is.
Nu is $0 \leqq \sum_0^\infty {}^n u_{z,n} \leqq 1$, want $\sum_0^\infty u_{z,n} = U_z(1) = q_z$ (zie 4.2.)
de wh, dat de speler met beginkapitaal z ooit geruïneerd zal worden, dus $\sum_0^\infty {}^n u_{z,n}$ is een wh. $U_z(A)$ zal dus zeker convergent zijn als $|A| \leqq 1$ is.
Uit (2) volgt:

(4) $\quad U_0(A) = u_{0,0} = 1\ \ ;\ U_S(A) = 0\ .$

$U_z(A)$ voldoet aan een gewone differentievergelijking als wij A constant houden. Volgens (2) is $u_{z,0}=0$ als $z>0$ is, dus dan is $U_z(A) = \sum_1^\infty {}^n u_{z,n} A^n = \sum_0^\infty {}^n u_{z,n+1} A^{n+1} = \sum_0^\infty {}^n (p\, u_{z+1,n} A^{n+1} + q\, u_{z-1,n} A^{n+1}) =$
$p\,A\,U_{z+1}(A) + q\,A\,U_{z-1}(A),$

dus:

(5) $\quad U_z(A) = A\left(p\,U_{z+1}(A) + q\,U_{z-1}(A)\right)$ ;

Als $A=1$ is, is dit dezelfde differentievergelijking als (4.2.1) en (4.2.2), zodat $U_z(1)=q_z$ is, zoals wij reeds opmerkten. Wezenlijk is echter, dat (5) oplosbaar is voor alle $A$ met $|A|\leqq 1$ Ter oplossing stellen wij weer $U_z = \omega^z$. Wij krijgen dan:

(6) $\quad 1 = A\left(p\omega + q\omega^{-1}\right)$,

een vierkantsvergelijking, welke in het algemeen twee oplossingen heeft, namelijk

(7) $$\omega_{1,2} = \frac{1 \pm \sqrt{1-4pqA^2}}{2pA}$$

(deze wortels zijn reëel, daar $1-4pqA^2 \geqq 0$ is; immers $pq \leqq \frac{1}{4}$ en $0 \leqq A \leqq 1$ ).
De algemeneoplossing is nu $U_z = a\omega_1^z + b\omega_2^z$, terwijl uit de randvoorwaarden $U_0 = 1$ en $U_s = 0$ volgt $a+b=1$ en $a\omega_1^s + b\omega_2^s = 0$, dus:

(8) $$U_z(A) = \frac{\omega_1^{s-z} - \omega_2^{s-z}}{\omega_1^s - \omega_2^s}\cdot\left(\frac{q}{p}\right)^z$$

waarbij $\omega_1$ en $\omega_2$ door (7) bepaald zijn (bedenk namelijk, dat $\omega_1\omega_2 = \frac{q}{p}$ is). Wij hebben nu dus $U_z(A)$ als functie van $A$ gekregen (want $\omega_1$ en $\omega_2$ zijn functies van $A$ ). Om de $u_{z,n}$ te vinden moeten wij $U_z(A)$ als functie van A in een machtreeks ontwikkelen. Hiertoe volgen wij de methode van Laplace en stellen

(9) $$\frac{1}{2A\sqrt{pq}} = \cos\varphi, \text{ dus}$$

$$\omega = \left(1 \pm \sqrt{1-\sec^2\varphi}\right).\cos\varphi\sqrt{\frac{q}{p}} = \sqrt{\frac{q}{p}}\left(\cos\varphi \pm \sqrt{\cos^2\varphi - 1}\right) = \sqrt{\frac{q}{p}}.e^{\pm i\varphi}$$

(10) $$U_z(A) = \left(\sqrt{\frac{q}{p}}\right)^z.\frac{e^{i(s-z)\varphi} - e^{-i(s-z)\varphi}}{e^{is\varphi} - e^{-is\varphi}} = \left(\frac{q}{p}\right)^{\frac{1}{2}z}\frac{\sin(s-z)\varphi}{\sin s\varphi}.$$

Nu is $U_z(A)$ een rationale functie van $A$ ,hetgeen wij inzien door in (8) teller en noemer door $\omega_1 - \omega_2$ te delen. Dan wordt $U_z(A)$ $U_z(A) = \frac{F_1(A)}{F_2(A)}$ , waarbij bijvoorbeeld $F_2(A) = \frac{\omega_1^s - \omega_2^s}{\omega_1 - \omega_2}$ een homogene symmetrische veelterm van de graad $s-1$ in $\omega_1$ en $\omega_2$ is, dus te schrijven is als een veelterm in $\omega_1+\omega_2$ en $\omega_1\omega_2$ , van de graad $s-1$ in $\omega_1 + \omega_2 = \frac{1}{pA}$ . Verder is $\omega_1\omega_2 = \frac{q}{p}$ ; dit bevat

$A$ niet, dus $F_2(A)$ is een veelterm van de graad $s-1$ in $A^{-1}$ dus een veelterm in $A$, gedeeld door een macht van $A$. Deze veelterm is van de graad $s-1$ als s oneven is en van de graad $s-2$ als s even is. Evenzo is $F_1(A)$ een veelterm in $A$, gedeeld door een macht van $A$, dus $U_z(A)$ is een rationale functie van $A$. Nu zijn de wortels van $F_2(A)$ eenvoudig te vinden met behulp van (10).
De noemer van (10) wordt 0 als $\sin s\varphi = 0$ dus $\varphi_k = \frac{k\pi}{s}$ is. Hiervan kunnen wij die waarden van $\varphi$ niet gebruiken, die ook de teller 0 maken, dus de $\varphi_k$ met $k=0$ en $k=s$ moeten wij weglaten

$$A_k = \frac{1}{2\sqrt{pq}\,\cos\varphi_k} \quad \text{volgens (9)}$$

Als $s$ even is, zal $\cos\varphi_k = 0$ zijn als $k=\frac{s}{2}$ is. Deze waarde van $k$ is dus ook niet bruikbaar.
Wij krijgen dus $s-1$ wortels $A_k$ van $F_2(A)$ als $s$ oneven is en $s-2$ wortels als $s$ even is. Steeds dus een even aantal.
Nu passen wij splitsing in partiële breuken toe:

$$(11) \qquad U_z(A) = \sum_1^{s-1}{}_k \frac{\rho_k}{A_k - A} \qquad \text{(als } s \text{ even is, moet } k=\tfrac{s}{2} \text{ uitgesloten worden)}$$

Hierbij is $\rho_k$ gelijk aan de teller van (10), gedeeld door de afgeleide naar $A$ van de noemer van (10) voor de waarde $A=A_k$ of $\varphi=\varphi_k$:

$$\rho_k = -\left(\frac{q}{p}\right)^{\frac{1}{2}z} \frac{\sin\frac{(s-z)k\pi}{s}}{s\cos\pi k\left(\frac{d\varphi}{dA}\right)_{A=A_k}} = \left(\frac{q}{p}\right)^{\frac{1}{2}z} \frac{\sin\frac{z\pi k}{s}\,\sin\frac{k\pi}{s}}{2s\sqrt{pq}\,\cos^2\frac{k\pi}{s}}$$

daar

$$\left(\frac{d\varphi}{dA}\right)_{A=A_k} = \frac{1}{\left(\frac{dA}{d\varphi}\right)_{\varphi=\varphi_k}} \quad \text{en} \quad \left(\frac{dA}{d\varphi}\right)_{\varphi=\varphi_k} = -\frac{s\sin\varphi_k}{2\sqrt{pq}\,\cos^2\varphi_k},$$

is

terwijl $-\dfrac{\sin\frac{(s-z)\pi k}{s}}{\cos\pi k} = \sin\dfrac{z\pi k}{s}$ is.

Uit (11) volgt:

$$U_z(A) = \sum_1^{s-1}{}_k \frac{\rho_k}{A_k} \sum_0^{\infty}{}_n \left(\frac{A}{A_k}\right)^n, \qquad \text{dus}$$

$$(12) \qquad u_{z,n} = \sum_1^{s-1}{}_k \frac{\rho_k}{A_k^{n+1}} = \frac{2^n p^{\frac{1}{2}(n-z)} q^{\frac{1}{2}(n+z)}}{s} \sum_1^{s-1}{}_k \cos^{n-1}\frac{k\pi}{s}\,\sin\frac{k\pi}{s}\,\sin\frac{k\pi z}{s}.$$

De term met $k=\frac{s}{2}$ , die wij als $s$ even is weg moeten laten, is zoals uit (12) blijkt, toch 0, zodat (12) zonder beperking ook geldig is als $s$ even is.

4.5. Waarschijnlijkheidstheoretische interpretatie van de voortbrengende functie.

De voortbrengende functie $U_z(A)=\sum_0^\infty n\, u_{z,n}A^n$ kunnen wij eenvoudig waarschijnlijkheidstheoretisch interpreteren als $0\leqq A\leqq 1$ is. Hiertoe stellen wij ons een ander spel voor, dat alleen daarin van het voorgaande verschilt, dat vóór iedere partij geloot zal worden of het spel zal worden voortgezet. Steeds is dan de kans, dat het spel voortgezet zal worden gelijk $A$ en de kans, dat dit niet zal gebeuren $1-A$ . Bij dit gewijzigde spel is $U_z(A)$ nu de totale kans, dat de speler met beginkapitaal geruïneerd zal worden, dus is een wh.
Immers de kans, dat de speler na precies $n$ partijen geruïneerd zal zijn, is gelijk aan het product van de kans, dat deze $n$ partijen gespeeld zullen worden (deze is $A^n$) en de kans, dat de speler in $n$ partijen geruïneerd zal worden als wij weten, dat deze gespeeld zullen worden (deze is $u_{z,n}$ ), dus deze kans is $u_{z,n}A^n$.. Voor verschillende $n$ hebben wij weer elkaar uitsluitende eventualiteiten, dus de totale wh, dat de speler geruïneerd zal worden, is $\sum_1^\infty n\, u_{z,n}A^n$ , dat is $U_z(A)$ als $0<z<s$ en $u_{z,0}=0$ .
De randen $z=0$ en $z=s$ moeten wij afzonderlijk bekijken. In overeenstemming met bovenstaande wh-theoretische interpretatie eisen wij $U_0(A)=1$ en $U_s(A)=0$ , waaruit de reeds bekende randvoorwaarden voor de $u_{z,n}$ 's volgen .
Met behulp van het gewijzigde spel kunnen wij nu de differentievergelijking voor de $U_z(A)$ rechtstreeks opstellen. Als namelijk $0<z<s$ is, dan is $U_z$ gelijk aan het product van twee whn:

1. de wh, dat er nog een partij gespeeld zal worden ($A$).
2. de wh, dat de speler uiteindelijk geruïneerd zal worden, als wij weten, dat er minstens één partij nog gespeeld zal worden. Deze wh vinden wij als volgt: Wint de speler de partij (kans daarop $p$ ), dan bezit hij $z+1$ en is zijn ruïneringskans $U_{z+1}$ geworden; verliest hij (kans $q$ ), dan bezit hij $z-1$ en is zijn ruïneringskans $U_{z-1}$ geworden. Deze eventualiteiten sluiten elkaar uit, dus krijgen wij voor de bedoelde wh $pU_{z+1}+qU_{z-1}$.

Wij vinden dus:

$$U_z = A\left(p U_{z+1} + q U_{z-1}\right),$$ dat is precies (4.4.5).

Het is dikwijls mogelijk voortbrengende functies op deze manier als whn te interpreteren. Wij zullen nog een voorbeeld geven in de volgende paragraaf en verwijzen verder naar [5] , [16] , alsmede naar paragraaf 5.7.

4.6. Aanvullende beschouwingen.

Wij beschouwen de ééndimensionale stochastische wandeling zonder grenzen: Laat $g_n$ de wh zijn, dat een bewegend punt, dat in O begint, na $n$ stappen voor het eerst in -1 aankomt. Dan is $g_n$ ook de wh, dat een punt, dat in $z$ begint na $n$ stappen voor het eerst in $z-1$ aankomt (er is immers geen rechtergrens).
Als

$$(1) \qquad G(A) \overset{def}{=} \sum_0^{\infty} g_n A^n$$

dan is $G(A)$ weer te interpreteren als de totalekans, dat een wandelend punt, uitgaande van O, ooit in -1 aankomt bij het gewijzigde spel met kansen $A$ en $1-A$ op voortzetting, respectievelijk beëindiging van het spel na elke partij.
Het punt kan alleen sprongen van de lengte 1 maken, dus kan alleen via $z-1$ van $z$ naar $z-2$ komen. Derhalve is de kans, dat het punt ooit van $z$ naar $z-2$ komt, gelijk aan de kans, dat het van $z$ naar $z-1$ komt, vermenigvuldigd met de kans, dat het van $z-1$ naar $z-2$ komt, dus is gelijk aan $G(A) \times G(A) = G(A)^2$.
In het algemeen is (bij het gewijzigde spel) de kans, dat het wandelende punt van $z$ uitgaande ooit in $z-k$ aankomt gelijk aan $G(A)^k$. Voor $G(A)$ geldt de betrekking:

$$(2) \qquad G = A\left(q + pG^2\right)$$

Namelijk als wij weten, dat er nog een spel gespeeld zal worden (kans daarop $A$ ), dan is de wh, dat $z-1$ direct bereikt wordt $q$ en de wh, dat dit via $z+1$ zal gebeuren $pG^2$ , waaruit (2) volgt.
Voor $G$ vinden wij dus een vierkantsvergelijking, in vorm gelijk aan (4.4.6). De wortels zijn reëel:

$$\omega_1 = \frac{1+\sqrt{1-4pqA^2}}{2pA}, \qquad \omega_2 = \frac{1-\sqrt{1-4pqA^2}}{2pA}, \quad \omega_1 \geq 0, \ \omega_2 \geq 0.$$

Tevens moet $0 \leq G \leq 1$ zijn.
Nemen wij $A>0$ ( $A=0$ is niet interessant en triviaal), dan is

$$pAG^2 - G + qA \begin{cases} >0 & \text{als } G=0 \text{ is.} \\ = A-1 \leq 0 & \text{als } G=1 \text{ is, dus moet, als } A<1 \end{cases}$$

is, er precies één wortel tussen 0 en 1 zijn. Dit is $\omega_2$ (de kleinste), dus $0 < \omega_2 < 1 < \omega_1$ en

$$(3) \qquad G(A) = \omega_2 = \frac{1-\sqrt{1-4pqA^2}}{2pA}.$$

Ontwikkelen wij $G(A)$ in een machtreeks, dan zal de coëfficiënt van $A^n$ juist gelijk $g_n$ zijn. De $g_n$'s zijn dus eenvoudig te vinden
Als $A=1$ is, dan is één wortel gelijk 1, de andere $\frac{q}{p}$ , dus:

$0 < \omega_2 = 1 \leq \omega_1 = \frac{q}{p}$ als $A=1$ en $q \geq p$ is en

$0 < \omega_2 = \frac{q}{p} \leq 1 = \omega_1$ als $A=1$ en $q \leq p$ is, dus $G(1) = \min\left(\frac{q}{p}, 1\right)$.

De kans $H(A)$ , dat het punt vanuit $z$ ooit in $z+1$ zal komen, verkrijgen wij door $p$ en $q$ te verwisselen, waaruit volgt, dat wij de reciproke wortels van (4.4.6) moeten hebben. Nu is $\frac{1}{\omega_2} > 1$, dus onbruikbaar.
Derhalve is:

$$(4) \qquad H(A) = \frac{1}{\omega_1} = \frac{p}{q}\,\omega_2 .$$

Definiëren wij tenslotte $f_n$ als de wh, dat het bewegende punt na $n$ stappen weer voor het eerst op zijn uitgangsplaats terug is - $f_n$ heet de terugkeerkans in $n$ stappen - en $F(A) \stackrel{def}{=} \sum_0^\infty f_n A^n$ ($f_0 = 0$), dan is $F(A)$ bij het gewijzigde spel te interpreteren als de wh, dat het punt ooit weer in zijn uitgangstoestand terugkeert en geldt:

$$(5) \qquad F(A) = A\big(pG(A) + qH(A)\big), \qquad \text{dus}$$

$$(6) \qquad F(A) = A\left(p\omega_2 + \frac{q}{\omega_1}\right) = 2pA\omega_2 = 1 - \sqrt{1-4pqA^2}.$$

Hieruit volgt, dat $\sqrt{1-4pqA^2} = 1 - F(A)$ de kans is, dat het punt nooit meer terugkeert. Zo krijgen $\omega_1$ en $\omega_2$ , die om analytische redenen worden ingevoerd, nu ook een wh-theoretische betekenis.
Wij hebben in het voorgaande voor de ééndimensionale onbegrensde stochastische wandeling gevonden (neem $A=1$):
Als $p>q$ is, dan zullen punten, rechts van het uitgangspunt $z$ gelegen, met wh 1 bereikt worden; het punt $z-1$ wordt met wh $\frac{q}{p}$

bereikt, het punt $z-k$ met wh $\left(\frac{q}{p}\right)^k$ $(k>0)$ . De totale terugkeerkans is $F(1) = 1-\sqrt{1-4pq} = 2q < 1$ . Als $p<q$ is, zullen links van z gelegen punten met zekerheid bereikt worden; het rechts van z gelegen punt $z+k$ zal met wh $\left(\frac{p}{q}\right)^k$ bereikt worden, terwijl de totale terugkeerkans gelijk $2p<1$ is.
Als $p = q = \frac{1}{2}$ is, zullen alle punten met zekerheid bereikt worden en zal ook de totale terugkeerkans gelijk 1 zijn.

# Deel II In de tijd discrete stochastische processen

## 5. Inleiding tot de theorie der oneindige Markof-ketens

### 5.1 Oneindige matrices

De theorie van de eindige Markof-ketens hebben wij teruggebracht tot een bestudering van eindige Markof-matrices en daardoor gebaseerd op de theorie van de lineaire vergelijkingen. Wanneer de verzameling $\mathcal{J}$ van mogelijke toestanden nu oneindig is, ontstaan grote moeilijkheden, die ten dele nog onopgelost zijn. Het blijkt namelijk, dat zelfs wanneer $\mathcal{J}$ aftelbaar oneindig is, de meest elementaire eigenschappen van stelsels lineaire vergelijkingen verloren gaan.
Duidelijk is, dat van de determinantentheorie niet veel overblijft. Men kan eventueel van een oneindige matrix eindige vierkante deelmatrices nemen, hiervan de determinantwaarde berekenen en limieten gaan onderzoeken, doch deze zullen meestal niet bestaan.
De determinantloze theorie van eindige stelsels lineaire vergelijkingen $Mx = y$, $y$ gegeven, berust op de volgende grondeigenschappen:

1. $\exists_x\, M \quad o,\ x \neq o \Leftrightarrow \exists_q\, qM = o,\ q \neq o$

(Ook het geadjungeerde stelsel heeft dus een niet triviale oplossing; deze eigenschap correspondeert met $|M| = o$ )

2. $\forall_x$ volgt uit $M_x = o$, dat $x = o \Leftrightarrow \forall_y\, \exists_x\, Mx = y$

(correspondeert met $|M| \neq o$ )

3. $\exists_{x_1, x_2}\, Mx_1 = y,\ Mx_2 = y,\ x_1 \neq x_2 \longrightarrow \exists_x\, Mx = o,\ x \neq o.$

Wanneer $\mathcal{J}$ nu aftelbaar oneindig is, dan blijft alleen eigenschap 3 behouden, terwijl 1 en 2 verloren gaan.
Ten aanzien van 3 is dit direct duidelijk. Als namelijk $Mx_1 = y$ en $Mx_2 = y$ en $x_1 \neq x_2$ , dan is $M(x_1 - x_2) = Mx_1 - Mx_2 = o$ , dus $x_1 - x_2 (\neq o)$ is een oplossing van $Mx = o$.
Dat de eigenschappen 1 en 2 in het algemeen verloren gaan, blijkt uit de volgende tegenvoorbeelden:

1. Neem $\mathcal{J} = N$ (de vz van de natuurlijke getallen) en neem $M_j^i = I_j^{i+1} - I_j^i$, dus

$$M = \begin{pmatrix} -1 & +1 & & & o \\ & -1 & +1 & & \\ o & & -1 & +1 & \\ & & & \ddots & \ddots \end{pmatrix}$$

Is $Mf=0$, dus $M^i_j f^j = 0$ , dan is $I^{i+1}_j f^j = I^i_j f^j$ , dus $f^{i+1} = f^i$ en omgekeerd, zodat $\forall^N_i f^i = 1$ een niet triviale oplossing is. Bekijk nu $gM=0$, d.w.z. $g_i M^i_j = 0$, dus $g_i I^{i+1}_j = g_i I^i_j = g_j$ , waarbij

$$\sum^i g_i I^{i+1}_j = \begin{cases} g_{j-1} & \text{voor } j \geqq 1 \text{ , waaruit volgt } g_0 = g_1 = g_2 = g_3 = \dots \\ 0 & \text{voor } j = 0 \text{ , dus } g_0 = g_1 = g_2 = \dots\dots = 0, \end{cases}$$

zodat 1 niet geldt.

2. Neem $\mathfrak{J} = N$, $M^i_j = I^{i+1}_j - I^{i+2}_j$.

Nu heeft $Mf=h$ voor willekeurige $h$ een oplossing. Immers $(Mf)^i = f^{i+1} - f^{i+2} = h^i$ , dus neem $f^{i+2} = f^{i+1} - h^i$ en verder $f^0 = 0$, $f^1 = b$ willekeurig $\neq 0$ . Ook $Mf=0$ heeft echter een niet nul-oplossing, namelijk $f = \begin{pmatrix} a \\ b \\ b \\ \vdots \end{pmatrix}$, $a$ en $b$ willekeurig $\neq 0$ . Dus 2 geldt niet van rechts naar links.

Dat 2 ook niet geldt van links naar rechts tonen wij aan door als matrix de gespiegelde van de vorige te nemen. Wij kunnen dus ook dezelfde $M$ nemen en deze dan links vermenigvuldigen met een vector $g$.

$$(gM)_j = g_i I^{i+1}_j - g_i I^{i+2}_j = \begin{cases} g_{j-1} - g_{j-2} & \text{voor } j \geqq 2 \\ g_0 & \text{voor } j = 1 \\ 0 & \text{voor } j = 0 \end{cases}$$

dus uit $gM=0$ volgt $g_0 = 0$ en $\forall^N_j\, g_{j+1} = g_j$ , dus $\forall_j\, g_j = 0$ . Het homogene stelsel heeft alleen maar de nuloplossing. Het niet homogene stelsel $(gM)_j = k_j$ heeft echter geen oplossing als $k_0 \neq 0$ is, daar altijd $(gM)_0 = 0$ moet zijn, dus inderdaad geldt in dit voorbeeld eigenschap 2 niet van links naar rechts.

## 5.2 Associativiteit van de matrixvermenigvuldiging

Als $M^i_j$ en $N^i_j$ matrices zijn met een aftelbaar oneindig aantal rijen en kolommen, dan is ook $M^i_j \pm N^i_j$ gedefinieerd, alsmede $(cM)^i_j$ , (c reëel of complex), doch het product $(MN)^i_j$ zal niet bestaan, wanneer $\sum^k M^i_k N^k_j$ divergeert. In het bijzonder behoeft dus $Mf$ ook niet gedefinieerd te zijn.

De matrixoptelling is commutatief en associatief, terwijl ook de distributieve eigenschap $M(L+N) = ML + MN$ behouden blijft, zodra tenminste twee van de drie producten bestaan.

De associatieve eigenschap van de matrixvermenigvuldiging blijft echter niet steeds behouden, zodat in het algemeen $(LM)N \neq L(MN)$ ook al bestaan de diverse producten. Het is dus mogelijk $L, M$

en N zo te kiezen, dat niet geldt:

$$\forall_i \forall_\ell \; \Sigma^k \left( \Sigma^j L^i_j M^j_k \right) N^k_\ell = \Sigma^j L^i_j \left( \Sigma^k M^j_k N^k_\ell \right).$$

Het is voldoende dit aan te tonen voor een bepaalde waarde van $i$ en $\ell$, zodat wij dus in plaats van de matrix $L^i_j$ een covector $g_j$ kunnen nemen en in plaats van de matrix $N^k_\ell$ een vector $f^k$. Wij zullen in het volgende voorbeeld dus laten zien, dat niet altijd geldt: $\Sigma^k \left( \Sigma^j g_j M^j_k \right) f^k = \Sigma^j g_j \left( \Sigma^k M^j_k f^k \right)$.

Neem hiertoe $\mathcal{J} = N$, $M^j_k = I^j_k - I^{j+1}_k$, $g = (1,1,1,\ldots..)$, $f = \begin{pmatrix} 1 \\ 1 \\ 1 \\ \vdots \end{pmatrix}$,

dan is

$$\Sigma^j g_j M^j_k = \begin{cases} g_k - g_{k-1} & \text{voor } k < 0 \\ g_0 \text{ voor } k = 0, \text{ dus} & = (1,0,0,\ldots..) = \delta_{k0} \end{cases}$$

$$\Sigma^k M^j_k f_k = f^j - f^{j+1} = 0.$$

Dus

$$\Sigma^k \left( \Sigma^j g_j M^j_k \right) f^k = \Sigma^k \delta_{k0} f^k = f^0 = 1 \quad \text{en} \quad \Sigma^j g_j \left( \Sigma^k M^j_k f^k \right) = \Sigma^j g_j . 0 = 0.$$

Derhalve geen gelijkheid.

Uit de stelling van paragraaf 9.6 der appendix volgt, dat de associatieve eigenschap $(LM)N = L(MN)$ wel behouden blijft, wanneer bestaan:

1. $(|L||M|)$ en $(|M||N|)$, waarbij $|M|^i_j = |M^i_j|$.

2. òf $(|L|.|M|)|N|$, òf $|L|(|M|.|N|)$ (Uit het bestaan van de een en eis 1 volgt het bestaan van de ander).

In ons tegenvoorbeeld is aan eis 2 niet voldaan.

## 5.3 Eigen waarden van oneindige Markof-matrices

Voor eindige Markof-matrices $P$ bewezen wij, dat uit $(P-\lambda)f = 0$ en $f \neq 0$ volgt $|\lambda| \leqq 1$.

Deze eigenschap blijft niet meer gelden als $P$ aftelbaar oneindig is.

Neem namelijk $E = I$ (de vz van alle gehele getallen) en $P^i_j = I^{i+1}_j$

Uit $Pf = \lambda f$ volgt nu $f^{i+1} = \lambda f^i$. Neem dus $\lambda \neq 0$ willekeurig, dan is een niet triviale oplossing $f^i = \lambda^i$ ($\lambda$ tot de macht $i$, $\lambda^{i+1} = \lambda.\lambda^i$).

Elke $\lambda \neq 0$ is dus een eigen waarde, dus geldt niet $|\lambda| \leqq 1$.

Wensen wij alleen begrensde eigen functies $f$, dan voldoen alleen eigen waarden met absolute waarde 1 en geldt wel $|\lambda| \leqq 1$.

Doch ook dan is het aantal eigen waarden niet-aftelbaar oneindig en heeft de vz van eigen waarden een grotere machtigheid dan de vz van mogelijke toestanden. (Bij eindige matrices is het aantal eigen waarden gelijk aan het aantal elementen in een rij of kolom, mits wij bij het tellen rekening houden met de multipliciteit).
Wij geven nu nog een tweede tegenvoorbeeld, dat interessant is, omdat het begrip eigen waarde van een matrix er door in verband wordt gebracht met het begrip karakteristieke functie uit de whr.
Neem $\mathcal{J} = I$ en hierop een willekeurig Markof-proces met de eigenschap $\forall_k \, P^{i+k}_{j+k} = P^i_j$

Laat $p_n$ de kans zijn op een sprong $n$ naar rechts, dus

$$p_n = \mathcal{P}\{\underline{i}_{k+1} = i+n \mid \underline{i}_k = i\}, \ \sum_{-\infty}^{\infty} p_n = 1, \ P^i_j = p_{j-i} = \sum_{-\infty}^{\infty}{}^n p_n \, I^{i+n}_j$$

(nemen wij $p_1 = 1$ en $p_k = 0$ als $k \neq 1$ is, dan hebben wij het vorige voorbeeld).
Als $Pf = \lambda f$ is, dan is $\sum^n p_n f^{i+n} = \lambda f^i$.

Wij willen begrensde eigen fcts. $f$ hebben, dus het ligt voor de hand voor $f$ een exponentiële functie met modulus 1 te substitueren.

Stel dus $f^h = e^{iht}$ (de rechtopgeschreven i stelt de imaginaire eenheid voor), dus $\sum^n p_n \, e^{i(n+i)t} = \lambda e^{iit}$, waaruit volgt, dat $f^h = e^{iht}$ een eigen vector is bij de eigen waarde $\lambda = \sum^n p_n e^{int} = \varphi(t)$ als $\varphi(t)$ de karakteristieke functie voorstelt, behorende bij de whn $p_n$.
Dus $\varphi(t)$ is eigen waarde met bijbehorende eigen vector $e^{iht}$
$\lambda = \varphi(t)$ is een gehele trancendente functie van $t$, neemt dus, zoals in de functietheorie bewezen wordt, elke waarde tussen -1 en +1 hoogstens aftelbaar vaak aan, dus er zijn niet aftelbaar veel eigen waarden.

## 5.4 Fuiken bij oneindige Markof-matrices

In 2.4 bewezen wij, dat een eindige Markof-matrix $P$ reducibel is als bij de eigen waarde 1 een eigen vector te vinden is, niet evenredig met de eenheidsvector. Dus:

$$\exists_f \ Pf = f, \ 0 < \|f\| < \infty, \ \exists_i \ f^i - f^1 \neq 0 \longrightarrow P \text{ reducibel.}$$

Deze eigenschap blijft niet meer gelden.
Neem namelijk $\mathcal{J} = I$ en voor $P$ de matrix, behorende bij de onbe-

grensde stochastische wandeling:

$$P^i_j = p_i I^{i+1}_j + q_i I^{i-1}_j \ , \ p_i + q_i = 1 \ , \quad q_i = \begin{cases} 2^{-i-1} \text{ voor } i \geqq 0 \\ 1 - 2^{-|i|-1} \text{ voor } i \leqq 0 \end{cases}$$

(Er is geen sommatie over $i$ bedoeld!)

$\prod_1^\infty p_i = \prod_1^\infty (1-q_i)$ is convergent, daar $\sum_1^\infty q_i$ convergent is, dus $\prod_1^\infty p_i > 0$ (alle $p_i > 0$ )

Evenzo is $\prod_1^\infty q_{-i} > 0$.

Beschouw $Pf = f$ , $p_i f^{i+1} + q_i f^{i-1} = f^i = (p_i + q_i) f^i$ (geen sommatie!),

dus $\quad p_i(f^{i+1} - f^i) = q_i(f^i - f^{i-1})$.

Neem $i > 0$ : $|f^{i+1} - f^i| = \frac{q_i}{p_i}|f^i - f^{i-1}| < 2^{-i}|f^i - f^{i-1}| <$

$$< 2^{-\{i+(i-1)+\dots+1\}}|f^1 - f^0| = 2^{-\frac{1}{2}i(i+1)}|f^1 - f^0| ,$$

dus $\quad |f^n| \leqq \sum_0^\infty |f^{i+1} - f^i| + |f^0| < |f - f^0| \sum_0^\infty 2^{-\frac{1}{2}i(i+1)} + |f^0| < \infty$ ,

dus begrensd, als $n > 0$.

Hetzelfde geldt als $n < 0$ is.
Neem verder $f^0 = 0, f^1 = \frac{1}{2}$ en $f^{-i} = f^i$ , dan is $f$ niet evenredig met de eenheidsvector, maar toch bevat $\mathfrak{I}$ geen afgesloten deelvz.
Stel namelijk $S \subset P, S \neq 0$ , $S$ stochastisch afgesloten en neem $i \in S, j \notin S$,

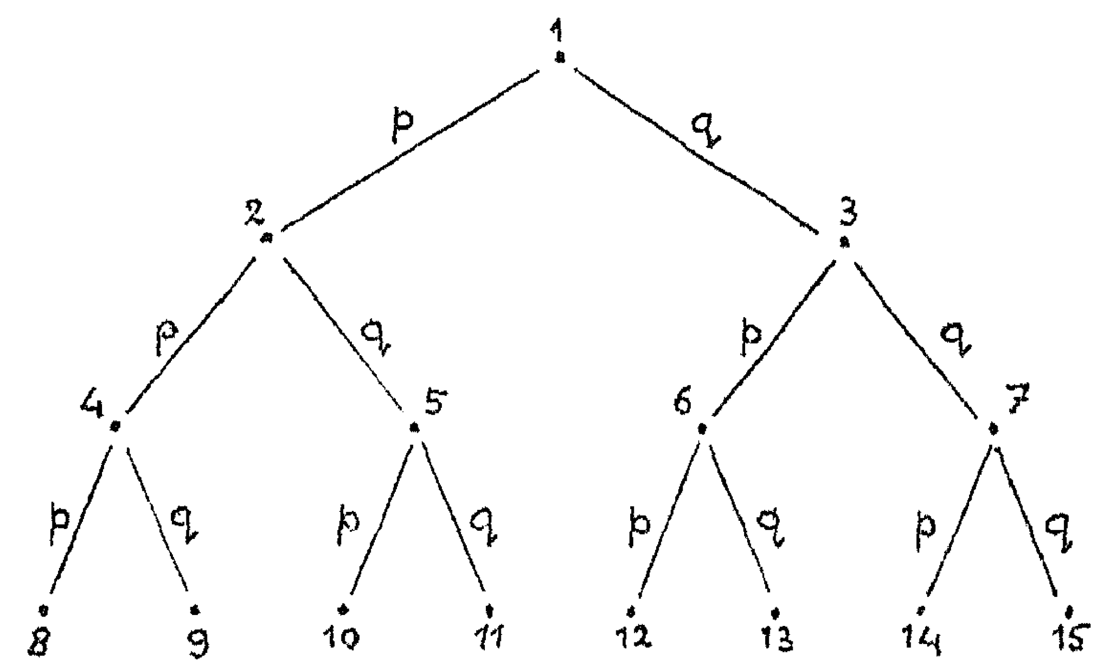
1. $\quad j > i \ , \ \mathcal{P}\{\underline{i}_{j-1} = j \mid \underline{i}_0 = i\} = p_i p_{i+1} \cdots \cdots p_{j-1} > 0$

2. $\quad j < i \ , \ \mathcal{P}\{\underline{i}_{i-j} = j \mid \underline{i}_0 = i\} = q_i q_{i-1} \cdots \cdots q_{j+1} > 0$ ,

waardoor wij een contradictie krijgen.
Hoewel bij de eigen waarde 1 dus een begrensde eigen vector hoort, niet evenredig met de eenheidsvector, zijn er toch geen fuiken.

W. Doeblin bestudeerde de asymptotische eigenschappen van Markof-ketens, die aan enkele speciale voorwaarden voldeden. De eerste voorwaarde luidt, dat niet iedere fuik weer twee disjuncte deelfuiken heeft. Zie [22].
Wij kunnen echter een voorbeeld geven, waarbij niet aan deze voorwaarde is voldaan.
Neem namelijk $P^i_j = p I^{2i}_j + q I^{2i+1}_j$ , $p + q = 1$ , $pq > 0$ ,

dus de mogelijke overgangen kunnen wij als volgt schematisch weergeven:




Zijn wij bijvoorbeeld in toestand 5 aangekomen, dan komen wij nooit meer uit de "deelboom" met 5 als top. Dit is dus een fuik, die weer minstens twee disjuncte deelfuiken heeft (bijv. de "deelbomen" met 10, resp. 11 als top). Elke fuik heeft dus weer minstens twee disjuncte deelfuiken.

## 5.5 Oneindige Markof-processen

Ondanks al deze weinig bemoedigende tegenvoorbeelden zijn er nog zeer interessante stellingen te bewijzen bij oneindige Markof-processen. Wij kunnen de theorie der stochastische processen zelfs geheel abstract opzetten, dus voor de vz van mogelijke toestanden $\mathfrak{J}$ een abstracte vz nemen.

In de appendix, hoofdstuk 9, wordt een generalisatie gegeven van het matrixbegrip. Een fct van tweeërlei soort $M^x_X$, meetbaar als fct van $x$ bij vaste $X$ en aftelbaar additief in $X$ bij vaste $x$, is te beschouwen als generalisatie van de gewone matrix $M^i_j$.
Een matrixproduct wordt gedefiniëerd door middel van:

$$(9.3.12) \qquad (MN)^x_Z = \int_Y M^x_{dy} N^y_Z$$

Onderzocht wordt onder welke voorwaarden deze integraal bestaat en onder welke voorwaarden de associativiteit van de matrixvermenigvuldiging behouden blijft (zie 9.5 en 9.6). Dit blijkt steeds het geval te zijn, wanneer de matrices begrensd zijn, dus eindige norm hebben.
De norm is hierbij gedefiniëerd door:

$$(9.3.4) \qquad \|M\| \stackrel{def}{=} \|M\|^{\mathfrak{J}}_{\mathfrak{J}}, \text{ waarbij}$$

$$\|M\|^X_Y \stackrel{def}{=} \sup_{x\in X} |M^x|_Y \stackrel{def}{=} \sup_{x\in X} \sup_{\{Y_i\}\in\mathfrak{D}(Y)} \sum^i |M^x_{Y_i}|,$$

waarin $\mathfrak{D}(Y)$ de klasse van alle "dissecties" van $Y$ in aftelbaar vele disjuncte deelvzn voorstelt.

In het volgende zullen wij nu steeds, tenzij anders wordt vermeld, onderstellen, dat de optredende matrices begrensd zijn. Dan kan dus zonder verder voorbehoud de gewone matrixrekening met Stieltjes-Lebesgue-Radon integratie in plaats van sommatie worden toegepast.
De in hoofdstuk 9 voorkomende $\sigma$-velden kunnen dan worden uitgebreid tot $\sigma$-velden. De stellingen en begrippen van hoofdstuk 9 zullen in de komende theorie bekend worden ondersteld.

## 5.6 Definities en grondbegrippen

In 1.4 definieerden wij een eindig wh-veld. Wij zagen daar, dat de wh-fct additief is. Een uitbreiding tot oneindige wh-velden ligt voor de hand.
Wij denken ons gegeven een vz $\pi$ en een $\sigma$-veld $\sigma_\pi$ van deelvzn van $\pi$, dat $\pi$ bevat, alsmede de vzn, bestaande uit één element (atomistisch is).
Wij noemen $\pi$ een waarschijnlijkheidsveld als op $\sigma_\pi$ tevens gedefiniëerd is een aftelbaar additieve verzamelingsfunctie $P$ (waarschijnlijkheidsmaat), waarvoor geldt:

(1) $\qquad \forall_{\Lambda}^{\sigma_\pi} \quad 0 \leq P_\Lambda \leq 1 = P_\pi$

(2) $\qquad \forall_{\{\Lambda_n\}}^{\mathfrak{D}(\Lambda)} \quad P_\Lambda = \sum_1^\infty P_{\Lambda_n}$

Wij eisen dus, dat $P$ aftelbaar additief is. Hierdoor wordt de theorie der wh-fcts een onderdeel van de maattheorie, hetgeen grote voordelen biedt van mathematische aard. Voor de theorie van de aftelbaar additieve vz-fcts zij weer verwezen naar hoofdstuk 9. Wij merken nog op, dat het nuttig is het $\sigma$-veld volledig te maken door er alle deelvzn van vzn met wh 0 aan toe te voegen.
In vele problemen uit de whr hebben wij te maken met verscheidene aftelbaar additieve vz-fcts, die alle op eenzelfde vz $\pi$ gedefiniëerd zijn; d.w.z. $P$ mag variëren over een andere vz $\Omega$, soms genaamd de vz van toegelaten hypothesen of parameterruimte. Kiezen wij een $\theta \in \Omega$, dan hoort daarbij een $P$, aan te geven met $P^\theta$, die op $\Lambda$ de waarde $P_\Lambda^\theta$ aanneemt. Wij onderstellen, dat $P_\Lambda^\theta$ bij vaste $\Lambda \in \sigma_\pi$ een fct van de eerste soort op $\Omega$ is, dus meetbaar is ten opzichte van een in $\Omega$ gegeven $\sigma$-veld $\sigma_\Omega$ en dat $P_\Lambda^\theta$ bij vaste $\theta \in \Omega$ een fct van de tweede soort op $\sigma_\pi$ is.
Als $\pi$ aftelbaar is, geldt $P_\Lambda^\theta = \sum_{\lambda \in \Lambda} P_\lambda^\theta$ en als $\pi$ en $\Omega$ beide

eindige vzn zijn, is $P^{\theta}_{\Lambda}$ bepaald door de gewone rechthoekige matrix met elementen $P^{\theta}_{\Lambda}$, $\lambda \in \pi$.
Wij kunnen $P^{\theta}_{\Lambda}$ dus als een (gegeneraliseerde)$(\Omega,\pi)$ matrix beschouwen en verwijzen voor de eigenschappen van deze matrices naar de appendix, hoofdstuk 9.
Een stochastische variabele $\underline{x}$ is een fct van de eerste soort op $\pi$, die waarden aanneemt, die tot een willekeurige vz $\mathfrak{J}$ behoren.
Dus als $\lambda \in \pi$ is, zal $x^{\lambda} \in \mathfrak{J}$ zijn.
Op $\mathfrak{J}$ moet ook een $\sigma$-veld gegeven zijn: $\sigma_{\mathfrak{J}}$
De meetbaarheidseis voor $x^{\lambda}$ betekent dan, dat

$\forall_S^{\sigma_{\mathfrak{J}}}$ $\mathcal{E}ns\{\lambda \mid x^{\lambda} \in S\} \in \sigma_{\pi}$ is, dus $x^{\lambda}$ beeldt de meetbare vzn van $\pi$ en $\mathfrak{J}$ op elkaar af.
Als $S \in \sigma_{\mathfrak{J}}$ is, dan bestaat

$$\mathcal{P}\{\underline{x} \in S\} \stackrel{\text{def}}{=} P(\lambda \in \pi \mid x^{\lambda} \in S),$$

waarbij $P$ de op $\sigma_{\pi}$ gedefinieerde wh-maat voorstelt.
Hebben wij twee stochastische variabelen $\underline{x}$ en $\underline{y}$, dan moet ook bestaan:

$$\mathcal{P}\{\underline{x} \in S_1 \wedge \underline{y} \in S_2\} \stackrel{\text{def}}{=} P(\lambda \in \pi \mid x^{\lambda} \in S_1,\ y^{\lambda} \in S_2).$$

Ook op de directe productvzn [1] van $\mathfrak{J}$ ($\mathfrak{J}\times\mathfrak{J}, \mathfrak{J}\times\mathfrak{J}\times\mathfrak{J}, \ldots$) onderstellen wij steeds een $\sigma$-veld gegeven, namelijk het veld, dat voortgebracht wordt door de vzn $S_1 \times S_2$, resp. $S_1 \times S_2 \times S_3$, waarbij $S_i \in \sigma$. Deze velden geven wij aan met $\sigma_{\mathfrak{J}^2}, \sigma_{\mathfrak{J}^3}$, enz.

-------------

1) Het directe product van twee vzn $A$ en $B$, $A\times B$ is gedefiniëerd d.m.v.

$$A \times B \stackrel{\text{def}}{=} \mathcal{E}ns\{(x,y) \mid x \in A,\ y \in B\}.$$

Bijvoorbeeld is het product van twee eendimensionale euclidische ruimten (reële getallenrechten) de tweedimensionale euclidische ruimte (platte vlak).
De definitie is alleen zinvol als in de productvz relaties van dezelfde soort als in de oorspronkelijke vzn gedefiniëerd kunnen worden.
Zo zal men in het product van twee topologische ruimten een topologische structuur trachten te definiëren, nauw verwant met die der oorspronkelijke ruimten. In het product van twee groepen kan men weer een groeprelatie definiëren. Hier hebben wij te maken met het product van meetbare ruimten (ruimten, elk met een $\sigma$-veld van deelvzn). In de productruimte kunnen wij eenvoudig een $\sigma$-veld definiëren, namelijk het $\sigma$-veld, voortgebracht door alle vzn van de vorm:

$$S_1 \times S_2 = \mathcal{E}ns\{(x,y) \mid x \in S_1 \in \sigma_{\mathfrak{J}_1},\ y \in S_2 \in \sigma_{\mathfrak{J}_2}\}$$

Zijn $\mathfrak{J}_1$ en $\mathfrak{J}_2$ de reële getallenrechte, $\sigma_{\mathfrak{J}_1}$ en $\sigma_{\mathfrak{J}_2}$ de stelsels van Borel-vzn in $\mathfrak{J}_1$, resp. $\mathfrak{J}_2$, dan is $\mathfrak{J}_1 \times \mathfrak{J}_2$ dus het platte vlak en is het volgens bovenstaande methode gedefiniëerde $\sigma$-veld in $\mathfrak{J}_1 \times \mathfrak{J}_2$ aequivalent met het stelsel van alle Borel-vzn in $\mathfrak{J}_1 \times \mathfrak{J}_2$ dus het $\sigma$-veld, dat voortgebracht wordt door de intervallen in $\mathfrak{J}_1 \times \mathfrak{J}_2$.

4. Een belangrijk bijzonder geval krijgen wij als $\mathcal{J}$ de vz $R$ van reële getallen voorstelt en het Markofproces invariant is bij translatie van $x$ en $X$ over de reële as, als dus:

(8) $\quad \forall_v^R \; P_{X+v}^{x+v} = P_X^x ,$

waarbij

(9) $\quad X+v \overset{def}{=} \mathcal{E}ns\{x+v \mid x \in X\}$ is, de over een afstand $v$ getranslateerde vz $X$.

Dus geldt dan:

(10) $\quad P_X^x = P_{X+(-x)}^o = p_{X-x} = \int p_{dv} \Big|_X^{x+v} ,$

waarbij

(11) $\quad p_z \overset{def}{=} P_z^o .$

Nu is:

$$\mathcal{P}\{\underline{x}_n \in X \mid \underline{x}_{n-1} = x\} = \mathcal{P}\{\underline{x}_n \in X - x \mid \underline{x}_{n-1} = o\} = \mathcal{P}\{(\underline{x}_n - \underline{x}_{n-1}) \in X - x \mid \underline{x}_{n-1} = v\},$$

waarbij $v$ willekeurig is.

Noemen wij $\underline{v}_k = \underline{x}_k - \underline{x}_{k-1}$ en $\underline{v}_o = \underline{x}_o$ , dus $\underline{x}_n = \underline{v}_o + \underline{v}_1 + \ldots\ldots + \underline{v}_n$ , dan volgt dus uit het voorgaande:

$\mathcal{P}\{\underline{v}_k \in X - x\} = p_{X-x}$ , dus de $\underline{v}_k$'s zijn o.o. verdeeld, ieder met verdelingsfct $p_X$.

Derhalve spreekt men van een <u>proces met onafhankelijke aangroeiingen</u> (process by independent increases, i.i. process).

Is de Markof keten niet stationair (maar wel enkelvoudig), dan moeten wij in (8), (10) en (11) aan $P$ en $p$ een index (n) toevoegen en zal dus $\underline{v}_k$ verdeeld zijn volgens $p_{(k)_X}$ , waarbij $P_{(k)_{X-x}} = P_{(k)_X}^x$.

5. Onlangs is bewezen, dat het hereditaire proces, bepaald dus door de in (3) gedefinieerde voorwaardelijke wh $P_{(n)_X}^{x_o,\ldots,x_{n-1}}$ een stationair Markof proces induceert in de ruimte van alle mogelijke wegen $x_o, x_1, \ldots\ldots, x_{n-1}$ en hier ook omgekeerd weer door bepaald is.

5.7 <u>Methode van de collectieve kenmerken</u>

Laat gegeven zijn een rij van whn $p_o, p_1, p_2, \ldots$ (anders gezegd een wh., afhangende van een index n), $o \leqq p_n \leqq 1$.

Voor vele problemen is het eenvoudiger om, in plaats van met de afzonderlijke whn $p_n$, te werken met de zogenaamde voortbrengende functie, gedefiniëerd door:

(1) $\varphi(z) = \sum_0^\infty p_n z^n.$

Deze methode is ingevoerd door Laplacè.
Wij voeren dus een hulpvariabele $z$ in. De reeks $\sum p_n z^n$ is in ieder geval convergent, als $|z|<1$ is, dus is binnen de eenheidscirkel in het complexe $z$-vlak een analytische fct van $z$.
Nu is een analytische fct volledig bepaald als de functiewaarden bekend zijn in een aftelbaar aantal punten met een verdichtingspunt (dat regulier punt van de fct is). In dat geval reeds zijn de getallen $p_n$ uit $\varphi(z)$ terug te vinden.
De volgende betrekking geldt namelijk:

$p_n = \frac{1}{2\pi i}\oint \varphi(z)\frac{dz}{z^{n+1}}$, waarbij geïntegreerd moet worden langs een gesloten kromme om de oorsprong, binnen de eenheidscirkel gelegen.
Wil men complexe integratie vermijden, dan zijn de $p_n$'s ook te berekenen uit de formule $p_n = \frac{\varphi^{(n)}(0)}{n!}$.
Het invoeren van de fct $\varphi(z)$ is natuurlijk alleen maar zinvol als wij de $p_n$'s weer uit $\varphi(z)$ kunnen berekenen. Uit het voorgaande volgt, dat dit zeker mogelijk is, als men de functiewaarden van $\varphi(z)$ kent in een willekeurig klein intervalletje om de oorsprong op de reële as.
Hebben wij een kategorisch systeem, is dus $\sum_0^\infty p_n = 1$, dan geldt $\varphi(1) = 1$.
Een generalisatie krijgen wij, wanneer wij elke $p_n$ met een andere variabele $z_n$ vermenigvuldigen en het geheel sommeren, dus:

(2) $\phi(z_0, z_1, \ldots) = \sum_0^\infty p_n z_n.$

Tussen de $z_n$'s kunnen wij allerlei afhankelijkheden laten bestaan. Ook geldt:

(3) $\varphi(z) = \phi(1, z, z^2, z^3, \ldots)$

Een convergentievoorwaarde is bij (2) weer nodig. Voldoende is bijvoorbeeld: $\forall_n\ |z_n| < \theta^n,\ 0<\theta<1$.
Wanneer de fct waarde van $\phi$ voor een voldoend groot aantal waarden van $z_0, z_1, z_2, \ldots$ bekend is, dan zijn de $p_n$'s weer te berekenen.
Is de functiewaarde bekend voor $z_n = 1$ en $z_i = 0$ als $i \neq n$, dan vin-

den wij direct:

(4) $\quad p_n = \phi(0,\dots,0,1,0,\dots)$

De $p_n$'s zijn dus direct terug te vinden als de waarde van $\phi$ bekend is voor de eenheidspunten van de aftelbaar vele coördinatenassen.

Hebben wij niet te maken met een aftelbaar aantal discrete whn, dan is een verdere generalisatie noodzakelijk en mogelijk:
Evenals in 5.6 zij gegeven een vz $\pi$ en een $\sigma$-veld $\sigma_\pi$ van deel-vzn van $\pi$, dat $\pi$ bevat, alsmede de vzn, bestaande uit één element. Tevens is gedefiniëerd een wh fct $P_\Lambda$, waarvoor (5.6.1) en (5.6.2) geldt.

Als generalisatie van (2) beschouwen wij nu:

(5) $\quad C(z) = \int_\pi P_{d\lambda} z^\lambda.$

Hierbij is $\lambda$ geen exponent, maar een element van $\pi$, $\lambda \in \pi$ ; $z^\lambda$ is dus een willekeurige fct van de eerste soort.
$C(z)$ bestaat in ieder geval als $z^\lambda$ begrensd is; $C(z)$ is een fct van een fct van $\lambda$, dus een lineaire <u>functionel</u>.
Wezenlijk is weer, dat $C(z)$ bekend is voor een niet al te kleine klasse van fcts $z^\lambda$, zodat de $P_\Lambda$'s terug te vinden zijn.
Is $C(z)$ bekend voor <u>alle</u> fcts van de eerste soort, dan vinden wij $P_\Lambda$ direct door voor $z^\lambda$ te substitueren $I_\Lambda^\lambda$, bij vaste $\Lambda$ een fct van de eerste soort:

(6) $\quad P_\Lambda = C(I_\Lambda)$

(zie de overeenkomst met (4)).
Tenslotte kunnen wij een klasse van wh fcts $P^\theta$ beschouwen, $\theta \in \Omega$, (zie het begin van 5.6) en definiëren:

(7) $\quad C(Z)_\Lambda = \int_\Omega Z_{d\theta} P_\Lambda^\theta,$

waarbij $Z$ aftelbaar additief is op $\sigma_\Omega$.
(Hoe is (7) te beschouwen als generalisatie van (2)?).
Als combinatie van (5) en (7) krijgen wij nu:

(8) $\quad C(Z,z) = \iint Z_{d\theta} P_{d\lambda}^\theta z^\lambda = \int_\Omega Z_{d\theta} \int_\pi P_{d\lambda}^\theta z^\lambda.$

De integratievolgorde doet niet ter zake onder weinig strenge voorwaarden (zie 9.6).
Als $C(Z,z)$ voor een voldoend grote klasse van fcts $Z$ en $z$ bekend is, dan zijn de $P$'s weer te berekenen, waaruit dan de $C$'s voor

een veel grotere klasse van fcts $Z$ en $z$ weer volgen.
Is $C(Z,z)$ bekend voor alle fcts van de eerste soort op $\Pi$ en alle fcts van de tweede soort op $\sigma_\Omega$ , waarvoor de integraal in (8) bestaat, dan zijn de $P$'s te vinden uit:

(9) $$P^\theta_\Lambda = C(1^\theta, 1_\Lambda).$$

De hulpfcts $z^\lambda$ en $Z_\Theta$ worden kenmerken genoemd, de erbij behorende functionel het collectieve kenmerk.

Wanneer $z^\lambda$ voor iedere $\lambda$ een wh is, wanneer dus $0 \leqq z^\lambda \leqq 1$ en wanneer $Z_\Theta$ een wh fct is, dus wanneer $Z_\Theta \geqq 0$ en $Z_\Omega = 1$ , dan kunnen wij aan $C(Z,z)$ een eenvoudige wh theoretische interpretatie geven:
Wij ~~interpreteren~~ $Z_\Theta$ ~~als de wh~~, dat met behulp van een lotingsmechanisme een $\theta \in \Theta$ gekozen wordt, waardoor dus een wh fct $P^\theta_\Lambda$ bepaald is.
Wordt nu met behulp van deze fct een $\lambda$ bepaald, dan onderstellen wij, dat er een kans $1-z^\lambda$ is, dat een gebeurtenis C (die niets te maken heeft met het beschouwde probleem) plaats vindt en dus een kans $z^\lambda$, dat C niet plaats vindt. De gebeurtenis C noemen wij een katastrophe.
Nu is $C(Z,z)$ de totale kans, dat geen katastrophe plaats vindt. Immers $\int_\Pi P^\theta_{d\lambda} z^\lambda$ is de voorwaardelijke kans, dat geen katastrophe plaats vindt als $\theta$ gekozen is. Vermenigvuldigen wij deze uitdrukking met de kans, dat $\theta$ gekozen wordt en integreren wij over $\theta$ , dan krijgen wij enerzijds $C(Z,z)$ , anderzijds ook de totale kans, dat geen katastrophe plaats vindt.
Wezenlijk is, dat de beide lotingsmechanismen voldoende onbepaald blijven. Hoe groter de klasse van fcts is, waarover $Z$ en $z$ kunnen variëren, des te groter is het aantal problemen, waarvoor het zin heeft de functionel $C(Z,z)$ in te voeren. $Z$ wordt hier dus beschouwd als een a priori-verdeling van hypothesen. De klasse van a priori-verdelingen moet ruim zijn.
Wij hebben $C$ en $z^\lambda$ geïnterpreteerd als de totale en voorwaardelijke wh, dat de katastrophe $C$ niet zal plaats vinden (in plaats van wel), daar dit een terminologisch voordeel biedt. Bezien wij immers m van elkaar onafhankelijke processen, waarbij $C_i$ de kans is, dat bij het $i$-de proces de katastrophe niet optreedt, dan is de totale kans, dat C niet d.w.z. geen enkele keer optreedt, gelijk aan $C_1 C_2 \ldots\ldots C_m$ , dus het product van de kansen $C_i$ . Dit is niet meer waar als $C_i$ de kans voorstelt, dat bij het $i$-de proces wel een katastrophe optreedt.

Tenslotte kunnen wij de beperkingen, die wij Z en z oplegden, waardoor de wh theoretische interpretatie mogelijk werd, weer opheffen, terwijl wij de interpretatie toch handhaven. We doen dan hetzelfde als in de analytische meetkunde gebeurt, wanneer men daar de meetkundige terminologie handhaaft en toch complexe getallen toelaat als waarden van de coördinaten van bijvoorbeeld punten.
Er is dus geen bezwaar tegen (zolang tenminste geen convergentiemoeilijkheden optreden) de wh theoretische terminologie te blijven gebruiken, ook al zijn de optredende whn niet "reëel" (waarmee wij bedoelen: ook al zijn zij geen reële getallen tussen 0 en 1).
Voor de methode van de collectieve kenmerken raadplege men in het bijzonder hoofdstuk 2 van de syllabus van de kadercursus Mathematische Statistiek (caput 3 Whr, zie [16] ), alsmede [18]. In de volgende twee hoofdstukken zullen wij met vrucht van deze methode gebruik maken.

## 6. Absorptieproblemen en Walds fundamentele identiteit bij stationaire Markof-processen

### 6.1 Collectieve matrix in verband met een absorberend gebied

Wij herinneren aan de in hoofdstuk 4 behandelde ruïneringsen absorptieproblemen, alsmede aan het probleem der sequente analyse - ook reeds in hoofdstuk 4 genoemd - waarbij als $(x,y)$ en $(x',y')$ punten in het XOY vlak voorstellen voor de matrix van overgangswhn geldt:

$$P^{(x,y)}_{(x',y')} = p \, /^{x}_{x'} \, /^{y+1}_{y'} + q \, /^{x+1}_{x'} \, /^{y}_{y'} \, .$$

Deze problemen zullen wij nu generaliseren.
Wij gebruiken weer de terminologie van de stochastische wandeling. Een wandelend, springend of bewegend punt springt, gaat of beweegt zich dus van de ene plaats naar een andere. De mogelijke plaatsen (toestanden) noemen wij ook wel punten. Een meetkundige illustratie is soms mogelijk.
In de sequente analyse vindt nu absorptie plaats als het bewegende punt in $A_1$ of $A_2$ gekomen is (Zie tekening 3 van 4.1). Wij kunnen ons dus indenken, dat in deze vzn "kleefstof" aangebracht is.
Generaliserend kunnen wij ons een vz $\mathcal{J}$ indenken met een $\sigma$-veld $\sigma_{\mathcal{J}}$ van deelvzn, alsmede een stationair Markof-proces op $\mathcal{J}$ , bepaald door de Markof-matrix $P^{x}_{X}$ (fct van tweeërlei soort, die

voldoet aan (5.6.7)). Wij denken ons verder een wandelend punt, dat, zijnde in $x$ met wh $P^x_X$ naar een punt van de vz $X$ kan springen. Absorptie vindt nu plaats, zodra het bewegende punt in een vz $A \subset \mathfrak{I}$ is aangeland, zodat de beweging alleen in $B = \mathfrak{I} - A$ kan plaats vinden. Hierbij is $\mathfrak{I}$ een willekeurige abstracte vz, boven was het de tweedimensionale ruimte.

Het is ook mogelijk, dat in elk punt $x \in \mathfrak{I}$ een klein beetje "kleefstof" is aangebracht, zodat in elk punt $x$ er een <u>kans</u> $A^x$ is, dat het springende punt, in $x$ aangeland, daar geabsorbeerd zal worden en een kans $B^x = 1 - A^x$, dat dit niet zal gebeuren. Nemen wij $A^x = I^x_A$, dan krijgen wij het vorige minder algemene geval terug.

Komt het bewegende punt in $x$ en vindt er absorptie in $x$ plaats (kans daarop $A^x$), dan stellen wij, dat er een kans $U^x$ is, dat een katastrophe $C$ niet zal plaats vinden. Wordt het bewegende punt daarentegen niet geabsorbeerd in $x$ (gaat het dus verder, kans $B^x$), dan stellen wij, dat er een kans $T^x$ is, dat de katastrophe niet geschiedt.

De kansen $U^x$ en $T^x$ zijn alleen ingevoerd om de in 5.7 beschreven methode van de collectieve kenmerken te kunnen gebruiken.

Wij noemen nu $C^x_X$ de totale wh, dat het wandelende punt, uitgaande van een willekeurig gegeven beginpunt $x$ ooit ergens in $X$ geabsorbeerd zal worden zonder dat de catastrophe van te voren heeft plaats gevonden.

$C^x_X$ is afhankelijk van de grootheden $U^x$ en $T^x$ en is bij vaste fcts $U$ en $T$ de collectieve matrix van het absorptieproces.

Bij gegeven $U^x$ en $T^x$ voldoet $C^x_X$ aan de integraalvergelijking:

$$(1) \qquad C^x_X = A^x U^x I^x_X + B^x T^x \int P^x_{dy} C^y_X .$$

Immers het punt start in $x$. Dan is er direct al een kans $A^x$ dat het geabsorbeerd zal worden, in welk geval er een kans $U^x$ is, dat de katastrophe niet plaats vindt. Deze eventualiteit geeft alleen een bijdrage tot $C^x_X$ als $x \in X$ is; vandaar de factor $I^x_X$. De andere mogelijkheid (met kans $B^x$) is, dat het punt niet geabsorbeerd wordt. In dat geval is er een kans $T^x$, dat in $x$ ook geen katastrophe plaats vindt, zodat $B^x T^x$ de kans is, dat het een sprong zal doen. Het springt dan naar een punt $y$ (overgangswh $P^x_{dy}$ als $dy$ een "kleine verzameling" voorstelt waar $y$ toebehoort) en heeft vervolgens een kans $C^y_X$, dat het in $X$ geabsorbeerd wordt zonder dat de katastrophe van te voren heeft plaats gevonden. Over $y$ moet nog geïntegreerd worden. Hiermee

is (1) juist bevonden.
Wij nemen nu verder aan, dat $\mathcal{U}_{x=1}$ is voor alle $x \in \mathfrak{I}$; dat wil dus zeggen, dat de katastrophe alleen plaats kan vinden in een punt als het bewegende punt daar niet is geabsorbeerd. Wij krijgen dan:

$$(2) \qquad C^x_X = A^x I^x_X + B^x T^x \int P^x_{dy} C^y_X .$$

Deze integraalvergelijking trachten wij door itereren op te lossen:

$$C^x_X = A^x I^x_X + B^x T^x \int P^x_{dy} A^y I^y_X + B^x T^x_X \int P^x_{dy} B^y T^y \int P^y_{dz} C^z_X$$

dus

$$(3) \qquad C^x_X = A^x I^x_X + \sum_1^{\infty} B^x T^x \int P^x_{dy_1} B^{y_1} T^{y_1} \int P^{y_1}_{dy_2} B^{y_2} T^{y_2} \int \ldots \int P^{y_{k-1}}_{dy_k} A^{y_k} I^{y_k}_X .$$

Wil (3) een oplossing zijn van (2), dan is nodig en voldoende, dat de oneindige reeks in (3) convergeert. Dit is zeker zo, als $\|T\| = \vartheta < 1$ is, daar dan de n-de term van de reeks in absolute waarde hoogstens gelijk is aan $1.\vartheta.1.1.\vartheta.1.1.\vartheta.\ldots.\vartheta.1.1.1 = \vartheta^n$ (immers $B^x \leq 1$, $P^x_X \leq 1$, $|T^x| \leq \vartheta$), zodat de convergente reeks $\sum \vartheta^n$ een majorant is.
In het speciale geval, dat $A^x = I^x_A$ is, krijgen wij:

$$(4) \qquad C^x_X = I^x_A I^x_X + I^x_B T^x \int P^x_{dy} C^y_X .$$

Hiervoor kunnen wij schrijven:

$$C^x_X = (I_A)^x_X + \int (I_B)^x_{dy} \int T^y_{dz} \int P^z_{dv} C^v_X ,$$

wanneer $T^x_X$ de door $T^x$ bepaalde diagonaalmatrix voorstelt, dus

$$(5) \qquad T^x_X \overset{\text{def}}{=} (TI)^x_X = T^x I^x_X .$$

Wij kunnen (4) dus in matrixvorm schrijven:

$$(6) \qquad C = I_A + I_B T P C$$

Itereren geeft:

$$C = I_A + I_B T P I_A + I_B T P I_B T P C = I_A + I_B T P I_A + I_B T P I_B T P I_A + I_B T P I_B T P I_B T P C .$$

De convergentievoorwaarde blijft dezelfde, dus er is absolute convergentie als $\|T\| < 1$.
Dan is dus:

$$(7) \qquad C = \sum_{o}^{\infty} (/_B TP)^n /_A = (/ - /_B TP)^{-1} /_A$$

(zie 9.7 voor machtreeksen in een matrix)
Dit resultaat kunnen wij veel eenvoudiger als volgt krijgen:
Uit (6) volgt:

$(/ - /_B TP)C = /_A$, dus als $(/ - /_B TP\ )$ een inverse heeft (en dit is zeker zo als $\|T\| < 1$ is, dan is

$$C = (/ - /_B TP)^{-1} /_A .$$

Ook (2) is in matrixvorm te schrijven, wanneer wij de volgende diagonaalmatrices invoeren:

$$(8) \qquad A^x_X \overset{def}{=} A^x /^x_X \quad \text{en} \quad B^x_X \overset{def}{=} B^x /^x_X .$$

Uit (2) volgt dan:

$$(9) \qquad C = A + BTPC$$

dus $(/ - BTP)C = A$. Heeft $/ - BTP$ een ondubbelzinnig bepaald inverse, dan is

$$(10) \qquad C = (/ - BTP)^{-1} A = \left\{ \sum_{o}^{\infty} (BTP)^m \right\} A .$$

Het laatste lid ontstaat door $(/ - BTP)$ formeel in een reeks te ontwikkelen.
Itereren van (9) geeft als identiteit:

$$(11) \qquad C = \sum_{o}^{N-1} (BTP)^n A + (BTP)^N C .$$

Als $o \leqq T^x \leqq \vartheta < 1$ is dan zal

$$\|(BTP)^N C\| \leqq \|BTP\|^N . \|C\| \leqq (1.\vartheta.1)^N . 1 = \vartheta^N \to o \quad \text{als} \quad N \to \infty$$

dus

$$(12) \qquad C = \sum_{o}^{\infty} (BTP)^n A \quad \text{voor} \quad \|T\| < 1 .$$

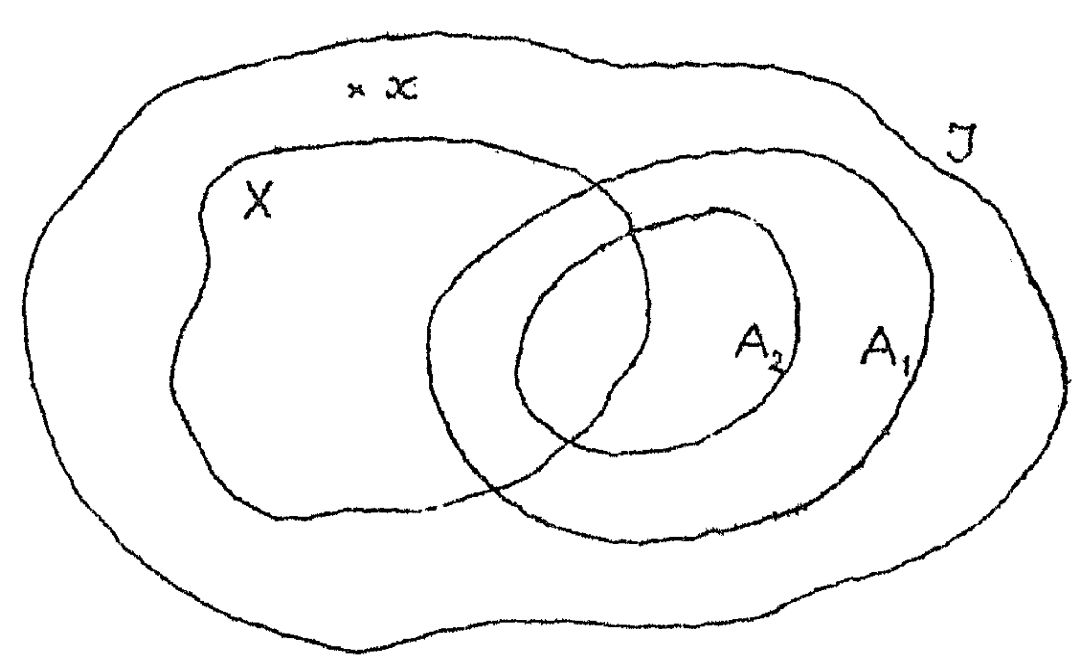
Wij weten echter alleen, dat $\|T\| \leqq 1$ is, zodat wij het geval $\|T\| = 1$ nog moeten onderzoeken. Als wij in (12) voor $T^x$ invullen $\theta T^x$, waarbij $\theta$ een reëel getal is, dan convergeert het rechterlid zeker als $o \leqq \theta < 1$ is. Daar C een wh is, dus $\leqq 1$, zal de reeks begrensd zijn als $o \leqq \theta < 1$ is en begrensd blijven als $\theta \to 1$. Verder zijn alle termen van de reeks $\geqq o$. Hieruit volgt, dat de reeks convergent blijft als $\theta = 1$ is, terwijl zijn

waarde $\leqq 1$ blijft. Dus geldt (12) niet alleen voor $\|T\| < 1$, doch ook voor $\|T\| = 1$.
Wij kunnen nu voor $T$ ook complexe waarden toelaten. De uitdrukking (12) blijft zeker gelden voor $\|T\| \leqq 1$.
Voor $\|T\| \leqq 1$ behoeft (10) niet meer waar te zijn. Hier is namelijk $A$ buiten haakjes gehaald, $\sum_0^\infty (BTP)^n$ is geen wh, dus behoeft niet begrensd te blijven. Het is dus mogelijk, dat $\sum (BTP)^n A$ convergent is, doch $\{\sum (BTP)^n\} A$ niet.

Voor $T^x = 1$, dus als de katastrophe in het geheel niet plaats vindt, stelt (10) de totale wh voor, dat een bewegend punt, beginnend in $x$, ooit ergens in $X$ geabsorbeerd zal worden.

## 6.2 Twee absorberende gebieden

Wij eisen nu weer: $A^x = I_A^x$, en beschouwen twee absorberende vzn $A_1$ en $A_2$ met $A_2 \subset A_1$ (zie figuur)




Wij kunnen ons nu twee soorten springende punten voorstellen; die van de eerste soort worden in $A_1$ geabsorbeerd, die van de tweede soort hebben geen last van $A_1$, doch worden in $A_2$ wel geabsorbeerd.
Zo men wil kan men hier denken aan twee soorten "vliegenlijm" en twee soorten springende "insecten". De collectieve matrices, behorende bij de vzn $A_1$ en $A_2$ en in beide gevallen dezelfde fct $T^x$ noemen wij $C_{(A_1)}$ en $C_{(A_2)}$.
Nu geldt de

Stelling

(1) $\quad C_{(A_1)} \cdot C_{(A_2)} = C_{(A_2)} = C_{(A_2)} \cdot C_{(A_1)}$ als $A_2 \subset A_1$.

Dus: het product van twee collectieve matrices, behorende bij twee absorberende vzn, waarvan de een de ander omvat, is commutatief en gelijk aan de collectieve matrix van de kleinste vz.
Bewijs:

1. De tweede gelijkheid van (1) is eenvoudig te bewijzen.
Wij hebben:

(6.1.6) $C_{(A)} = I_A + I_B T P C_{(A)}$,

waaruit volgt:

(2) $I_A C_{(A)} = I_A$.

(Bedenk namelijk, dat $I_A . I_A = I_A$ en $I_A . I_B = I_{A \cap B} = 0$ want $A \cap B = 0$.)

(3) $I_B C_{(A)} = I_B T P C_{(A)}$,

dus

(4) $C_{(A)} = I_A + I_B C_{(A)}$.

Verder volgt uit (6.1.7)

(5) $C_{(A)} I_A = C_{(A)}$.

Met behulp van deze betrekkingen vinden wij:

$$C_{(A_2)} C_{(A_1)} = C_{(A_2)} I_{A_2} \{ I_{A_1} + I_{B_1} C_{(A_1)} \} = C_{(A_2)} I_{A_2} = C_{(A_2)},$$

daar $I_{A_2} I_{A_1} = I_{A_2 \cap A_1} = I_{A_2}$ $(A_2 \subset A_1)$ $\quad I_{A_2} . I_{B_1} = 0$ $(A_2 \cap B_1 = 0)$.

2. De eerste gelijkheid van (1) zullen wij op twee manieren bewijzen.

Het tweede bewijs zal in paragraaf 6.4 gegeven worden en is algebraïsch. Nu geven wij eerst een bewijs, waarbij gebruik wordt gemaakt van de wh theoretische interpretatie.

Hiertoe beschouwen wij de volgende gebeurtenis $\mathcal{A}$ : het wandelende punt komt vanuit $x$ in $A_2 \cap X$ zonder dat de katastrophe $\mathcal{C}$ plaats gevonden heeft.

Dan is de wh, dat plaats vindt, en dus geldt:

(6) $C_{(A_2)X}^{\;x} = C_{(A_1)A_2 \cap X}^{\;x} + \int_{D} C_{(A_1)dy}^{\;x} C_{(A_2)X}^{\;y}$, waarbij $D \stackrel{def}{=} A_1 -- A_2$.

Om dit te bewijzen beschouwen wij een wandelend punt van de tweede soort, dat dus niet in $A_1$, maar wel in $A_2$ wordt geabsorbeerd.

Als $\mathcal{A}$ plaats vindt, dan zijn er twee elkaar uitsluitende mogelijkheden:

a. Het wandelende punt $W$ komt vanuit $x$ in $A_2 \cap X$ en wordt daar geabsorbeerd zonder dat het ooit in $D$ is geweest.

b. Het wandelende punt komt vanuit $x$ in $A_2 \cap X$ nadat het reeds een of meer keren in $D$ is geweest.

De kans, dat mogelijkheid a. zich voordoet is $C_{(A_1)A_2 \cap X}^{\;x}$.

Doet b. zich daarentegen voor, dat is er dus een punt $y \in D$ waar $W$ de eerste keer, dat het in $D$ was, verbleef.
De wh, dat het zal komen in een punt $y$ van een kleine vz $Y$ is weer gelijk aan de wh, dat het in $Y$ geabsorbeerd zou worden als $A_1$ de absorberende vz was, dus is $C_{(A_1)Y}^{x}$.
Immers $C_{(A_1)Y}^{x}$ is op te vatten als de wh, dat $W$, uitgaande van $x$ in $Y$ komt zonder eerst ergens anders in $A_1$ te zijn geweest. En deze wh moeten wij precies hebben, daar $y$ het eerste punt van $D$, dus van $A_1$ is, waar $W$ komt. (Zou $W$ nog eerder in $A_2 = A_1 -- D$ gekomen zijn, dan zou mogelijkheid b. zich niet voorgedaan hebben, hetgeen wij onderstellen).
Is $W$ eenmaal in $y$, dan is de wh, dat $\mathcal{A}$ zich voordoet, gelijk aan $C_{(A_2)X}^{y}$ geworden. Over alle $Y$ van een dissectie van $D$ moet nog gesommeerd worden, dus na limiet overgang moet over $D$ geïntegreerd worden. Hieruit volgt (6).
In matrixnotatie luidt (6):

(7) $$C_{(A_2)} = C_{(A_1)} I_{A_2} + C_{(A_1)} I_D C_{(A_2)}.$$

Uit $D = A_1 -- A_2 = B_2 -- B_1 = B_2 \cap A_1$ volgt:

$$I_D = I_{B_2} I_{A_1} = I_{B_2}(I - I_{B_1}) = I_{B_2} - I_{B_1} \quad (B_1 \subset B_2).$$

Met behulp van (5) en het bovenstaande krijgen wij nu:

$$C_{(A_1)} I_D = C_{(A_1)} I_{A_1} I_D = C_{(A_1)}(I_{A_1} I_{B_2} - I_{A_1} I_{B_1}) = C_{(A_1)} I_{A_1} I_{B_2} = C_{(A_1)} I_{B_2}. \quad (A_1 \cap B_1 = 0)$$

Dus: $$C_{(A_2)} = C_{(A_1)} I_{A_2} + C_{(A_1)} I_{B_2} C_{(A_2)} = C_{(A_1)}\{I_{A_2} + I_{B_2} C_{(A_2)}\} = C_{(A_1)} C_{(A_2)} \quad \text{(zie (4))}$$

Hiermee is de eerste gelijkheid van (1) bewezen.

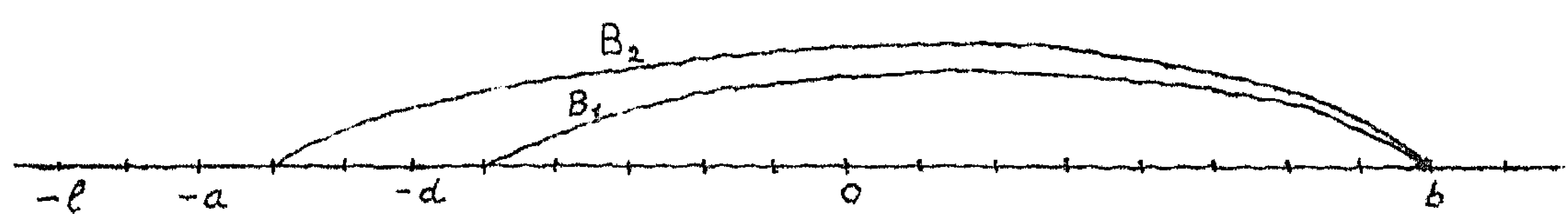
## 6.3 Aanvullende opmerkingen

Nemen wij $A_1 = A_2$, dan volgt uit (6.2.1):

(1) $$C_{(A)}^2 = C_{(A)},$$

dus $C_{(A)}$ is een idempotente matrix. (Dit volgt reeds uit de tweede gelijkheid van (6.2.1)). In zijn dissertatie beschouwde J.H.B. Kemperman het volgende speciale geval:

1. $\mathfrak{J} = I$ (de vz van alle gehele getallen);
2. $B_1$ en $B_2$ zijn eindige "intervallen". Zie de tekening, waarbij de gehele getallen door aequidistante punten op een rechte lijn voorgesteld zijn:




$B_1 = \mathcal{E}ns\,\{x \in I \mid -d < x < b\}, \quad 0 \in B_1$

$B_2 = \mathcal{E}ns\,\{x \in I \mid -a < x < b\}, \quad 0 < b, \quad 0 < d < a.$

3. $\forall_x^I \; P^{x+k}_{y+k} = P^x_y = P^0_{y-x}.$

Wij hebben dus te maken met een proces met onafhankelijke aangroeiingen.

4. $T^x = T$ (constant).

Als uitgangspunt $x$ kiezen wij 0.

$$D = A_1 -- A_2 = B_2 -- B_1 = \mathcal{E}ns\,\{-j \in I \mid -a < -j \leq -d \text{ dus } d \leq j < a\}.$$

Wh theoretisch volgt uit voorwaarde 3. direct:

(2) $\quad C_{(A)}{}^{x}_{y} = C_{(A-x)}{}^{0}_{y-x},$

waarbij:

(3) $\quad A - x \overset{def}{=} \mathcal{E}ns\,\{y \mid x + y \in A\}.$

De collectieve matrix $C_{(A)}{}^{x}_{y}$ stelt immers een wh voor (zie 6.1). Deze wh verandert niet in grootte - daar wij met een proces met onafhankelijke aangroeiingen te maken hebben - wanneer wij alle punten over een afstand $x$ verschuiven, mits ook de absorberende vz $A$ over dezelfde afstand $x$ verschoven wordt.
Uit (6.2.6) volgt nu als wij $X = \{-\ell\}$ nemen, $-\ell \leq -a$ :

(4) $\quad C_{(A_2)}{}^{0}_{-\ell} = C_{(A_1)}{}^{0}_{-\ell} + \sum_{d}^{a-1} C_{(A_1)}{}^{0}_{-j}\, C_{(A_2)}{}^{-j}_{-\ell},$

dus met behulp van (2):

(5) $\quad C_{(A_2)}{}^{0}_{-\ell} = C_{(A_1)}{}^{0}_{-\ell} + \sum_{d}^{a-1} C_{(A_1)}{}^{0}_{-j}\, C_{(A_2+j)}{}^{0}_{-\ell+j}.$

Noem

(6) $\quad C_{(A_2)}{}^{0}_{-\ell} = \varphi(b, a, \ell),$

dan geldt dus:

(7) $\quad \varphi(b, a, \ell) = \varphi(b, d, \ell) + \sum_{d}^{a-1} \varphi(b, d, j)\, \varphi(b+j, a-j, \ell-j).$

Zie verder [42], blz. 71.

Met behulp van (6.2.6) of (6.2.7) kan $C_{(A_2)}{}^x_X$ worden berekend voor alle $x$ en $X$, wanneer de matrix $C_{(A_1)}{}^x_X$ volledig bekend is. Wij kunnen echter $C_{(A_2)}$ ook direct in $C_{(A_1)}$ uitdrukken:
Daar $D = A_1 \cap B_2$, dus $I_D = I_{A_1} \cdot I_{B_2}$ is en daar (6.2.3) en (6.2.5) gelden, krijgen wij:

$$C_{(A_1)} I_D C_{(A_2)} = C_{(A_1)} I_{A_1} I_{B_2} C_{(A_2)} = C_{(A_1)} I_{A_1} I_{B_2} TPC_{(A_2)} = C_{(A_1)} I_D TPC_{(A_2)}.$$

Hierdoor gaat (6.2.7) over in:

$$C_{(A_2)} = C_{(A_1)} I_{A_2} + C_{(A_1)} I_D TPC_{(A_2)} = C_{(A_1)} \{ I_{A_2} + I_D TPC_{(A_2)} \}.$$

Door op de bekende manier iteratie toe te passen krijgen wij:

$$(8) \qquad C_{(A_2)} = C_{(A_1)} \sum_{n=0}^{\infty} \{ I_D TPC_{(A_1)} \}^n I_{A_2},$$

waarvoor wij, als $\|T\| < 1$ is, kunnen schrijven:

$$(9) \qquad C_{(A_2)} = C_{(A_1)} (I - I_D TPC_{(A_1)})^{-1} I_{A_2}$$

Als $\|T\| = 1$ is, blijkt (8) convergent te blijven.
Nemen wij in (9) $A_1 = \mathfrak{J}$ (dus $C_{(A_1)}{}^x_X = C_{(\mathfrak{J})}{}^x_X = I^x_X$, daar absorptie direct in $x$ al plaats vindt, $A_2 = A$, dan is $I_D = I_{\mathfrak{J}-A} = I_B$, zodat (9) een generalisatie blijkt te zijn van (6.1.7).
Vervolgens beschouwen wij het geval, dat $D$ uit één enkel punt bestaat: $D = \{d\}$.
Wij definiëren:

$$(10) \qquad Q \stackrel{\text{def}}{=} (I - I_D TPC_{(A_1)})^{-1} - I.$$

Dan is:

$$(11) \qquad (I - I_D TPC_{(A_1)})(I + Q) = I,$$

of

$$(12) \qquad Q = I_D TPC_{(A_1)} I + I_D TPC_{(A_1)} Q,$$

dus

$$(13) \qquad Q^x_X = I^x_D T^x (PC_{(A_1)} + PC_{(A_1)} Q)^x_X.$$

Hieruit volgt, dat $Q^x_X = 0$ is, als $x \neq d$ is.
Als $x = d$ is, dan krijgen wij:

$$(14) \qquad Q^d_X = \frac{T^d (PC_{(A_1)})^d_X}{1 - T^d (PC_{(A_1)})^d_d}.$$

Uit (9), (10) en (14) volgt nu:

$$(15)\qquad C_{(A_2)\,X}^{\;x} = C_{(A_1)\,A_2\cap X}^{\;x} + C_{(A_1)\,d}^{\;x}\, Q_{A_2\cap X}^{\;d} =$$

$$= C_{(A_1)\,A_2\cap X}^{\;x} + C_{(A_1)\,d}^{\;x}\,\frac{T^d\,(P C_{(A_1)})_{A_2\cap X}^{d}}{1 - T^d\,(P C_{(A_1)})_d^d}$$

Wanneer $\mathfrak{I}$ aftelbaar is, kunnen wij door recursie met behulp van (15) de $C_{(A_n)}$ successievelijk berekenen, indien wij $\mathcal{B}_0 = 0$ nemen (dus $C_{(A_0)} = I$ ) en vervolgens steeds één element aan $\mathcal{B}_n$ toevoegen.

Ter vereenvoudiging van de notatie nemen wij aan, dat de elementen $x, y, z, \ldots$ van $\mathfrak{I}$ de natuurlijke getallen $1, 2, 3, \ldots$ zelf zijn, zo genummerd, dat $\mathcal{B}_n = \{1, \ldots\ldots, n\}$ dus $A_n = \{n+1, n+2, \ldots\}$ is. Verder schrijven wij in plaats van $C_{(A_n)}$ kortweg $C_n$.

Voor $y \geqq n$ zal

$$(P C_{n-1})_y^n = \sum_1^{\infty} {}^{z}\, P_z^n\, (C_{n-1})_y^z = \sum_1^{n-1} {}^{z}\, P_z^n\, (C_{n-1})_y^z + P_y^n\,,$$

daar $(C_{n-1})_y^x = 0$ is voor $x \geqq n$ , tenzij $x = y$ is ($(C_{n-1})_x^x = 1$ voor $x \geqq n$).

De recursieve betrekking wordt nu:

$$(C_n)_y^x = 0 \text{ voor } y \leqq n \text{ en}$$

$$(16)\qquad (C_n)_y^x = (C_{n-1})_y^x + (C_{n-1})_n^x \Big\{1 - T^n P_n^n - T^n \sum_1^{n-1} {}^{z}\, P_z^n (C_{n-1})_n^z\Big\}^{-1} T^n \Big\{P_y^n + \sum_1^{n-1} {}^{z}\, P_z^n (C_{n-1})_y^z\Big\}$$

$$\text{voor } y \geqq n+1.$$

Wij herinneren er aan, dat $(C_{n-1})_y^x$ en $(C_n)_y^x$ functies zijn van de oneindige reeks variabelen $T^1, T^2, T^3, \ldots$ (de bovenindices zijn geen exponenten).

Tenslotte de volgende opmerking:

In 8.5 hebben wij de matrices $E_\rho$ ingevoerd. Deze speelden een rol bij de constructie van de normaalvorm van een matrix. Er bestaat een zekere analogie tussen de matrices $E_\rho$ en $C_{(A)}$. Definiëren wij $E_\pi = \sum_{\rho \in \pi} E_\rho$ , dan volgt uit (8.5.6), dat $E_\pi^2 = E_\pi$ , zodat $E_\pi$ evenals $C_{(A)}$ idempotent is.

Als $\pi' \subset \pi$ is, dan is $E_\pi E_{\pi'} = E_{\pi'} = E_{\pi'} E_\pi$ (analoog aan (6.2.1).

Voor de $E_\rho$'s geldt verder nog:

$\sum_{\rho \in \mathfrak{I}} E_\rho = I$ (8.5.5), benevens als $\pi' = \mathfrak{I} - \pi$ is: $E_\pi E_{\pi'} = 0$ en $E_\pi + E_{\pi'} = I$.

Deze eigenschappen vinden wij bij de $C_{(A)}$ niet terug. Of de analogie tussen $C_{(A)}$ en $E_\pi$ nog verder gaat, is vooralsnog onbekend.

## 6.4 Tweede bewijs van (6.2.1). Verdere resultaten

Wij zullen nu dus een algebraïsch bewijs geven van de eerste gelijkheid van (6.2.1):

$$C_{(A_1)}\, C_{(A_2)} = C_{(A_2)} \quad \text{als } A_2 \subset A_1 \text{ is.}$$

Wij gaan uit van (6.2.4). Hiervoor kunnen we schrijven:

$$I - C_{(A)} = I - I_A - I_B TPC_{(A)} = I_B(I - TPC_{(A)}) = I_B(I-TP) + I_B TP(I - C_{(A)}),$$

waaruit volgt:

(1) $$(I - I_B TP)(I - C_{(A)}) = I_B(I - TP),$$

dus voor $\|T\| < 1$ :

(2) $$I - C_{(A)} = (I - I_B TP)^{-1} I_B (I - TP).$$

Dus geldt ook:

(3) $$I - C_{(A)} = \left\{ \sum_0^{\infty} (I_B TP)^n \right\} I_B (I - TP),$$

welke betrekking ook door iteratie uit (1) te verkrijgen is en dus geldt voor alle $T$ , waarvoor het rechterlid convergent is. Wij vermenigvuldigen (1) en (2) nu van rechts met een willekeurige begrensde matrix $D_X^*$ , dus nemen voorlopig $\|T\| < 1$ .
Daar wij met begrensde matrices te doen hebben is de matrixvermenigvuldiging associatief, zodat uit (1) en (2) volgt:

(4) $$C_{(A)} D = D \Longleftrightarrow I_B D = I_B TPD.$$

Dus als het linkerlid van (4) waar is voor een begrensde matrix $D$ , dan volgt daaruit het rechterlid en omgekeerd.
Neem nu $A = A_1, B = B_1$ en $D = C_{(A_2)}$ (inderdaad begrensd), dan volgt uit (4):

(5) $$I_{B_1} C_{(A_2)} = I_{B_1} TPC_{(A_2)} \longrightarrow C_{(A_1)} C_{(A_2)} = C_{(A_2)}.$$

Wij moeten dus alleen bewijzen, dat de eerste gelijkheid juist is. Dit is direct duidelijk, wanneer wij (6.2.3) opschrijven voor $A = A_2$ , $B = B_2$ en dan beide leden van links met $I_{B_1}$ vermenigvuldigen ( $B_1 \subset B_2$ , dus $I_{B_1} I_{B_2} = I_{B_1}$ ). Hiermee is het gestelde bewezen.
Wij kunnen uit (4) nog andere interessante resultaten afleiden. Daar de vz $X$ , waar van $D_X^*$ afhangt, er niet toe doet, zolang $D$ in de vergelijkingen alleen als laatste, meest rechtse factor

voorkomt, kunnen wij $D_x^x$ vervangen door een begrensde fct van de eerste soort $f^x$.
Uit (4) volgt dan:

$$(6) \qquad C_{(A)} f = f \rightleftarrows 1_B f = 1_B T P f ,$$

dus

$$(7) \qquad \int C_{(A)}{}^x_{dy} f^y = f^x \rightleftarrows 1_B^x f^x = 1_B^x T^x \int P_{dy}^x f^y .$$

De linkergelijkheid houdt in, dat $f$ een eigen functie van $C$ is bij de eigen waarde 1. Als in de rechtergelijkheid de factor $1_B^x$ zou ontbreken en als $T^x = T$ (constant) was, dan was $f$ dus tevens een eigen functie van $P$ bij de eigen waarde $\frac{1}{T}$ . De factor $1_B^x$ maakt, dat de bewering slechts geldt voor $x \in B$ . Als $x \notin B$ gaat de rechtergelijkheid over in de triviale gelijkheid 0=0.
In het algemeen is $T^x$ echter niet constant.
In de belangrijkste toepassing echter is $f$ niet begrensd. Wij zullen de associativiteit van de matrixvermenigvuldiging, die wij in het bovenstaande gebruikten, nu moeten bewijzen. Hiertoe zullen wij gebruik maken van stelling 3 van 9.5.
Wij willen namelijk de volgende stelling bewijzen:

Stelling

Wanneer $f_0^x \geq 0$ en $T_0^x \geq 0$ bestaan, zodanig dat:

$$(8) \qquad \forall_x^B f_0^x \geq T_0^x (P f_0)^x \quad \text{en}$$

$$(9) \qquad \| 1_B P T_0 1_B \| = \sup_{x \in B} \int_B P_{dy}^x T_0^y \leq 1 ,$$

dan zal (7) gelden voor alle $f^x$ en $T^x$, die voor $x \in B$ voldoen aan $|f^x| \leq c f_0^x$ en $|T^x| \leq \theta T_0^x$ ($0 \leq \theta \leq 1$, $c$ een constante).[1]
Bewijs: De volgende betrekkingen gelden:

$$(6.1.7) \qquad C = \sum_0^{\infty} {}_n (1_B T P)^n 1_A ,$$

$$(6.4.3) \qquad 1 - C = \left\{ \sum_0^{\infty} {}_n (1_B T P)^n \right\} 1_B (1 - T P) ,$$

wanneer tenminste de rechterleden convergeren. Dit is inderdaad het geval, zoals wij nu eerst zullen bewijzen:

N.l. is: $\quad [(1_B |T| P)^n]_X^x = \theta^n [(1_B T_0 P)^n]_X^x \leq \theta^n 1_B^x T_0^x .$

-------------

1) Wij beperken ons dus niet tot fcts $T^x$ , waarvoor $\forall_x \; 0 \leq T^x \leq$ maar laten alle mogelijke fcts van de eerste soort voor $T$ toe.

Immers voor $n=1$ is:

$$(1_B T_0 P)^x_X = 1^x_B T_0^x P^x_X \leqq 1^x_B T_0^x .$$

Stel nu, dat $\left[(1_B T_0 P)^n\right]^x_X \leqq 1^x_B T_0^x$ is, dan volgt hieruit:

$$\left[(1_B T_0 P)^{n+1}\right]^x_X = \left[(1_B T_0 P)(1_B T_0 P)^n\right]^x_X \leqq \{(1_B T_0 P)(1_B T_0)\}^x =$$

$$= \int 1^x_B T_0^x P^x_{dy} 1^y_B T_0^y = 1^x_B T_0^x \int_B P^x_{dy} T_0^y \leqq 1^x_B T_0^x$$

volgens (9), zodat het gestelde met volledige inductie bewezen is.

Hieruit volgt:

$$\forall^{\sigma_y}_X \left\{ \sum_0^\infty n (1_B |T| P)^n \right\}^x_X \leqq \sum_0^\infty n \theta^n 1^x_B T_0^x = (1-\theta)^{-1} 1^x_B T_0^x ,$$

zodat $C$ en het rechterlid van (4) absoluut genomen bestaan.

Noemen wij $R \overset{def}{=} \sum_0^\infty n (1_B |T| P)^n$, dan bestaat $R$ dus, alsmede $\sum_0^\infty n (1_B T P)^n$.

Eveneens bestaat $(R|f|)^x$.

Immers $(R|f|)^x = \left\{ \sum_0^\infty n (1_B |T| P)^n . |f| \right\}^x \leqq c . \sum_0^\infty n \theta^n . \{(1_B T_0 P)^n f_0\}^x .$

(De verwisseling van sommatie en integratie is toegestaan, daar $R$ bestaat en positief is).

Verder is $\{(1_B T_0 P)^n f_0\}^x \leqq (1_B f_0)^x$ voor alle n, hetgeen weer door inductie te bewijzen is:

$n=1$: $\{(1_B T_0 P) f_0\}^x = 1^x_B T_0^x (P f_0)^x \leqq 1^x_B f_0^x$ volgens (8).

Stel het gestelde is bewezen voor n.

$$\{(1_B T_0 P)^{n+1} f_0\}^x = \{[(1_B T_0 P)^n (1_B T_0 P)] f_0\}^x = \{(1_B T_0 P)^n (1_B T_0 P) f_0\}^x$$ volgens

stelling 3 van 9.5, daar $[(1_B T_0 P) f_0]^x$ en $\{(1_B T_0 P)^n (1_B T_0 P)\}^x_X$ volgens het voorgaande bestaan voor alle $x \in \mathfrak{J}$, respectievelijk $X \in \sigma_{\mathfrak{J}}$ en:

$$\{(1_B T_0 P)^n (1_B T_0 P) f_0\}^x \leqq \{(1_B T_0 P)^n 1_B f_0\}^x \leqq \{(1_B T_0 P)^n f_0\}^x \leqq (1_B f_0)^x < \infty$$

is volgens inductie-onderstelling.

Dus $(R|f|)^x \leqq c\,(1-\theta)^{-1} (1_B f_0)^x$, bestaat dus.

Uit (6.1.7) volgt nu, dat $(Cf)^x$ bestaat en dus ook

$$\{(1-C)f\}^x = f^x - \int C^x_{dy} f^y .$$

Volgens (6.4.3) geldt:

(10) $$\{(1-C)f\}^x = \left[\left\{ \sum_0^\infty n (1_B T P)^n 1_B (1-TP) \right\} f\right]^x = \left[ \sum_0^\infty n (1_B T P)^n 1_B (1-TP) f \right]^x ,$$

weer volgens stelling 3 van 9.5, daar:

1. $\left[\sum_{0}^{\infty}{}^{n}(1_B T P)^n 1_B (1-TP)\right]_X^x$ bestaat voor alle $X\epsilon\sigma_{\mathfrak{J}}$ en $x\epsilon\mathfrak{J}$.

2. $(1_B(1-TP)f)^x$ bestaat voor alle $x\epsilon\mathfrak{J}$.

Immers $(|1_B TPf|)^x \leq (1_B |T|.P.|f|)^x \leqq c\,\theta\, 1_B^x\, T_0^x (Pf_0)^x \leqq c\,\theta\, 1_B^x f_0^x$.

3. $\forall_x^{\mathfrak{J}} \left\{\sum_{0}^{\infty}{}^{n}(1_B TP)^n 1_B (1-TP)f\right\}^x < \infty$.

Immers $\left|\left(\sum_{0}^{\infty}{}^{n}(1_B TP)^n 1_B f\right)^x\right| \leq (R|f|)^x < \infty$ en

$\left|\left(\sum_{0}^{\infty}{}^{n}(1_B TP)^n 1_B TPf\right)^x \leqq (R 1_B |T| P |f|)^x \leqq\right.$

$\leqq c\,\theta \sum_{0}^{\infty}{}^{n} \theta^n \{(1_B T_0 P)^n 1_B T_0 P f_0\}^x \leqq c\,\theta \sum_{0}^{\infty}{}^{n} \theta^n \{(1_B T_0 P)^n f_0\}^x$ volgens $(\theta)$ $\leqq$

$\leqq c\,\theta (1-\theta)^{-1} (1_B f_0)^x$ volgens het voorgaande.

Als de tweede gelijkheid van (7) nu waar is, zal het derde lid van (10) nul zijn, dus ook het eerste lid van (10), waaruit dus volgt, dat ook de eerste gelijkheid van (7) waar is.
In omgekeerde richting bewijst men (7) door uit te gaan van

(6.4.1) $(1-1_B TP)(1-C) = 1_B(1-TP)$.

(11) $1_B(1-TP)f = \{(1-1_B TP)(1-C)\}f = (1-1_B TP)(1-C)f$,

want

1. $(1-1_B TP)(1-C)$ bestaat.
2. $(1-C)f$ bestaat (reeds bewezen).
3. $\{(1-1_B TP)(1-C)\}f = 1_B(1-TP)f < \infty$.

Is dus de eerste gelijkheid van (7) waar, dan is het derde lid van (11) nul, dus ook het eerste lid van (11), waaruit volgt, dat ook de tweede gelijkheid van (7) waar is. Hiermee is de stelling volledig bewezen.

## 6.5 Walds fundamentele identiteit

Wij beschouwen een proces met onafhankelijke aangroeiingen (zie 5.6).
Wij nemen $\mathfrak{J} = \mathbb{R}$, de vz der reële getallen.

(1) $P_X^x = P_{X-x}^0 = \int p_{dv}\, 1_X^{x+v}$

waarbij:

(2) $\quad p_x \overset{\text{def}}{=} P^{o}_{x}$

Wij nemen verder $T^x = T$, constant, en beschouwen de integraalvergelijking:

(3) $\quad \int P^{x}_{dy} f^{y} = \frac{1}{T} f^{x}.$

Nu zal

(4) $\quad \int P^{x}_{dy} f^{y} = \int \left( \int p_{dv} I^{x+v}_{dy} \right) f^{y} = \int p_{dv} \int I^{x+v}_{dy} f^{y} = \int p_{dv} f^{x+v},$

wanneer tenminste de verwisseling der integratievolgorde is toegestaan. Wij zoeken een oplossing van de vorm

(5) $\quad f^{x} = e^{\xi x}$

dus

(6) $\quad \int p_{dv} e^{\xi(x+v)} = \frac{1}{T} e^{\xi x},$

of daar $e^{\xi x} \neq 0$ is:

(7) $\quad \int p_{dv} e^{\xi v} = \frac{1}{T}$ of $\varphi(\xi) = \frac{1}{T}$

met

(8) $\quad \varphi(\xi) \overset{\text{def}}{=} \int p_{dv} e^{\xi v},$

de karakteristieke functie van de verdelingsfunctie $p$.
De verwisseling van de integratievolgorde in (4) is zeker toegestaan als de integralen in (4) absoluut convergeren, dus als (8) absoluut convergeert, dus als $\varphi(\mathcal{Re}\,\xi)$ bestaat.
Wij hebben nu bewezen de

Stelling

Voor alle $\xi$, waarvoor $\varphi(\mathcal{Re}\,\xi)$ bestaat, is $f^{x} = e^{\xi x}$ een eigen functie van de matrix $P$ van het proces met onafhankelijke aangroeiingen bij de eigen waarde $\varphi(\xi)$.
Aan de tweede gelijkheid van (6.4.7) is nu voldaan (wanneer $B$ tenminste niet leeg is). De relatie (7) geldt uiteraard voor alle $x \in \mathfrak{J}$ (niet alleen $\in B$).
Om tot de eerste gelijkheid van (6.4.7) te besluiten, moeten wij nagaan of aan de voorwaarden van de stelling van 6.4 is voldaan ( $f^{x} = e^{\xi x}$ is niet begrensd als $\mathcal{Re}\,\xi \neq 0$ is).
Wij kiezen $f_0^{x} = e^{\xi_0 x}$, $\xi_0$ reëel, zodanig dat $Pf_0$, dus $\varphi(\xi_0)$ bestaat. Verder kiezen wij $T_0^{x} = T_0$, constant, reëel en $> 0$. Voorwaarde

(6.4.9) is vervuld voor $T_0 \leqq 1$. Aan (6.4.8) is voldaan als

(9) $\qquad T_0\,\varphi(\xi_0) \leqq 1\,.$

Wij kiezen $T$ en $\xi$ verder zodanig, dat voldaan is aan (7), dat $|T|\,T_0^{-1} = \theta < 1$ is en dat $|e^{\xi x}| \leqq c\,e^{\xi_0 x}$, dus dat $(\xi_0 - \mathcal{R}e\,\xi)\,x \geqq -\ln c$ is voor alle $x \in \mathcal{B}$.
Stellen we nu verder

(10) $\qquad C_{(A)} = \sum_0^\infty T^n\,{}_nC_{(A)}\,,$

zodat ${}_nC_{(A)}{}^x_X$ de kans op absorptie is in $A \cap X$, direct volgend op de n-de stap, dan is:

(11) $\qquad (f^{-1}Cf)^x = e^{-\xi x}\sum_0^\infty{}^n\, T^n \int {}_nC_{(A)}{}^x_{dy}\, e^{\xi y}.$

De eerste gelijkheid van (6.4.7) houdt nu in (als wij $x = o$ stellen):

(12) $\qquad \sum_1^\infty{}^n\, \varphi(\xi)^{-n}\int {}_nC_{(A)}{}^o_{dy}\, e^{\xi y} = 1\,.$

Deze gelijkheid staat bekend als Walds fundamentele identiteit. Wij hebben dus gezien, dat het een speciaal geval is van de gelijkheid $Cf = f$, die geldt voor alle $f$, die op $\mathcal{B}$ voldoen aan $f = TPf$, waarbij de voorwaarden van de stelling van 6.4 vervuld moeten zijn. Wald had deze gelijkheid nodig bij de ontwikkeling van de theorie der sequente analyse. Zie [58], blz. 159-160.

## 7. Lussen in de stochastische wandeling

### 7.1 Inleiding

In de organische chemie treden problemen op over de aard van molecuulstructuren, die in verband gebracht kunnen worden met de theorie van de stochastische wandeling. Het gaat hier om het statistisch chemisch onderzoek naar de opbouw van zeer grote en lange moleculen. Wij kunnen ons zo'n molecule namelijk ontstaan denken doordat eerst twee atomen ten opzichte van elkaar een vaste positie innemen. Hierna voegt een derde atoom zich er bij en wel zodanig, dat het aantal mogelijke posities ten opzichte van de twee reeds aanwezige (op rotaties na) eindig is. Het is bijvoorbeeld mogelijk, dat de verbindingslijn van het 2e en 3e atoom een vaste hoek moet maken met de ver-

bindingslijn van atoom 1 en 2. Men kan zich nu de mogelijke posities van alle vervolgens tot de keten toetredende atomen a priori als punten van een rooster gegeven denken. In dit rooster vormt zich nu dus een keten van atomen. Bij iedere stap is er nu een bepaalde wh, dat het bijkomende atoom zich in een bepaald roosterpunt plaatst. Wij krijgen zo een stochastische wandeling, die echter gecompliceerd wordt door het feit, dat terugkeer niet mogelijk is. Op één plaats kan immers slechts één atoom aanwezig zijn. Het is dus geen enkelvoudig, doch een hereditair Markof-proces. Via deze beschouwing komen wij echter op het probleem van het optreden van terugkeringen (lussen) bij een stochastische wandeling. Dit probleem zullen wij in de volgende paragrafen bestuderen. Dit is echter nog geenszins voldoende om de eigenschappen der processen met "verboden terugkeer" te leren kennen, een probleem dat nog bij lange na niet is opgelost.
Laat gegeven zijn een aftelbaar aantal whn $p_n$ , zodanig dat $\sum^n p_n = 1$ is. Wij beschouwen nu de voortbrengende functie volgens Laplace:

$$(1) \qquad C = \sum_0^{\infty}{}^{n} p_n T^n , \sum p_n = 1 .$$

$T^n$ is hierin de n-de macht van $T$ . Wij voeren verder de stochastische variabele $\underline{n}$ in, gedefinieerd door

$$(2) \qquad \mathcal{P}\{\underline{n} = n\} = p_n .$$

Nu geven de afgeleiden van $C$ naar $T$ voor $T=1$ de factoriële momenten van $\underline{n}$ :

$$(3) \qquad \left(\frac{dC}{dT}\right)_{T=1} = \mathcal{E}\,\underline{n} , \quad \left(\frac{d^kC}{dT^k}\right)_{T=1} = \mathcal{E}\,\underline{n}^{!k} .$$

Immers $\left(\frac{dC}{dT}\right)_{T=1} = \left[\sum_1^{\infty} n p_n T^{n-1}\right]_{T=1} = \sum_0^{\infty} n p_n = \mathcal{E}\,\underline{n}$

$$\left(\frac{d^2C}{dT^2}\right)_{T=1} = \left[\sum_2^{\infty} n^{!2} p_n T^{n-2}\right]_{T=1} = \sum_0^{\infty} n^{!2} p_n = \mathcal{E}\,\underline{n}^{!2} , \text{ enz.}$$

(Zie voor de factoriële momenten [15] , caput II, blz.81).
Voor de gewone momenten geldt:

$$(4) \qquad \mu_k = \mathcal{E}\,\underline{n}^k = \left[\left(T\frac{d}{dT}\right)^k C\right]_{T=1} .$$

Hebben wij een stelsel whn, dat van twee indices afhangt, $p_{m,n}$ dan voeren wij een voortbrengende fct $C$ , alsmede de stochastische variabelen $\underline{m}$ en $\underline{n}$ in door middel van:

(5) $$C = \sum\nolimits^{m,n} p_{m,n} T^m U^n , \quad p_{m,n} = \mathcal{P}\{\underline{m} = m \wedge \underline{n} = n\} .$$

Nu is

(6) $$\mu_{k,\ell} = \mathcal{E}\, \underline{m}^k \underline{n}^\ell = \left[\left(T\frac{\partial}{\partial T}\right)^k \left(U\frac{\partial}{\partial U}\right)^\ell C\right]_{T=U=1} .$$

Hangt $p$ van $k$ indices af, dan voeren wij $k$ stochastische variabelen $\underline{m}_1,\ldots\ldots,\underline{m}_k$ in, die wij opvatten als de $k$ componenten van een vector $\underline{m}$ :

(7) $$\underline{m} = (\underline{m}_1,\ldots\ldots,\underline{m}_k) .$$

Nu zal

(8) $$C \stackrel{def}{=} \sum p_{m_1,\ldots,m_k} (T^1)^{m_1} \ldots\ldots (T^k)^{m_k} ,$$

waarin $T^1,\ldots\ldots,T^k$ hulpvariabelen zijn en

(9) $$p_{m_1,\ldots,m_k} = \mathcal{P}\{\underline{m}_1 = m_1 \wedge \ldots\ldots \wedge \underline{m}_k = m_k\} .$$

Wij definiëren voorts

(10) $$\nabla_i \stackrel{def}{=} \frac{\partial}{\partial T_i}$$ (wordt uitgesproken als: nabla)

(11) $$\nabla\varphi \stackrel{def}{=} \left(\frac{\partial\varphi}{\partial T^1},\ldots\ldots,\frac{\partial\varphi}{\partial T^k}\right),$$

wanneer $\varphi(T^1,\ldots\ldots,T^k)$ een fct van $T^1,\ldots\ldots,T^k$ is. (Er worden hier geen exponenten bedoeld). $\nabla\varphi$ is dus een vector. Schrijven wij verder $T=1$ ter afkorting van $T^1 = T^2 = \ldots\ldots = T^k = 1$ , dan is

(12) $$\mathcal{E}\, \underline{m}_i = \left[\nabla_i C\right]_{T=1} ,$$

(13) $$\mathcal{E}\, \underline{m} = \left[\nabla C\right]_{T=1} .$$

Voorts geldt

(14) $$\left[\nabla_i \nabla_j C\right]_{T=1} = \mathcal{E}\, \underline{m}_i \underline{m}_j - \delta_{ij} \mathcal{E}\, \underline{m}_i = \begin{cases} \mathcal{E}\, \underline{m}_i^{!2} & als\ i=j . \\ \mathcal{E}\, \underline{m}_i \underline{m}_j & als\ i \neq j . \end{cases}$$

Definiëren wij tenslotte

(15) $\quad \Box \stackrel{\text{def}}{=} \sum^i \nabla_i \nabla_i = \sum^i \nabla_i^2,$

de gegeneraliseerde Laplace-operator, dan geldt

(16) $\quad \frac{1}{2}\left[\Box C\right]_{T=1} = \sum^i \frac{1}{2}\,\mathcal{E}\, m_i'^2 .$

## 7.2 Het lussenprobleem bij aftelbare Markof-processen

Wij beschouwen een Markof-proces (stochastische wandeling) over een aftelbare vz $\mathfrak{J}$ . Als uitgangspunt kiezen wij een punt $a \in \mathfrak{J}$. Voor de collectieve matrix $C_X^x$ van het absorptieproces leidden wij in 6.1 de betrekking (6.1.1) af. Wij nemen nu $A^x = A$, constant. De kans op absorptie is dus in ieder punt even groot. Een andere interpretatie, die op hetzelfde neerkomt, is, dat wij vóór elke stap loten, òf wij nog verder zullen gaan. De kans op het doen van een nieuwe stap is dan $B = 1 - A$.
De fct $T$ nemen wij niet constant. We nemen $U^x = T^x$, zodat de katastrophe C dus ook kan plaats vinden als het punt reeds geabsorbeerd is, als er dus geen stap meer zal worden gedaan. De wh, dat de katastrophe in het punt $x$ , waar het wandelende punt verblijft, plaats vindt, is dus steeds $1 - T^x$ . Het is duidelijk, dat $T^x$ $(x \in \mathfrak{J})$ een vector voorstelt met een aftelbaar oneindig aantal componenten.
In matrixvorm wordt (6.1.1) nu, gezien het bovenstaande:

(1) $\quad C = AT + BTPC,$ $A$ *en* $B$ *constanten.*

Wij merken verder op, dat $C_{\mathfrak{J}}^a(T)$ de wh is, dat het wandelende punt, uitgaande van $a$ , ooit in $\mathfrak{J}$ geabsorbeerd zal worden zonder dat $\mathcal{C}$ heeft plaats gevonden; dat wil zeggen, dat $C_{\mathfrak{J}}^a(T)$ de wh is, dat $\mathcal{C}$ nooit plaats vindt. Wij schrijven in plaats van $C_{\mathfrak{J}}^a(T)$: $C_{\mathfrak{J}}^a$ .
Door itereren vinden wij uit (1):

(2) $\quad C_{\mathfrak{J}}^a = \sum_0^\infty{}^n B^n(1-B)\left((TP)^n T\right)^a = \sum_0^\infty{}^n B^n(1-B)\left(T(PT)^n\right)^a .$

Het derde lid volgt uit het tweede vanwege de associativiteit der matrixvermenigvuldiging.
Wij krijgen dus een machtreeks in de oneindig dimensionale vector $T$:

(3) $$C_{\mathfrak{J}}^a = \sum_0^\infty{}^n B^n(1-B)\sum^{x_1,\dots,x_n} T^a P_{x_1}^a T^{x_1} P_{x_2}^{x_1} T^{x_2} \dots\dots P_{x_n}^{x_{n-1}} T^{x_n}$$
$$= \sum_0^\infty{}^n B^n(1-B)\sum^{x_1,\dots,x_n} T^a T^{x_1} \dots\dots T^{x_n}\left(P_{x_1}^a P_{x_2}^{x_1} \dots\dots P_{x_n}^{x_{n-1}}\right).$$

$P^a_{x_1} P^{x_1}_{x_2} \ldots\ldots\ldots P^{x_{n-1}}_{x_n} B^n(1-B)$ is de kans op het optreden van de bepaalde weg $(a, x_1, \ldots\ldots x_n)$.

Het aantal keren, dat het wandelende punt in $x$ terechtkomt, noemen wij $n_x$. Dan komt $T^x$ $n_x$ maal voor in $T^a T^{x_1} \ldots\ldots T^{x_n}$, zodat $C^a$ de vorm heeft:

$$(4) \qquad C^a_y = \sum p_{n_{x_1}, n_{x_2}, \ldots} (T^{x_1})^{n_{x_1}} (T^{x_2})^{n_{x_2}} (T^{x_3})^{n_{x_3}} \ldots\ldots,$$

dus $C^a_y$ heeft de vorm (7.1.8) met alleen $k = \infty$.

Dus, analoog aan (7.1.12)

$$(5) \qquad \left[\nabla_x C^a_y(T)\right]_{T=1} = \mathcal{E}\, \underline{n}_x .$$

Dus: de partiële afgeleide van $C^a_y$ naar $T^x$ is voor $T=1$ de verwachting van het aantal malen, dat het wandelende punt in het punt $x$ terecht komt.

Analoog aan (16) vinden wij verder:

$$(6) \qquad \tfrac{1}{2}\left[\Box\, C^a_y(T)\right]_{T=1} = \sum_{x \in \mathfrak{I}} \tfrac{1}{2}\, \mathcal{E}\, \underline{n}_x^{!2} = \mathcal{E} \sum_{x \in \mathfrak{I}} \tfrac{1}{2}\, \underline{n}_x^{!2} .$$

Als $\underline{n}_x = 0$ of $\underline{n}_x = 1$ is, dan is $\underline{n}_x^{!2} = 0$; als $\underline{n}_x = 2$ is, dus als $x$ een dubbelpunt van de weg is, dan is $\frac{1}{2}\underline{n}_x^{!2} = 1$. Dus geeft (6) de verwachting van het aantal dubbelpunten van de weg, waarbij een 3,4, algemeen $k$-voudig punt als 3,6, algemeen $\frac{1}{2}k(k-1)$ dubbelpunten geteld wordt, zoals in de algebraïsche meetkunde gebruikelijk is.

Noemen wij de stochastische variabele, die het aantal dubbelpunten, geteld zoals boven beschreven, aangeeft $\underline{D}$, dan volgt dus uit (6):

$$(7) \qquad \mathcal{E}\, \underline{D} = \left[\tfrac{1}{2}\, \Box\, C^a_y(T)\right]_{T=1} .$$

Wij hebben nu bewezen de

Stelling:

De verwachting van het aantal dubbelpunten of lussen in de stochastische wandeling is gelijk aan $\frac{1}{2}$ maal het resultaat van toepassing van de gegeneraliseerde operator van Laplace op $C^a_y$ bij de waarde $T=1$.

Daar wij bij de afleiding van (7) alleen van het feit gebruik maakten, dat $C^a_y$ een machtreeks is in $T^x$, maar niet van de speciale vorm (2), geldt de stelling ook voor willekeurige niet stationaire Markof-processen, waarbij $\mathcal{J}$ echter aftelbaar oneindig blijft.

## 7.3 Toepassing bij stationaire Markof-processen

Wanneer het Markof-proces stationair is, kunnen wij de linkerleden van (5), (6) en (7) van de vorige paragraaf eenvoudig berekenen.
Uit (7.2.3) volgt als wij $i$ in plaats van $x_i$ schrijven:

$$(1) \qquad \nabla_i C^a_{\mathcal{J}} = \sum_0^{\infty} {}^n B^n(1-B) \sum_0^n {}^h \left((TP)^h\right)^a_i \left((PT)^{n-h}\right)^i$$

Nogmaals differentiëren geeft:

$$(2) \qquad \nabla_j \nabla_i C^a_{\mathcal{J}} = \sum_1^{\infty} {}^n B^n(1-B) \sum_{f+g+h=n-1} {}^{f,g,h} \left((TP)^f\right)^a_j \left((PT)^g P\right)^j_i \left((PT)^h\right)^i +$$
$$+ \sum_1^{\infty} {}^n B^n(1-B) \sum_{f+g+h=n-1} {}^{f,g,h} \left((TP)^f\right)^a_i \left((PT)^g P\right)^i_j \left((PT)^h\right)^j .$$

Voor $i=j$ krijgen wij:

$$(3) \qquad \tfrac{1}{2} \nabla_i^2 C^a_{\mathcal{J}} = \sum_1^{\infty} {}^n B^n(1-B) \sum_{f+g+h=n-1} {}^{f,g,h} \left((TP)^f\right)^a_i \left((PT)^g P\right)^i_i \left((PT)^h\right)^i ,$$

zodat na sommatie over $i$ en nadat $T=1$ gesteld is, het resultaat is:

$$(4) \qquad \mathcal{E}\,\underline{D} = \sum_1^{\infty} {}^n B^n(1-B) \sum_{f+g+h=n-1} {}^{f,g,h} \sum {}^i (P^f)^a_i (P^{g+1})^i_i (P^h 1)^i_{\mathcal{J}} .$$

Nu is $P$ een Markof-matrix, dus $P1=1$, zodat

$$(5) \qquad \mathcal{E}\,\underline{D} = \sum_1^{\infty} {}^n B^n(1-B) \sum_{f+g \leqq n-1} {}^{f,g} \sum {}^i (P^f)^a_i (P^{g+1})^i_i ,$$

hetgeen ook direct in te zien is: $(P^{g+1})^i_i$ is de terugkeerwh in $g+1>0$ stappen, uitgaande van punt $i$.
De voorwaardelijke verwachting van $\underline{D}$ onder de voorwaarde, dat de weg een lengte n heeft, is gelijk aan:

$$(6) \qquad \mathcal{E}\{\underline{D} \mid n\} = \sum_{f+g \leqq n-1} {}^{f,g} \sum {}^i (P^f)^a_i (P^{g+1})^i_i .$$

Verwisselen wij in (5) de sommatie over n en die over $f$ en $g$ dan krijgen wij:

$$\mathcal{E}\,\underline{D} = \sum_0^{\infty} {}^{f,g} (1-B) \sum_{f+g+1}^{\infty} {}^n \sum {}^i \left((BP)^f\right)^a_i \left((BP)^{g+1}\right)^i_i B^{n-f-g-1} =$$
$$= \sum_0^{\infty} {}^{f,g} (1-B) \sum {}^i \left((BP)^f\right)^a_i \left((BP)^{g+1}\right)^i_i (1-B)^{-1} =$$

$$= \sum_{0}^{\infty} {}_{f,g} \sum^{i} \left( (BP)^{f} \right)_{i}^{a} \left( (BP)^{g+1} \right)_{i}^{i} .$$

Noemen wij

(7) $$\varphi_{y}^{x} \overset{def}{=} \left( \sum_{0}^{\infty} {}_{n} (BP)^{n} \right)_{y}^{x} = \left( (I - BP)^{-1} \right)_{y}^{x} ,$$

dan geldt:

(8) $$\mathcal{E}\,\underline{D} = \sum_{x \in \mathfrak{Y}} \varphi_{x}^{a} (\varphi_{x}^{x} - 1)$$

Wanneer wij te maken hebben met een proces met onafhankelijke aangroeiingen, kan (8) nog verder worden vereenvoudigd. In dat geval zal $\varphi_{x}^{x}$ immers onafhankelijk van $x$ en $= \varphi_{a}^{a}$ of ook $\varphi_{0}^{0}$ zijn, zodat in (8) alleen de $\varphi_{x}^{a}$ gesommeerd moet worden over $x$ . Daar $\sum_{x \in \mathfrak{Y}} \varphi_{x}^{a} = \varphi_{\mathfrak{Y}}^{a} = (1 - B)^{-1}$ is, krijgen wij dus:

(9) $$\mathcal{E}\,\underline{D} = (1 - B)^{-1} (\varphi_{0}^{0} - 1) .$$

Daar voorts

(10) $$\mathcal{E}\,\underline{D} = \sum_{0}^{\infty} {}_{n} B^{n} (1 - B)\, \mathcal{E} \{ \underline{D} | n \}$$

zal $\mathcal{E} \{ \underline{D} | n \}$ dus gelijk zijn aan de coëfficiënt van $B^{n}$ in $(1-B)^{-2}(\varphi_{0}^{0}-1)$.

Nu is $(1 - B)^{-2} (\varphi_{0}^{0} - 1) = \sum_{1}^{\infty} k B^{k-1} \sum_{1}^{\infty} {}_{\ell} B^{\ell} (P^{\ell})_{0}^{0} = (k + \ell - 1 = n) \sum_{1}^{\infty} {}_{n} B^{n} \sum_{1}^{n} {}_{\ell} (n - \ell + 1) (P^{\ell})_{0}^{0} ,$

dus

(11) $$\mathcal{E} \{ \underline{D} | n \} = \sum_{1}^{n} {}_{\ell} (n - \ell + 1) (P^{\ell})_{0}^{0} ,$$

welke formule ook direct uit (6) is af te leiden.
Zoals zoveel eenvoudige toepassingen van algemene stellingen is dit resultaat triviaal: Immers $(P^{\ell})_{0}^{0}$ is de wh, dat een weg van de lengte $\ell$ gesloten zal zijn; dat wil zeggen, dat zijn eerste en laatste punt samenvallen. Heeft de weg de lengte n, dan kan een lus de grootte $\ell$ hebben, waarbij $1 \leqq \ell \leqq n$ is. Met een "lus van de lengte 1" bedoelen wij dan een "stap", waarbij het springende punt niet springt, doch op dezelfde plaats blijft. Is dit zo, dan kan de lus beginnen bij het 1e, 2e,...,$(n-\ell+1)^{e}$ punt. Daar het een proces met onafhankelijke aangroeiingen betreft, hebben deze $n-\ell+1$ mogelijke gevallen gelijke whn om het beginpunt van een lus ter grootte $\ell$ te zijn, namelijk $(P^{\ell})_{0}^{0}$.
Hieruit volgt (11) direct.

## 7.4 Differentialen van functionellen

Hebben wij een niet aftelbare vz $\mathcal{T}$ en continue verdelingsfcts $\mathcal{P}$ , dan wordt het lussenprobleem triviaal: De terugkeerwh is dan immers nul. Wij kunnen echter verwante problemen beschouwen, bijvoorbeeld de vraag, hoe dikwijls een wandelend punt binnen een bepaalde afstand zal komen van een van zijn vorige verblijfplaatsen. Dit is slechts mogelijk in metrische ruimten.[1) Algemeen kunnen wij terugkeer definiëren door het nul worden van een fct van uitgangspunt en wandelend punt (waarbij deze fct dus niet aan de afstandsrelaties behoeft te voldoen). Alvorens dit probleem te behandelen willen wij eerst enkele opmerkingen maken over differentialen. Zie voor het onderstaande ook [12] . In de gewone analyse is het mogelijk het begrip differentiaal in te voeren, onafhankelijk van het begrip differentiaalquotiënt, namelijk op de volgende wijze:

(1) $$df(x) \overset{def}{=} \lim_{\varepsilon \downarrow 0} \frac{f(x+\varepsilon dx) - f(x)}{\varepsilon},$$

als de limiet bestaat. Hierin stelt $dx$ een eindige aangroeiing van $x$ voor.
Het differentiaalquotiënt $f'(x)$ wordt dan gedefiniëerd met behulp van

(2) $$df(x) = f'(x)\,dx,$$

dus

(3) $$f'(x) = \left[df(x)\right]_{dx=1}$$

In meer dimensies krijgen wij:

(4) $$df(\vec{x}) = \lim_{\varepsilon \downarrow 0} \frac{f(\vec{x}+\varepsilon d\vec{x}) - f(\vec{x})}{\varepsilon} = (\nabla f)\, d\vec{x}.$$

De limiet bestaat onder zekere regulariteitseisen en definiëert de gradiënt $\nabla f$.

-------------

1) Een metrische ruimte is een ruimte, waarin een "afstandsbegrip" gedefiniëerd is; dat wil zeggen, dat aan elk tweetal tot de ruimte behorende punten $x$ en $y$ is toegevoegd een getal $\rho^{xy}$ met de volgende eigenschappen:

1. $\rho^{xy} \geqq 0$
2. $\rho^{xx} = 0$
3. $\rho^{xy} = \rho^{yx}$
4. $\rho^{xy} + \rho^{yz} \geqq \rho^{xz}$

Men kan in een metrische ruimte $M$ een topologische structuur definiëren door de vzn, welke voortgebracht worden door de vzn. $S(x;\varepsilon) \overset{def}{=} \mathcal{E}ns\{y \in M | \rho^{xy} < \varepsilon\}, x \in M, \varepsilon > 0$ , open te verklaren. De vzn $S(x;\varepsilon)$ worden sferen of $\varepsilon$-omgevingen van $x$ genoemd.

Uit (4) volgt nu:

$$(5) \qquad \nabla_i f = \lim_{\varepsilon \downarrow 0} \frac{f(\vec{x}+\varepsilon\vec{e}_i)-f(\vec{x})}{\varepsilon} =$$

waarbij $\vec{e}_i$ dus de vector is, die op de $i$-de plaats een 1 heeft, en verder uit nullen bestaat.
Voor een functionel van de eerste soort definiëren wij, analoog aan (1):

$$(6) \qquad d\,C(T) \overset{def}{=} \lim_{\varepsilon \downarrow 0} \frac{C(T^x+\varepsilon dT^x)-C(T^x)}{\varepsilon}$$

als de limiet bestaat.
Analoog aan (5) krijgen wij dan

$$(7) \qquad \left(\frac{\partial}{\partial T}\right)_x C = \nabla_x C = \lim_{\varepsilon \downarrow 0} \frac{C(T+\varepsilon I_x)-C(T)}{\varepsilon}.$$

Onder zekere regulariteitsvoorwaarden, die bijvoorbeeld vervuld zijn als de functionel $C$ analytisch is, is $\nabla_x C$ aftelbaar additief.
Is $\varphi$ een functionel van de tweede soort (fct van de fct $F_x$ van de tweede soort), dan definiëren wij:

$$(8) \qquad \left(\frac{\partial}{\partial F}\right)^x \varphi \overset{def}{=} \lim_{\varepsilon \downarrow 0} \frac{\varphi(F+\varepsilon I^x)-\varphi(F)}{\varepsilon},$$

terwijl als $\psi$ een functionel van tweeërlei soort is (fct van de matrix $M_x^x$ ), wij kunnen definiëren:

$$(9) \qquad \left(\frac{\partial}{\partial M}\right)_A^a \psi \overset{def}{=} \lim_{\varepsilon \downarrow 0} \frac{\psi(M+\varepsilon I_A I^a)-\psi(M)}{\varepsilon}$$

Hierin is $I_A I^a$ een matrix: $(I_A I^a)_x^x = I_A^x I_x^a$ . Dit moet niet verward worden met $I^a I_A = \int I_{dx}^a I_A^x = I_A^a$ , zoals ook algemeen de matrix $fF$ met kengetallen $f^x F_x$ verschillend is van het product $Ff$, d.i. $\int F_{dx} f^x$.
Van al deze definities zullen wij alleen (7) in de volgende paragrafen nodig hebben.

## 7.5 Generalisatie van het lussenprobleem bij niet aftelbare Markof-processen

Wij beschouwen een algemeen in de tijd discreet stochastisch proces, waarbij dus gedefiniëerd is de wh.

$$P_{(n)X_0,\ldots,X_n} = \mathcal{P}\{\underline{x}_0 \in X_0 \wedge \underline{x}_1 \in X_1 \wedge \ldots\ldots \wedge \underline{x}_n \in X_n \mid n\}.$$

Het collectieve kenmerk van het algemene stochastische proces is:

(1) $$C=(1-B)\sum_{0}^{\infty n}B^n\iint\cdots\int P_{(n)dx_0,\ldots,dx_{n-1}}T^{x_0}\ldots T^{x_{n-1}},$$

waaruit (7.2.2) weer terug te krijgen is door de eisen (5.6.5), (5.6.6) en $P_{(o)x}=I_x^o$ te stellen.
Differentiatie geeft:

(2) $$\nabla_x C=(1-B)\sum_{0}^{\infty n}B^n\iint\cdots\int P_{(n)dx_0,\ldots,dx_{n-1}}\sum_{0}^{n}{}^{k}T^{x_0}\ldots T^{x_{k-1}}I_x^{x_k}T^{x_{k+1}}\ldots T^{x_n};$$

hetgeen volgt uit de productregel voor het differentiëren, wanneer wij bedenken, dat

$$\nabla_x T=\lim_{\varepsilon\downarrow 0}\frac{(T+\varepsilon I_x)-T}{\varepsilon}=I_x.$$

Noemen wij

(3) $$n_x \overset{def}{=} \sum_{0}^{n}{}^{k} I_x^{x_k},$$

dat is dus het aantal malen, dat het wandelende punt in X is, dan volgt uit (2):

(4) $$\left[\nabla_x C\right]_{T=1}=\mathcal{E}\,\underline{n}_x.$$

Als $X=\mathfrak{J}$ is, zal $\underline{n}_x=\underline{n}$ zijn, het aantal stappen, dat het wandelende punt gedaan heeft voordat het proces eindigt. (4) is een generalisatie van (7.2.5).
De voorwaardelijke verwachting is:

(5) $$\mathcal{E}\{\underline{n}_x|n\}=\int\cdots\int P_{(n)dx_0,\ldots,dx_n}\sum_{0}^{n}{}^{k}I_x^{x_k}=\sum_{0}^{n}{}^{k}P_{(n)\underbrace{\underbrace{\mathfrak{J},\ldots,\mathfrak{J}}_{k},X,\mathfrak{J},\ldots,\mathfrak{J}}_{n+1}}$$

Nogmaals differentiëren geeft als generalisatie van (7.1.14):

(6) $$(\nabla_y\nabla_x C)_{T=1}=\mathcal{E}\{\underline{n}_x\underline{n}_y-\underline{n}_{x\cap y})$$

Immers

$$(\nabla_y\nabla_x C)_{T=1}=(1-B)\sum_{0}^{\infty n}B^n\int\cdots\int P_{(n)dx_0,\ldots,dx_n}\sum_{k\neq \ell}^{k,\ell}I_x^{x_k}I_y^{x_\ell}=$$
$$=\mathcal{E}\sum_{k\neq\ell}^{k,\ell}I_x^{x_k}I_y^{x_\ell}=\mathcal{E}(\underline{n}_x n_y-\underline{n}_{x\cap y}).$$

Thans gaan wij over tot een generalisatie van het lussenprobleem. Daar bij continue verdelingen de wh op een lus nul is, beschouwen wij algemener "quasi-lussen" door invoering van een tweetoestandsfunctie $f^{xy}$. Bijvoorbeeld: $f^{xy}=\rho^{xy}$, de afstand der punten $x$ en $y$ in de metrisch verondersteld ruimte $\mathfrak{J}$, of $f^{xy}=1$

als $\rho^{xy} \leqq a$ en $=0$ als $\rho^{xy} > a$ is. De fct $f$ moet zowel ten opzichte van $x$ als ten opzichte van $y$ van de eerste soort zijn.
We definiëren:

(7) $$\square_{(f)} \overset{\text{def}}{=} \iint f^{xy} \nabla_{dx} \nabla_{dy} .$$

Is $\mathfrak{J}$ aftelbaar, dan wordt dit $\square_{(f)} = \sum^{ij} f^{ij} \nabla_i \nabla_j$ , hetgeen in (7.1.15) overgaat als $f^{ij} = \delta_{ij}$.
De operator heeft constante coëfficiënten daar $f^{ij}$ niet afhangt van $T$.
Onder zekere regulariteitsvoorwaarden geldt, dat

$$\iint f^{xy} \nabla_{dx} \nabla_{dy} = \iint f^{yx} \nabla_{dy} \nabla_{dx}$$ , zodat (7) identiek 0 is als $f$ antisymmetrisch is ($f^{xy} = -f^{yx}$). Wij kunnen ons daarom beperken tot symmetrische fcts $f$ ($f^{xy} = f^{yx}$).

Mits de verwisseling der integratievolgorde is toegestaan, volgt uit (6):

$$\tfrac{1}{2}\left[\square_{(f)} C\right]_{T=1} = \tfrac{1}{2} \iint \mathcal{E}\left(\underline{n}_{dx}\underline{n}_{dy} - \underline{n}_{dx \cap dy}\right) f^{xy} = \tfrac{1}{2} \iint \mathcal{E}\, \underline{n}_{dx}\underline{n}_{dy} f^{xy} - \tfrac{1}{2} \int \mathcal{E}\, \underline{n}_{dx} f^{xx},$$

dus

(8) $$\tfrac{1}{2}\left[\square_{(f)} C\right]_{T=1} = \tfrac{1}{2} \mathcal{E} \iint \underline{n}_{dx}\underline{n}_{dy} f^{xy} - \tfrac{1}{2} \mathcal{E} \int \underline{n}_{dx} f^{xx}.$$

Uit (3) volgt:

(9) $$\int \underline{n}_{dx} f^{xx} = \sum_{0}^{n}{}^{k} \int l_{dx}^{\underline{x}_k} f^{xx} = \sum_{0}^{n}{}^{k} f^{\underline{x}_k \underline{x}_k}$$

en

(10) $$\iint \underline{n}_{dx}\underline{n}_{dy} f^{xy} = \sum_{0}^{n}{}^{k,l} \iint l_{dx}^{\underline{x}_k}\, l_{dy}^{\underline{x}_l} f^{xy} = \sum_{0}^{n}{}^{k,l} f^{\underline{x}_k , \underline{x}_l}.$$

Nu kunnen onder $\underline{x}_0, \dots\dots, \underline{x}_n$ gelijken voorkomen. Stel het aantal verschillende is $r$: $\underline{x}_{(1)}, \dots\dots, \underline{x}_{(r)}$. Het aantal malen, dat het wandelende punt $\underline{x}_{(i)}$ passeert, noemen wij $\underline{n}_i$ .

Nu gaan (9) en (10) over in:

(11) $$\int \underline{n}_{dx} f^{xx} = \sum_{1}^{r}{}^{i} \underline{n}_i f^{\underline{x}_{(i)} \underline{x}_{(i)}}$$

en

(12) $$\iint \underline{n}_{dx}\underline{n}_{dy} f^{xy} = \sum_{1}^{r}{}^{i,j} \underline{n}_i \underline{n}_j f^{\underline{x}_{(i)} x_{(j)}} .$$

Wij definiëren voorts:

(13) $$\underline{S}_{(f)} \overset{\text{def}}{=} \frac{1}{2}\sum_{i\neq j}^{i,j} \underline{n}_i \underline{n}_j f^{\underline{x}(i)\underline{x}(j)} + \frac{1}{2}\sum^{i} \underline{n}_i(\underline{n}_i - 1) f^{\underline{x}(i)\underline{x}(i)}.$$

Deze grootheid is gelijk aan de som van de $f^{\underline{x}(i)\underline{x}(j)}$ $(i,j \in \{1,\dots,n\})$ over alle paren $i \neq j$, vermeerderd met de som van de $f^{\underline{x}_i \underline{x}_i}$ over alle meervoudige punten, geteld als meervoudige dubbelpunten zoals beschreven in 7.2.

Uit (8) volgt nu:

(14) $$\mathcal{E}\,\underline{S}_{(f)} = \frac{1}{2}\left[\Box_{(f)} C\right]_{T=1} = \frac{1}{2}\,\mathcal{E}\sum_{i\neq j}^{i,j} \underline{n}_i \underline{n}_j f^{\underline{x}(i)\underline{x}(j)} + \frac{1}{2}\,\mathcal{E}\sum^{i} \underline{n}_i(\underline{n}_i - 1) f^{\underline{x}(i)\underline{x}(j)}.$$

De tweede term wordt nul:

1° als $\forall_x^{\mho} f^{xx} = 0$, bijvoorbeeld als $f^{xy} = \rho^{xy}$, de afstand van $x$ en $y$ is,

2° $\forall_i\ \underline{n}_i = 0$ of $\underline{n}_i = 1$ (spr o). 1)

Het laatste geval treedt steeds op als de overgangswhn $P_{(n)x_0,\dots,x_n}$ continue verdelingen hebben. In dat geval volgt uit (13):

(15) $$\underline{S}_{(f)} = \frac{1}{2}\sum_{i\neq j}^{i,j} f^{\underline{x}(i)x(j)},$$

waarin $\underline{x}_{(i)}$ de $\underline{n}$ punten zijn, die het wandelende punt bereikt. Hiervoor kunnen wij ook schrijven:

(16) $$\underline{S}_{(f)} = \frac{1}{2}\underline{n}(\underline{n}-1)\,\mathop{\mathcal{M}}_{i\neq j} f^{\underline{x}(i)\underline{x}(j)},$$

als $\mathop{\mathcal{M}}_{i\neq j} f^{\underline{x}(i)\underline{x}(j)}$ het gemiddelde voorstelt der waarden $f^{\underline{x}(i)\underline{x}(j)}$ genomen voor alle $i \neq j$.

Wij kunnen $f^{xy}$ de "quasiafstand" der punten $x$ en $y$ noemen. Als $f^{xy} = \rho^{xy}$ krijgen wij de gewone afstand.

Gaan wij daarentegen uit van een relatie $R(x,y)$, die voor alle paren $(x,y)$ òf wel òf niet geldt (bijvoorbeeld $R(x,y) \overset{\text{def}}{=} (\rho^{xy} \leqq a), a \geqq 0$) en nemen wij voor $f$ de karakteristieke fct van deze relatie, dus

$$f^{xy} = \begin{cases} 1 & \text{als } R(x,y) \text{ geldt,} \\ 0 & \text{als } R(x,y) \text{ niet geldt,} \end{cases}$$

-------------

1) Spr (salva probabilitate) betekent: behoudens een wh.

dan wordt volgens (15) $\underline{S}_{(f)}$ gelijk aan een half maal het aantal geordende paren van verschillende punten van de weg, waarvoor $R(x,y)$ geldt. Bij een symmetrische relatie is dit het aantal geordende paren.

## 7.6 Toepassing bij stationaire Markof-processen

Als het stochastisch proces een stationair enkelvoudig Markof-proces is, kunnen wij het linkerlid van (7.5.8) weer eenvoudig berekenen. Wij nemen weer aan, dat $f^{xy}$ symmetrisch is en krijgen:

(1) $$\tfrac{1}{2}\square_{(f)}C = \sum_{0}^{\infty} {}^{n}\, B^n(1-B) \sum_{k+\ell+m=n-1} {}^{k,\ell,m} \iint ((TP)^k)^{\circ}_{dx}((PT)^{\ell}P)^{x}_{dy}((PT)^m)^y f^{xy},$$

hetgeen voor $T^x = 1$, daar dan $(PT)^y = 1$ is, wordt:

(2) $$\left[\tfrac{1}{2}\square_{(f)}C\right]_{T=1} = \sum_{0}^{\infty} {}^{n}\, B^n(1-B) \sum_{k+\ell\leq n-1} {}^{k,\ell} \iint (P^k)^{\circ}_{dx}(P^{\ell+1})^{x}_{dy} f^{x,y}.$$

Is er sprake van een proces met onafhankelijke aangroeiingen en is bovendien $f^{x,y} = f^{0,y-x} = f^{y-x}$, dan volgt uit (5.6.10):

$$\iint (P^k)^{\circ}_{dx}(P^{\ell+1})^{x}_{dy} f^{x,y} = \int p^{(k)}_{dz} \int p^{(\ell+1)}_{dv} \int l^{z}_{dx} \int l^{x+v}_{dy} f^{y-x} = \int p^{(\ell+1)}_{dv} f^{v},$$

waarbij:

(3) $$p^{(k)}_{X} \overset{\mathrm{def}}{=} (P^k)^{\circ}_{X}.$$

Immers $\int l^{x+v}_{dy} f^{y-x} = f^{v}$, $\int l^{z}_{dx} = 1$, $\int p^{(k)}_{dz} = 1$.

Het rechterlid van (2) wordt dus:

$$\sum_{0}^{\infty} {}^{n}\, B^n(1-B) \sum_{0}^{n-1} {}^{\ell} \sum_{0}^{n-\ell-1} {}^{k} \int p^{(\ell+1)}_{dv} f^{v} = \sum_{0}^{\infty} {}^{n}\, B^n(1-B) \sum_{1}^{n} {}^{\ell} (n-\ell+1) \int p^{(\ell)}_{dv} f^{v},$$

dus

(4) $$\left[\tfrac{1}{2}\square_{(f)}C\right]_{T=1} = \sum_{0}^{\infty} {}^{n}\, B^n(1-B) \sum_{1}^{n} {}^{\ell} (n-\ell+1) \int p^{(\ell)}_{dv} f^{v}.$$

Ook deze relatie kan direct ingezien worden.
Wanneer $f^{v}$ een Fourier-getransformeerde heeft:

(5) $$f^{v} = \int e^{ivt} X(t)\, dt,$$ dan zal als

(6) $$\varphi(t) \overset{\mathrm{def}}{=} \int p_{dv}\, e^{ivt},$$ dus $\int p^{(\ell)}_{dv}\, e^{ivt} = (\varphi(t))^{\ell}$ gelden:

$$\left[\tfrac{1}{2}\square_{(f)}C\right]_{T=1} = \sum_{0}^{\infty} {}^{n}\, B^n(1-B) \sum_{1}^{n} {}^{\ell} (n-\ell+1) \int (\varphi(t))^{\ell} X(t)\, dt =$$

$$= \sum_{0}^{\infty} {}^{n}\, B^n(1-B) \int \frac{n-(n+1)\varphi(t)+\varphi(t)^{n+1}}{(1-\varphi(t))^2}\, \varphi(t) X(t)\, dt.$$

## Deel III. Appendix

### 8. Eigenschappen van matrices en vectoren:

Wij zullen thans enige eigenschappen van vierkante matrices bewijzen, die in de voorgaande hoofdstukken genoemd en gebruikt zijn.

Wij zullen zonder dit steeds expliciet te vermelden in dit hoofdstuk vierkante matrices beschouwen met $\tau$ rijen en kolommen, waarvan de elementen complexe getallen zijn. Zie hiervoor ook [11], [52] en [60], waar ook de bewijzen te vinden zijn van stellingen, die wij slechts noemen.

#### 8.1 Matrixring

Stellen $M$ en $N$ matrices voor, dan definiëren we:

(1) $$(M \pm N)^i_j \overset{\text{def}}{=} M^i_j \pm N^i_j$$

(2) $$(MN)^i_j \overset{\text{def}}{=} M^i_k N^k_j$$

(3) $$(cM)^i_j \overset{\text{def}}{=} cM^i_j \quad \text{als } c \in R.$$

De matrixoptelling is associatief en commutatief; de matrixvermenigvuldiging is wel associatief, doch niet commutatief. Verder geldt de distributieve eigenschap van de vermenigvuldiging ten opzichte van de optelling.

De nulmatrix is de matrix, geheel bestaande uit elementen $o$. Deze geven we aan met $O$. Dus geldt $M + O = O + M = M$.

De eenheidsmatrix is de matrix, die op de hoofddiagonaal elementen 1 heeft staan, en verder uit nullen bestaat. Noemen we deze $I$ (iota), dus:

$$I^i_j = \begin{cases} 1 & \text{als } i = j \\ 0 & \text{als } i \neq j \end{cases} \quad , \text{ dan geldt } MI = IM = M.$$

De eenheidsmatrix is dus met elke andere matrix verwisselbaar. Uit (2) volgt, dat ook $M^n$ gedefinieërd is. Zo is $(M^2)^i_j = (MM)^i_j = M^i_k M^k_j$.

Voor een bepaalde $i$ en $j \in \{1, \ldots\ldots, \tau\}$ beschouwen we de $\tau$-1 bij $\tau$-1 matrix, die ontstaat door van de oorspronkelijke matrix de $i$-de rij en de $j$-de kolom weg te laten. Van de nieuwe matrix berekenen we de determinantwaarde. Onder de minor van het element $M^i_j$ van de oorspronkelijke matrix $M$ verstaan we nu deze determinantwaarde, voorzien van een + teken als $i+j$ even en van een − teken als $i+j$ oneven is. We beschouwen nu de matrix, die ontstaat, als we de elementen van $M$ door hun minoren vervangen en de zo verkregen matrix ten opzichte van zijn hoofddiagonaal spiegelen. Deze matrix heet de geadjungeerde van $M$ . We geven hem aan met $M_a$ . Nu geldt:

(4) $MM_a = M_aM = I.|M|$ als $|M|$ de determinantwaarde van $M$ voorstelt. Als $|M| \neq 0$ kunnen we dus een inverse van $M$ definiëren.

(5) $M^{-1} = \frac{M_a}{|M|}$ .

Er geldt nu:

(6) $M^{-1}M = MM^{-1} = I$.

Matrices met determinantwaarde $0$ hebben dus geen inverse.

Ook zijn er nuldelers. Uit $MN = 0$ hoeft dus <u>niet</u> te volgen $M = 0$ of $N = 0$ . Neem bijvoorbeeld $r = 2$ , dan geldt $\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$.

Eigenschap (3) definiëert een externe vermenigvuldiging.
Uit het voorgaande volgt nu, dat de vierkante matrices met $r$ rijen en kolommen een ring vormen met het lichaam der reële getallen als operatoren (als operatoren kunnen we natuurlijk evengoed de complexe getallen of de gehele getallen, enz. nemen).
Voor de determinantwaarden gelden nog de eigenschappen:

1. Een <u>ring</u> $R$ is een stelsel elementen $a, b, c, \ldots\ldots$, waarbij aan ieder tweetal elementen $a$ en $b$ eenduidig is toegevoegd een som $a+b$ en een product $ab$ met de volgende eigenschappen:

1. $a+(b+c)=(a+b)+c$ , associatieve eigenschap van de optelling
2. $a+b = b+a$ , commutatieve eigenschap van de optelling
3. $\forall_a^R \forall_b^R \exists_x^R \; a+x=b$
4. $a(bc)=(ab)c$ , associativiteit van de vermenigvuldiging
5. $\left.\begin{array}{l} a(b+c)=ab+ac \\ (b+c)a=ba+ca \end{array}\right\}$ distributieve eigenschap van de vermenigvuldiging t.o.v. de optelling.

Uit 3 volgt de existentie van een (en slechts een) nulelement, terwijl uit 3 ook volgt, dat ieder element een tegengestelde bezit (bij elke $a$ bestaat een element $-a$ ,zodanig, dat $a+(-a)=0$ ).
De vermenigvuldiging behoeft niet commutatief te zijn. Is dit wel zo, dan spreekt men van een <u>commutatieve ring</u>. Ook wordt de oplosbaarheid van de vergelijking $ax=b$ niet geëist.
Een ring kan nuldelers hebben; het is dus mogelijk, dat $ab=0$ is, maar dat toch $a \neq 0$ en $b \neq 0$ is.
Voorbeelden van ringen: alle gehele getallen, alle even getallen, alle polynomen met reële coefficienten, alle polynomen met gehele coefficienten.

Er kan ook nog een externe vermenigvuldiging gedefiniëerd zijn tussen de ringelementen en de elementen van een andere ring $\Omega$: $\alpha, \beta, \gamma, \ldots\ldots$, operatoren genaamd. Dat wil zeggen, dat $\alpha a$ gedefiniëerd is en weer een element van $R$ is. De volgende twee voorwaarden moeten bovendien vervuld zijn:

$$\alpha(a+b) = \alpha a + \alpha b,$$
$$(\alpha+\beta)a = \alpha a + \beta a.$$

We spreken van een ring met operatoren. B.v.: de ring van alle even getallen bezit de gehele getallen als operatoren.
Geldt bovendien:

6. $\quad \forall_{a\neq 0} \forall_b \exists_x \; ax = b,$

dan krijgen we een lichaam. Meestal beschouwt men commutatieve lichamen, zodat dan ook geldt:

7. $\quad ab = ba.$

Uit 6 volgt, dat een lichaam ook een éénelement bezit, terwijl ieder element $\neq 0$ een inverse bezit $(aa^{-1} = 1)$. Nuldelers kunnen in een lichaam niet voorkomen, want stel $ab=0$, dan is als $a \neq 0$:

$$a^{-1}(ab) = (a^{-1}a)b = 1.b = b, \text{ dus } b = 0.$$

1. Een ideaal $\mathfrak{a}$ in een ring $R$ is een deelverzameling van $R$, waarvoor geldt:

1. Als $a \in \mathfrak{a}$ en $b \in \mathfrak{a}$, dan $a-b \in \mathfrak{a}$.

2. Als $a \in \mathfrak{a}$ en $r \in R$, dan $ar \in \mathfrak{a}$.

Dit is een linksideaal. Bij een rechtsideaal wordt geëist:

2'. Als $a \in \mathfrak{a}$ en $r \in R$, dan $ra \in \mathfrak{a}$.

Geldt zowel 2 als 2', dan spreekt men van een tweezijdig ideaal. Bij commutatieve ringen vallen de drie begrippen samen. Uit 1,2, resp. 1,2' is duidelijk, dat elk ideaal tevens een ring is.
Een hoofdideaal is een ideaal, dat voortgebracht wordt door een element van de ring, b.v. door $a \in R$, d.w.z. dat bestaat uit alle elementen van de vorm $ra + na$ ($r \in R$, $n$ een geheel getal).
Ringen, waarbij ieder ideaal hoofdideaal is, worden Euclidisch genoemd. B.v. de ring van alle gehele getallen, van alle veeltermen met coefficienten uit een lichaam. Euclidische ringen bezitten steeds een éénelement. Men kan bewijzen, dat, als $(a,b)=1$, er elementen $p \in R$ en $q \in R$ gevonden kunnen worden, zodanig, dat $ap + bq = 1$.

(7) $\quad |MN| = |M|.|N|$, waaruit volgt $|M^n| = |M|^n$.

(8) $|M^{-1}| = |M|^{-1}$

(9) $|cM| = c^{r}|M|$.

Uit de eigenschappen (1),(2) en (3) volgt verder, dat een willekeurige veelterm in $M$ met coëfficiënten uit de operatorenring weer een matrix is. Hierbij definiëren we $M^{o} = I$.

Bij iedere veelterm $f(x) = \sum_{o}^{m} c_i x^{m-i}$ met $c_i \in R$ kunnen we dus definiëren een matrixveelterm $f(M) = \sum_{o}^{m} c_i M^{m-i}$ , waarbij $f(M)$ weer een matrix is.

Zijn $f(M)$ en $g(M)$ twee zulke polynomen, dan geldt:

(10) $f(M)\,g(M) = g(M)f(M)$.

Alle polynomen in $M$ vormen uiteraard weer een ring (aequivalentie met veeltermring). Bij deze ring is de vermenigvuldiging volgens (10) wel commutatief. Deze ring is dus een commutatieve deelring van de ring van alle matrices en bezit dezelfde operatoren als de oorspronkelijke ring.

Een andere commutatieve deelring krijgen we door alle diagonaalmatrices te beschouwen, dat zijn die matrices, die alleen elementen $\neq 0$ hebben op de hoofddiagonaal.

Duidelijk is, dat sommen, verschillen en producten van diagonaalmatrices weer diagonaalmatrices zijn. De vermenigvuldiging in deze deelring is commutatief.

Immers stel $A = \begin{pmatrix} a_1 & & o \\ & \ddots & \\ o & & a_n \end{pmatrix}$ en $B = \begin{pmatrix} b_1 & & o \\ & \ddots & \\ o & & b_n \end{pmatrix}$, dan is :

$$AB = \begin{pmatrix} a_1b_1 & & o \\ & \ddots & \\ o & & a_nb_n \end{pmatrix} = BA.$$

## 8.2 Eigen waarden en eigen vectoren van matrices

We beschouwen de volgende $r$ vergelijkingen in de $r$ componenten $f_1, \ldots, f_r$ van een vector $f$ :

(1) $\forall_i^{r}\, M_j^i f^j = \lambda f^i$ of in matrixnotatie: $Mf = \lambda f$.

De determinant van dit stelsel lineaire vergelijkingen in $f^i$ is:

(2) $$D(\lambda) \overset{def}{=} \begin{vmatrix} M_1^1-\lambda & M_2^1 & \cdots & M_r^1 \\ M_2^2 & M_2^2-\lambda & \cdots & M_r^2 \\ \vdots & \vdots & \ddots & \vdots \\ M_1^r & M_2^r & \cdots & M_r^r-\lambda \end{vmatrix} = |M - \lambda I|.$$

$D(\lambda)$ is een veelterm van de graad $r$ in $\lambda$ en heet de karakteristieke veelterm. Wanneer we de som van de hoofdminoren van de orde $k$ $J_k$ noemen (en $J_0 = 1$ ), dan geldt:

$$(3) \qquad D(\lambda) = \sum_{0}^{r} {}^{k} (-\lambda)^{r-k} J_k .$$

De wortels van de vergelijking $D(\lambda) = 0$ worden de eigen waarden van de matrix $M$ genoemd. Alleen voor die waarden van $\lambda$, waarvoor $D(\lambda) = 0$ is, heeft het stelsel (1) een niet-triviale oplossing. Deze oplossing wordt een eigen vector genoemd, behorende bij de eigen waarde $\lambda$ .
Zijn de $r$ wortels $\lambda_\alpha$, $\alpha \in \{0, 1, \ldots, r-1\}$ van $D(\lambda) = 0$ alle verschillend, dan kunnen we schrijven:

$$(4) \qquad D(\lambda) = (-1)^r \prod_{\alpha=0}^{r-1} (\lambda - \lambda_\alpha) .$$

In dit geval hoort bij elke $\lambda_\alpha$ precies één eigen vector $f_\alpha$ .
Komt de eigen waarde $\lambda_\rho$ echter voor met de veelvuldigheid $n_\rho$ , dus is:

$$(5) \qquad D(\lambda) = (-1)^r \prod_{\rho=0}^{s-1} (\lambda - \lambda_\rho)^{n_\rho} \quad \text{met} \quad \sum_{0}^{s-1} {}^{\rho} n_\rho = r \quad (\textit{alle } \lambda_\rho \textit{ ongelijk}),$$

dan is het mogelijk, dat bij de eigen waarde $\lambda_\rho$ meerdere l.o. eigen vectoren (maximaal $n_\rho$ ) gevonden worden. Dit hoeft echter niet, zoals uit de in 8.6 te geven theorie volgt.

Bij iedere eigen waarde behoort ook minstens een niet-triviale eigen covector $g_i$ , waarvoor geldt:

$$(6) \qquad \forall_j^r \, g_i M_j^i = \lambda g_j \quad \text{of} \quad g' M = \lambda g' .$$

Als $f$ een eigen vector is van $M$ bij de eigen waarde $\lambda$ , dan geldt dus $Mf = \lambda f$ en zal $M^2 f = MMf = M\lambda f = \lambda Mf = \lambda^2 f$, *dus*

$$(7) \qquad M^2 f = \lambda^2 f ,$$

terwijl in het algemeen geldt:

$$(8) \qquad \forall_n^N \, M^n f = \lambda^n f \text{ (door inductie te bewijzen)},$$

alsmede voor eigen covectoren:

$$(9) \qquad \forall_n^N \, g' M^n = \lambda^n g' \quad \text{als} \quad g' M = \lambda g' .$$

Is dus $f$ een eigen vector van $M$ bij eigen waarde $\lambda$ , dan zal $f$ tevens een eigen vector van $M^n$ zijn bij de eigen waarde $\lambda^n$ .

Is tenslotte $\varphi(x)$ een willekeurig polynoom en is $f$ , resp. $g'$, een eigen vector, resp. eigen covector bij de eigen waarde $\lambda$, dan is uit (7), (8) en (9) direct af te leiden, dat geldt:

(10) $$\varphi(M)f = \varphi(\lambda)f \quad \text{en}$$
$$g'\varphi(M) = \varphi(\lambda)g',$$

dus $f$ , resp. $g'$ is ook eigen vector, resp. eigen covector van de matrix $\varphi(M)$ bij de eigen waarde $\varphi(\lambda)$.

Hiermee is bewezen de stelling:

Alle matrices, die behoren tot de ring der polynomen in $M$ , hebben dezelfde eigen vectoren en eigen covectoren.

## 8.3 Enkele stellingen over eigen waarden en eigen vectoren van matrices

Stelling 1 Als $f$ een eigen vector en $g$ een eigen covector is van $M$ , behorende bij verschillende eigen waarden, dan geldt $g'f=0$.

Bewijs: Stel $f$ hoort bij eigen waarde $\lambda$, $g'$ bij $\mu$, dus $\lambda \neq \mu$.

$$\left.\begin{aligned} Mf &= \lambda f \\ g'M &= \mu g' \end{aligned}\right\} \text{dus } \lambda g'f = g'(\lambda f) = g'(Mf) = (g'M)f = \mu g'f.$$

Daar $\lambda \neq \mu$ is, volgt hieruit $g'f = g_i f^i = 0$ , q.e.d.
Men spreekt wel eens van de orthogonaliteitsrelatie.

Stelling 2 Als geen twee eigen waarden $\lambda_\alpha$ , $\alpha \in \{0,.....r-1\}$ van $M$ aan elkaar gelijk zijn, als $f_\alpha$ de eigen vector en $g_\alpha'$ de eigen covector voorstelt bij $\lambda_\alpha$ , dan geldt na normering der $g_{\alpha i}$ :

$$g_{\alpha i} f_\beta^i = \delta_{\alpha\beta}.$$

Bewijs: Het bewijs is geleverd, wanneer we aantonen, dat niet geldt: $\forall_\alpha\, g_{\alpha i} f_\beta^i = 0$.
Stel hiertoe, dat $\forall_\alpha\, g_{\alpha i} f_\beta^i = 0$ is, dan zijn er dus constanten $c^i$ te vinden, waarvan er minstens één $\neq 0$ is, zodanig dat $\forall_\alpha\, g_{\alpha i} c^i = 0$ is, waaruit volgt det.$(g_{\alpha i}) = 0$.
Dan zouden er echter ook getallen $b_\alpha$ te vinden zijn, zodanig dat $\forall_i \sum^\alpha b_\alpha g_{\alpha i} = 0$ is, terwijl minstens één $b_\alpha \neq 0$ is.

Dan is:

$$\forall_j \quad 0 = \sum^\alpha b_\alpha g_{\alpha i} (M^k)^i_j = \sum^\alpha \lambda_\alpha^k g_{\alpha j} b_\alpha \quad \text{volgens (8.2.9)}$$

Neem $j=1$ en geef $k$ de waarden $0,1,\ldots,r-1$ (de relatie geldt voor iedere natuurlijke $k$ ), dan hebben we $r$ vergelijkingen met $r$ onbekenden $g_{\alpha 1} b_\alpha$ , $\alpha \in \{0,\ldots,r-1\}$.

De determinant van dit stelsel wordt de determinant van Vandermonde :

det $(\lambda_\alpha^k, \alpha, k \in \{0, \ldots, r-1\}$ en deze determinant is $\neq 0$ , aangezien de $\lambda_\alpha$ verschillend zijn.
Dus geldt voor $j=1$ en analoog voor iedere $j$ :

$$\forall_j \, \forall_\alpha \; g_{\alpha j} b_\alpha = 0.$$

Voor minstens één $\alpha$ is $b_\alpha \neq 0$ ; voor deze $\alpha$ is $g_{\alpha 1} = g_{\alpha 2} = \ldots = g_{\alpha r} = 0$ dit is echter geen niet-triviale eigen covector. We hebben dus een tegenstrijdigheid gekregen.
Er is dus minstens één $\alpha$ , waarvoor geldt $g_{\alpha i} f^i_\beta \neq 0$ .Volgens stelling 1 moet deze gelijk $\beta$ zijn.
Normeren we nu nog $g$ , dan geldt dus $g_{\alpha i} f^i_\beta = \delta_{\alpha\beta}$ , q.e.d.

Stelling 3

Onder de voorwaarden van stelling 2 geldt na normering:

$$M^i_j = \sum^\alpha \lambda_\alpha f^i_\alpha g_{\alpha j}.$$

Bewijs: Stel $N^i_j = \sum^\alpha \lambda_\alpha f^i_\alpha g_{\alpha j}$ , dan is

$$\forall_\beta \; N^i_j f^j_\beta = \sum^\alpha \lambda_\alpha f^i_\alpha g_{\alpha j} f^j_\beta = \lambda_\beta f^i_\beta$$

dus iedere $f_\beta$ is eveneens een eigen vector van $N$ bij dezelfde eigen waarde $\lambda_\beta$ .
Hieruit volgt: $\forall_\beta (M-N) f_\beta = 0$ , dus daar de $f_\beta$ l.o. zijn:

$M = N$ , dus $M^i_j = \sum^\alpha_1 \lambda_\alpha f^i_\alpha g_{\alpha j}$ , q.e.d.

Stelling 4

Onder de voorwaarden van stelling 2 geldt algemeen na normering:

$$\forall^{N'}_k (M^k)^i_j = \sum^\alpha \lambda^k_\alpha f^i_\alpha g_{\alpha j}$$

Bewijs: Voor $k=1$ is het gestelde juist. Nemen we aan, dat het bewezen is voor $k=n$ , dan geldt voor $k=n+1$ :

$$(M^{n+1})^i_j = (M^n)^i_h M^h_j = \sum^\alpha \lambda^n_\alpha f^i_\alpha g_{\alpha h} \sum^\beta \lambda_\beta f^h_\beta g_{\beta j} = \sum^{\alpha,\beta} \lambda^n_\alpha \lambda_\beta f^i_\alpha g_{\alpha h} f^h_\beta g_{\beta j} =$$

$$= \sum^\alpha \lambda^{n+1}_\alpha f^i_\alpha g_{\alpha j} \text{, daar } g_{\alpha h} f^h_\beta = \delta_{\alpha\beta} \text{ is volgens stelling 2.}$$

Dus ook voor $k=n+1$ is het gestelde juist, waarmee de stelling door inductie bewezen is.

## 8.4 Minimumveelterm

In (8.2.2) definiëerden wij reeds: $\mathcal{D}(\lambda) = |M - \lambda I|$.
$\mathcal{D}(\lambda)$ bleek een veelterm van de graad $\tau$ in $\lambda$ te zijn en was ook te schrijven in de vorm

$$(8.2.5) \qquad \mathcal{D}(\lambda) = (-1)^{\tau} \prod_{\rho=0}^{s-1} (\lambda - \lambda_\rho)^{n_\rho} \quad met \quad \sum_{0}^{s-1} n_\rho = \tau \quad en \quad \lambda_\rho \neq$$

(geen twee $\lambda$'s aan elkaar gelijk).

Tevens is gedefiniëerd $\mathcal{D}(M)$, een veelterm in de matrix $M$, zelf een matrix, behorende tot de commutatieve ring van alle polynomen in $M$.
Nu geldt:

Stelling 1: Iedere matrix voldoet aan zijn eigen karakteristieke vergelijking.

Dus $\mathcal{D}(M) = 0$ (Stelling van Cayley-Hamilton-Frobenius).

Bewijs: De geadjungeerde matrix van $M - \lambda I$ noemen we $N$, zijn elementen $N^i_j(\lambda)$, veeltermen van de graad $\tau - 1$ of lager in $\lambda$, dus

$$(1) \qquad (M^i_j - \lambda I^i_j) N^j_k(\lambda) = I^i_k \mathcal{D}(\lambda) \quad of \quad (M - \lambda I) N(\lambda) = I . \mathcal{D}(\lambda).$$

Links en rechts staan veeltermen van de graad $\tau$ in $\lambda$. De coëfficiënten van deze veeltermen zijn getallen. Daar de gelijkheid (1) voor iedere $\lambda$ geldt, moeten de coëfficiënten van overeenkomstige machten van $\lambda$ links en rechts aan elkaar gelijk zijn. Deze gelijkheid blijft uiteraard gelden als we voor $\lambda$ de matrix $M$ invullen. Formeel verandert er aan het rekenschema immers niets. Aan de identiteit (1) blijft dus voldaan als wij voor $\lambda$ de matrix $M$ invullen. Wij krijgen dan:

$$(M - MI) N(M) = I \mathcal{D}(M), \text{ dus } \mathcal{D}(M) = 0, \text{ q.e.d.}$$

Alle veeltermen in $\lambda$ met coëfficiënten uit een lichaam vormen een commutatieve ring. De veeltermen $R$ uit deze ring met de eigenschap $R(M) = 0$ vormen een ideaal[1] in deze ring. Daar in een veeltermring met coëfficiënten uit een lichaam elk ideaal tevens hoofdideaal[1] is, is er een kleinste veelterm in $\lambda$, $\Psi(\lambda)$, die de eigenschap $\Psi(M) = 0$ heeft, en de minimumveelterm wordt genoemd. Dit is de grootste gemene deler van alle polynomen in $\lambda$, die deze eigenschap hebben.
In $\Psi(\lambda)$ nemen we de coëfficiënt van de hoogste macht van $\lambda$ gelijk 1. De graad van $\Psi(\lambda)$ kan lager zijn dan van $\mathcal{D}(\lambda)$.

Noemen wij de wortels van $D(\lambda)=0$ $\lambda_\rho$, hun multipliciteit in $D(\lambda)$ $n_\rho$ en in $\Psi(\lambda)$ $m_\rho$, waarbij we er zorg voor dragen, dat alle $\lambda_\rho$'s verschillend zijn, dan geldt dus:

$$D(\lambda)=(-1)^r\prod(\lambda-\lambda_\rho)^{n_\rho} \quad, n_\rho>0 \quad (8.2.5) \text{ en}$$

(2) $$\Psi(\lambda)=\prod(\lambda-\lambda_\rho)^{m_\rho}, \quad m_\rho \leqq n_\rho \text{ , alsmede de}$$

Stelling 2 :

Als $\lambda_\rho$ een eigen waarde is van $M$ , dan bevat de minimum-veelterm de factor $(\lambda-\lambda_\rho)$ minstens éénmaal.

Dus $\forall_\rho \; n_\rho>0 \longrightarrow m_\rho>0$ .

Bewijs: Stel, dat er een $\rho$ is, waarvoor $n_\rho>0$ en $m_\rho=0$ is. Dan zijn de veeltermen $\Psi(\lambda)$ en $\lambda-\lambda_\rho$ dus onderling ondeelbaar. Daar wij hier te maken hebben met polynomen in 1 variabele over het lichaam der complexe getallen, die dus een ring met hoofdidealen [1] vormen, volgt uit de ondeelbaarheid van $\Psi(\lambda)$ en $\lambda-\lambda_\rho$, dat er veeltermen $Q_1(\lambda)$ en $Q_2(\lambda)$ bestaan, zodanig dat

$\Psi(\lambda)\,Q_1(\lambda)+(\lambda-\lambda_\rho)\,Q_2(\lambda)=1$ , dus na substitutie van de matrix $M$ voor $\lambda$ :

$\Psi(M)\,Q_1(M)+(M-\lambda_\rho I)\,Q_2(M)=I$ of, daar $\Psi(M)=0$ en $(M-\lambda_\rho I)$ en $Q_2(M)$ veeltermen in dezelfde matrix $M$ , dus verwisselbaar zijn:

(3) $$Q_2(M)\,(M-\lambda_\rho I)=I.$$

Gegeven is, dat $D(\lambda_\rho)=0$ . Hieruit volgt, dat er een niet-triviale oplossing bestaat van de vergelijkingen $(M-\lambda_\rho I)f=0$ .

Passen we nu (3) toe op deze oplossing $f$, dan krijgen we:

$Q_2(M)(M-\lambda_\rho I)f=f$ of volgens de associativiteit der matrixvermenigvuldiging:

$Q_2(M)\{(M-\lambda_\rho I)f\}=f$ , maar $(M-\lambda_\rho I)f=0$ , zodat $f=0$ , waarmee we op een tegenspraak komen. $\Psi(\lambda)$ en $\lambda-\lambda_\rho$ zijn dus niet ondeelbaar, dus $m_\rho>0$ , q.e.d.

## 8.5 Splitsing van een matrix

Voor willekeurige $\lambda$ beschouwen we:

(1) $$g_\rho(\lambda) \overset{\text{def}}{=} \frac{\Psi(\lambda)}{(\lambda-\lambda_\rho)^{m_\rho}}.$$

Door breuksplitsing toe te passen kunnen wij schrijven:

(2) $$\frac{1}{\Psi(\lambda)}=\sum^{\rho}\frac{h_\rho(\lambda)}{(\lambda-\lambda_\rho)^{m_\rho}},$$

waarbij de graad van $h_\rho(\lambda)$ kleiner dan $m_\rho$ is.

Uit (1) en (2) volgt:

$$1 = \sum^\rho \frac{h_\rho(\lambda)}{(\lambda-\lambda_\rho)^{m_\rho}}\,\psi(\lambda) = \sum^\rho h_\rho(\lambda)\,g_\rho(\lambda)\;, \text{ dus}:$$

(3) $$\sum^\rho h_\rho(\lambda)\,g_\rho(\lambda) = 1\,,$$

een identiteit in $\lambda$ , waarbij de $h$'s en de $g$'s veeltermen in $\lambda$ zijn. In deze identiteit substitueren we voor $\lambda$ weer de matrix $M$ . Definiëren wij:

(4) $$E_\rho \overset{\text{def}}{=} h_\rho(M)\,g_\rho(M)\,,$$

dan is $E_\rho$ dus een product van matrices, dus zelf een matrix en volgt uit (3):

(5) $$\sum E_\rho = I \quad , \text{ de eenheidsmatrix.}$$

Er geldt:

(6) $$E_\rho E_\sigma = \delta_{\rho\sigma} E_\rho\,, \text{ dus}$$
$$E_\rho^2 = E_\rho$$
$$E_\rho E_\sigma = 0 \quad \text{voor } \rho \neq \sigma\,.$$

Immers als $\rho \neq \sigma$ is, dan geldt, daar veeltermen in $M$ verwisselbaar zijn:

$$E_\rho E_\sigma = h_\rho(M)\,g_\rho(M)\,h_\sigma(M)\,g_\sigma(M) = h_\rho(M)\,h_\sigma(M)\,g_\rho(M)\,g_\sigma(M) = 0\,,$$

daar uit (1) volgt $g_\rho(\lambda)\,g_\sigma(\lambda) \equiv 0\,(\psi(\lambda))$ [1], dus $g_\rho(M)\,g_\sigma(M) = 0$.

$$E_\rho^2 = E_\rho E_\rho = E_\rho\,(I - \sum_{\sigma\neq\rho} E_\sigma) = E_\rho\,.$$

De $E$'s zijn dus <u>idempotent</u> [2] en orthogonaal .

Wij hebben in (5) de eenheidsmatrix additief gesplitst in idempotente, elkaar annulerende matrices. In 8.6 zullen we zien, dat de rang van $E_\rho$ gelijk $n_\rho$ is.

Definiëren wij:

(7) $$F_\rho \overset{\text{def}}{=} (M-\lambda_\rho)E_\rho \quad \text{(waarbij wij } M-\lambda_\rho \text{ i.p.v. } M-\lambda_\rho I \text{ schrijven),}$$

dan zijn $M-\lambda_\rho$ en $E_\rho$ weer verwisselbaar en geldt:

$$F_\rho^{m_\rho} = (M-\lambda_\rho)^{m_\rho} E_\rho = \psi(M)\,h_\rho(M) = 0\;, \text{ dus}:$$

(8) $$F_\rho^{m_\rho} = 0\,,\; F_\rho^k \neq 0 \text{ voor } 0 < k < m_\rho\,,\; F_\rho^k = 0 \text{ voor } k \geq m_\rho\,.$$

---

1) $a \equiv 0(b)$ wil zeggen: $a$ is deelbaar door $b$ .

2) $A$ is idempotent als $A^2 = A$ .

Bovendien is direct duidelijk:

(9) $F_\rho E_\rho = E_\rho F_\rho = F_\rho$.

(10) $F_\rho E_\sigma = 0$ voor $\rho \neq \sigma$,

(11) $F_\rho F_\sigma = 0$ voor $\rho \neq \sigma$.

De F 's zijn dus orthogonaal en nilpotent.[1]
Voor M geldt nu:

$$M = MI = \Sigma^\rho M E_\rho = \Sigma^\rho \lambda_\rho E_\rho + \Sigma^\rho (M - \lambda_\rho) E_\rho \text{, dus}$$

(12) $$M = \Sigma^\rho \lambda_\rho E_\rho + \Sigma^\rho F_\rho .$$

Verder:

$$M^2 = \left(\Sigma^\rho(\lambda_\rho E_\rho + F_\rho)\right)^2 = \Sigma^\rho (\lambda_\rho E_\rho + F_\rho)^2 + \sum_{\rho \neq \sigma}^{\rho,\sigma} (\lambda_\rho E_\rho + F_\rho)(\lambda_\sigma E_\sigma + F_\sigma).$$

Nu is de tweede term van het rechterlid o vanwege de relaties (6), (10) en (11), dus:

$$M^2 = \Sigma^\rho (\lambda_\rho E_\rho + F_\rho)^2 .$$

In het algemeen vindt men:

(13) $$\forall_n^{N'} M^n = \Sigma^\rho (\lambda_\rho E_\rho + F_\rho)^n .$$

Ontwikkelen we het rechterlid van (13) in een binomiaalreeks, dan krijgen we:

$$M^n = \Sigma^\rho \sum_0^{n} {}^k \binom{n}{k} \lambda_\rho^{n-k} F_\rho^k \text{, mits :}$$

(14) $$F_\rho^0 \overset{\text{def}}{=} E_\rho .$$

De index $k$ kunnen we laten lopen van o tot $m_\rho - 1$, daar voor $m_\rho \leqq n$ geldt $F_\rho^k = 0$ zodra $k \geqq m_\rho$ is, terwijl voor $m_\rho > n$ de binomiaalcoëfficiënten nul worden, zodra $k > n$ is.
Dus:

(15) $$M^n = \Sigma^\rho \sum_0^{m_\rho - 1} {}^k \binom{n}{k} \lambda_\rho^{n-k} F_\rho^k .$$

## 8.6 Normaalvorm van een matrix

Wij gaan $r$ l.o. vectoren zoeken met de eigenschap, dat er een lineaire transformatie bestaat, die deze vectoren overvoert in eenheidsvectoren, zodanig, dat de matrix M getransformeerd wordt in de nog nader te definiëren normaalvorm. Hiertoe gaan wij als volgt te werk:

Wij beschouwen een bepaalde eigen waarde $\lambda_\rho$ van M , die de

---

1) A is nilpotent als $\exists_k A^k = 0$.

multipliciteit $n_\rho$ heeft en in de minimumveelterm met veelvuldigheid $m_\rho$ voorkomt. De index $\rho$ is dus vast. Deze zullen wij weglaten.

$$F = (M-\lambda)E,\ F^0 = E,\ F^m = 0,\ F^k \neq 0 \text{ als } k < m.$$

De rang van $F^{m-1}$ noemen we $r_1$, $r_1 > 0$.
Er zijn nu precies $r_1$ vectoren $f_{0\alpha}^{(1)}$, $\alpha \in \{1,\ldots\ldots,r_1\}$ te vinden, die l.o. zijn mod $(M-\lambda)^{m-1}$. 1)

Verder definiëren we:

$$f_{1\alpha}^{(1)} \overset{\text{def}}{=} (M-\lambda) f_{0\alpha}^{(1)},\ f_{2\alpha}^{(1)} \overset{\text{def}}{=} (M-\lambda)^2 f_{0\alpha}^{(1)},\ \ldots\ldots,\ f_{m-1,\alpha}^{(1)} \overset{\text{def}}{=} (M-\lambda)^{m-1} f_{0\alpha}^{(1)}.$$

Voor deze vectoren geldt $\forall_{k \in \{0,\ldots,m-1\}}$:

$$(1)\qquad \begin{cases} F^{m-k} f_k = 0 \\ F^{m-k-1} f_k \neq 0, \text{ hangt niet af van } k \\ (M-\lambda) f_k = f_{k+1}, \text{ waarbij wij } f_m \equiv 0 \text{ definiëren.} \end{cases}$$

De $f_{k\alpha}^{(1)}$ zijn voor iedere vaste $k$ l.o. mod $(M-\lambda)^{m-k-1}$, want stel $\sum_1^{r_1} c_\alpha (M-\lambda)^{m-k-1} f_{k\alpha}^{(1)} = 0$, dan is dus

$$\sum_1^{r_1} c_\alpha (M-\lambda)^{m-1} f_{0\alpha}^{(1)} = 0, \text{ dus } \forall_\alpha\, c_\alpha = 0.$$

Totaal vinden we $m r_1$ vectoren, die l.o. zijn, want stel:

$$\sum_1^{r_1} \sum_0^{m-1} c_{\alpha k} f_{k\alpha}^{(1)} = 0$$
, dan blijkt door de operator $(M-\lambda)^{m-1}$ toe te passen $\forall_\alpha\, c_{\alpha 0} = 0$.

Neem aan, dat voor $k < \ell$ ($\ell \in \{1,\ldots\ldots,m-1\}$) reeds bewezen is $c_{\alpha k} = 0$ en pas de operator $(M-\lambda)^{m-\ell-1}$ toe:

$$\sum_1^{r_1} \sum_\ell^{m-1} c_{\alpha k} (M-\lambda)^{m-\ell-1} f_{k\alpha}^{(1)} = 0, \text{ dus } \sum_1^{r_1} c_{\alpha\ell} (M-\lambda)^{m-\ell-1} f_{\ell\alpha}^{(1)} = 0,$$

$$\text{dus } \forall_\alpha\, c_{\alpha\ell} = 0, \text{ daar de } f_{\ell\alpha}^{(1)} \text{ l.o. mod } (M-\lambda)^{m-\ell-1} \text{ zijn.}$$

Door inductie is dus bewezen $\forall_\alpha \forall_k\, c_{\alpha k} = 0$.

Vervolgens zoeken we vectoren $f^{(2)}$, waarvoor geldt $F^{m-1} f^{(2)} = 0$ en die mod $(M-\lambda)^{m-2}$ onderling en met de $f^{(1)}$'s l.o. zijn. Zo er dergelijke vectoren zijn, geven wij ze aan met $f_{1\beta}^{(2)}$ en definiëren wij:

$$f_{2\beta}^{(2)} = (M-\lambda) f_{1\beta}^{(2)},\ \ldots\ldots,\ f_{m-1,\beta}^{(2)} = (M-\lambda)^{m-2} f_{1\beta}^{(2)}.$$

De vectoren $f_{k\beta}^{(2)}$ zijn onderling en met de $f_{k\alpha}^{(1)}$ l.o. en voldoen voor alle $k \in \{1,\ldots,m-1\}$ aan (1).

---

1) De vectoren $f_\alpha$, $\alpha \in \{1,\ldots,r\}$ zijn l.o. mod.(modulo) $(M-\lambda)^k$, wanneer uit $\sum_1^{r} a_\alpha (M-\lambda)^k f_\alpha = 0$ volgt $\forall_\alpha\, a_\alpha = 0$. Vanzelfsprekend zijn de vectoren $f_\alpha$ dan ook l.o.

Hebben we vectoren $f^{(i-1)}$ gevonden, dan zoeken wij vervolgens vectoren $f^{(i)}$, waarvoor geldt $F^{m-i+1}f^{(i)}=0$ en die mod.$(M-\lambda)^{m-i-j+1}$ onderling en met de $f^{(j)}$ 's l.o. zijn $(\forall_{j\leq i-1})$. Met behulp hiervan definiëren we dan weer vectoren:

$$(M-\lambda)f^{(i)},\ (M-\lambda)^2f^{(i)},\ldots,(M-\lambda)^{m-i}f^{(i)}.$$

Deze voldoen eveneens aan (1).
De laatste vectoren, die we zo vinden zijn vectoren $f^{(m)}$. Al deze vectoren zijn l.o. Het totaal aantal, dat wij bij de eigen waarde $\lambda_\rho$ vinden, noemen wij $n'_\rho$.

Voor deze $n'_\rho$ vectoren geldt dus:

$$(2)\qquad \forall_{i\in\{1,\ldots,m_\rho\}}\forall_{k\in\{i-1,\ldots,m_\rho-1\}}\ (M-\lambda_\rho)f_{\rho_{k\iota}}^{(i)}=f_{\rho_{k+1,\iota}}^{(i)}.$$

Nu is $\sum^\rho n'_\rho=\tau$. Immers als $f$ een willekeurige vector is (natuurlijk met $\tau$ componenten), dan is dus $f=If=\sum E_\rho f$ ; $E_\rho f=F_\rho^\circ f$ is lineair uit te drukken in de vectoren $f_\rho$, dus ook $f$ is lineair uit te drukken in de gevonden vectoren. Wij vinden dus inderdaad een volledig stelsel vectoren, dus $\sum^\rho n'_\rho=\tau$.

De vectoren $f_{\rho_{k\iota}}^{(i)}$ ordenen wij nu zodanig, dat voldaan is aan:

(3)
- a. Als $\rho_1<\rho_2$ is, dan staan de $f_{\rho_1}$ voor de $f_{\rho_2}$.
- b. Als $i_1<i_2$ is bij vaste $\rho$, dan staan de $f_\rho^{(i_1)}$ voor $f_\rho^{(i_2)}$
- c. Als $\iota_1<\iota_2$ is bij vaste $\rho$ en $i$, dan staan de $f_{\rho_{\iota_1}}^{(i)}$ voor de $f_{\rho_{\iota_2}}^{(i)}$
- d. Als $k_1<k_2$ is bij vaste $\rho$, $i$ en $\iota$, dan staat $f_{\rho_{k_1\iota}}^{(i)}$ <u>achter</u> $f_{\rho_{k_2\iota}}^{(i)}$.

Het begin van de rij ziet er dus als volgt uit:

$$f_{1_{m_1-1\,1}}^{(1)},f_{1_{m_1-2\,1}}^{(1)},\ldots,f_{1_{11}}^{(1)},f_{1_{01}}^{(1)};\ f_{1_{m_1-1\,2}}^{(1)},\ldots,f_{1_{02}}^{(1)};\ldots,f_{1_{0m_1}}^{(1)};f_1^{(2)},\ldots;f_{2_{m_2-1\,1}}^{(1)},\ldots$$

Door middel van een lineaire transformatie $T$ zorgen wij nu, dat deze $\tau$ vectoren eenheidsvectoren worden en wel zodanig, dat de $f$ op de $j$-de plaats in de rij overgaat in $e_j$ ( $e_j$ is de uit $\tau$ componenten bestaande eenheidsvector met de 1 op de $j$-de plaats).

$$M^*\overset{def}{=}T^{-1}MT,\ E_\rho^*\overset{def}{=}T^{-1}E_\rho T,\ F_\rho^*\overset{def}{=}T^{-1}F_\rho T.$$

Uit $(M-\lambda_1)f_{1_{m_1-1\,1}}^{(1)}=0$ volgt nu $M^*e_1=\lambda_1e_1$, dus $M^{*i}_je_1^j=\lambda_1e_1^i$, dus $M^{*1}_1=\lambda_1, M^{*i}_1=0$ voor $i\neq1$, waarmee de eerste kolom van $M^*$ gevonden is.

Uit $(M-\lambda_1) f^{(1)}_{1\,m_1-2\,1} = f^{(1)}_{1\,m_1-1\,1}$ volgt $M^* e_2 = \lambda_1 e_2 + e_1$, dus $M^{*i}_{\ \ j}\, e_2^j = \lambda_1 e_2^i + e_1^i$, dus $M^{*1}_{\ \ 2} = 1$, $M^{*2}_{\ \ 2} = \lambda_1$, $M^{*i}_{\ \ 2} = 0$ voor $i \geq 3$, waarmee de tweede kolom van $M^*$ gevonden is.
Wij vinden zo voor $M^*$ een matrix van de vorm:

$$(4)\qquad M^* = \begin{pmatrix} \boxed{\begin{matrix}\lambda_1 & 1 & & 0\\ & \lambda_1 & 1 & \\ & & \ddots & 1\\ 0 & & & \lambda_1\end{matrix}} & & & 0 \\ & \boxed{\begin{matrix}\lambda_1 & 1 & & 0\\ & \ddots & \ddots & \\ & & \ddots & 1\\ 0 & & & \lambda_1\end{matrix}} & & \\ & & \boxed{\begin{matrix}\lambda_2 & 1 & & 0\\ & \ddots & \ddots & \\ & & \ddots & 1\\ 0 & & & \lambda_2\end{matrix}} & \\ 0 & & & \ddots \end{pmatrix}$$

$M^*$ bestaat dus uit een aantal vierkanten van de vorm

$$(5)\qquad \left|\begin{matrix}\lambda_\rho & 1 & & 0\\ & \ddots & \ddots & \\ & & \ddots & 1\\ 0 & & & \lambda_\rho\end{matrix}\right|$$

Het totaal aantal keren, dat op de hoofddiagonaal van voorkomt, bedraagt
Achteraf laten wij nu zien:
Immers $|M-\lambda I| = |T^{-1}|\,|M-\lambda I|.|T| = |T^{-1}(M-\lambda I)T| = |M^* - \lambda T^{-1} I T| = |M^* - \lambda I|$.

Nu bevat $|M^* - \lambda I|$, zoals uit (4) volgt de factor $(\lambda-\lambda_\rho)^{n'_\rho}$, $|M-\lambda I|$ bevat de factor $(\lambda-\lambda_\rho)^{n_\rho}$. Dus $n'_\rho = n_\rho$. Dit geldt voor alle $\rho$.
$M^*$ wordt de <u>normaalvorm</u> van de matrix $M$ genoemd.
$M^*$ en $M$ hebben dezelfde karakteristieke vergelijking, dus dezelfde eigen waarden. Is $f$ (respectievelijk $f^*$) een eigen vector van $M$ (resp. $M^*$) bij eigen waarde $\lambda$, dan zal $T^{-1}f$ (resp. $Tf^*$) eigen vector van $M^*$ (resp. $M$) zijn bij dezelfde eigen waarde $\lambda$. Immers stel $Mf = \lambda f$, $M^* = T^{-1}MT$, dus $M = TM^*T^{-1}$, dus $TM^*T^{-1}f = \lambda f$, waaruit wij zien $M^*(T^{-1}f) = \lambda(T^{-1}f)$. Evenzo volgt uit $M^*f^* = \lambda f^*$, dat $M(Tf^*) = \lambda(Tf^*)$.

Het aantal l.o. eigen vectoren, dat bij een bepaalde eigen waarde gevonden kan worden, is dus invariant bij lineaire transformatie (wij maakten geen gebruik van de speciale vorm van $M^*$ ).

Dit aantal is bij $M^*$ eenvoudig te berekenen. Beschouw namelijk een speciale eigen waarde $\lambda_\rho$ van $M^*$ (en $M$ ) en stel, dat er in de normaalvorm $a_\rho$ vierkanten zijn met $\lambda_\rho$ op de hoofddiagonaal. Het is direct duidelijk, dat $(M^*-\lambda_\rho I)$ de rang $r-a_\rho$ heeft, zodat er dus $a_\rho$ l.o. vectoren te vinden zijn met de eigenschap $(M^*-\lambda_\rho I)f=0$ . Het aantal eigen vectoren van een matrix bij een bepaalde eigen waarde is dus gelijk aan het aantal vierkanten, dat in de normaalvorm van de matrix voorkomt met die eigen waarde op de hoofddiagonaal.

We hebben dus: $1 \leqq a_\rho \leqq n_\rho$ voor alle $\rho$ , waarbij $a_\rho$ het bedoelde aantal is. Zodra er in de normaalvorm een $\lambda_\rho$ -vierkantje is, waarin $\lambda_\rho$ meer dan éénmaal voorkomt (en waarin dus ook minstens één element 1 voorkomt) zal $a_\rho < n_\rho$ zijn. Dat dit mogelijk was beweerden wij reeds in 8.2. Er zal maar één eigen vector gevonden worden bij eigen waarde $\lambda_\rho$ met multipliciteit $n_\rho > 1$ als er in de normaalvorm maar één vierkant is met $\lambda_\rho$ op de hoofddiagonaal (dus met lengte $n_\rho$ en $n_\rho - 1$ elementen 1 op de nevendiagonaal).

Uit $M = \sum(\lambda_\rho E_\rho + F_\rho)$ , dus $M^* = \sum(\lambda_\rho E_\rho^* + F_\rho^*)$ volgt verder, dat $E_\rho^*$ een matrix is met buiten de hoofddiagonaal elementen 0 en op de hoofddiagonaal alleen elementen 1 op die plaatsen, waar in $M^*$ $\lambda_\rho$ staat en verder nullen. $E_\rho^*$ , dus ook $E_\rho$ , heeft derhalve de rang $n_\rho$ .(De rang van een matrix is namelijk invariant bij lineaire transformatie).

$F_\rho^*$ is een matrix met alleen elementen 1 op die plaatsen, waar in $M^*$ een 1 voorkomt in een $\lambda_\rho$ -vierkant (naast $\lambda_\rho$ op de hoofddiagonaal). De overige elementen van $F_\rho^*$ zijn $0$ .

Dus:

$$(6) \qquad E_1^* = \begin{pmatrix} 1 & & & & 0 \\ & \ddots & & & \\ & & 1 & & \\ & & & 0 & \\ 0 & & & & \ddots \\ & & & & & 0 \end{pmatrix}, \; n_1 \text{ elementen } 1$$

$$(7) \qquad F_1^* = \begin{pmatrix} 0 & 1 & & & & 0 \\ & \ddots & \ddots & & & \\ & & 0 & 1 & & \\ & & & 0 & 1 & \\ & & & & \ddots & \ddots \\ 0 & & & & & 0 \\ & & & & & & 0 \end{pmatrix}, \; n_1 - a_1 \text{ elementen } 1$$

Laat zien, dat de betrekkingen (5) tot en met (11) van 8.5 gelden. Uit (8.5.7) volgt $F_\rho^* = (M^* - \lambda_\rho) E_\rho^*$ . Laat zien, dat dit inderdaad het geval is.

## 8.7 Elementaire Delers

Terwille van de volledigheid noemen wij in deze paragraaf nog enkele stellingen uit de theorie van de elementaire delers. Deze stellingen worden echter niet door ons gebruikt bij de theorie van de stochastische processen. De bewijzen zullen wij dan ook meestal achterwege laten.

Wij beschouwen matrices $A(\lambda)$ , waarvan de elementen veeltermen zijn in $\lambda$ met coëfficiënten uit het lichaam der complexe getallen. Dus:

$$A(\lambda) = \begin{pmatrix} a_{11}(\lambda) & a_{12}(\lambda) & \cdots & a_{1r}(\lambda) \\ a_{21}(\lambda) & & & \vdots \\ \vdots & & & \vdots \\ a_{r1}(\lambda) & \cdots & \cdots & a_{rr}(\lambda) \end{pmatrix},$$ waarbij $a_{ij}(\lambda)$ een veelterm in $\lambda$ is.

Onder elementaire transformaties van een dergelijke matrix verstaan wij de volgende operaties:

1. Vermenigvuldiging van een rij of kolom met een willekeurig complex getal $\neq 0$ .
2. Optelling van een met een willekeurige veelterm vermenigvuldigde rij (kolom) bij een andere rij (kolom).

Door middel van elementaire transformaties kunnen wij altijd bereiken, dat twee rijen of kolommen van de matrix verwisseld worden.

Elementaire transformaties zijn lineair.

Wij noemen twee matrices aequivalent als ze door elementaire transformaties in elkaar over te voeren zijn.

Nu geldt:

Stelling 1: Iedere matrix $A(\lambda)$ is aequivalent met een matrix van de vorm:

(2) $$\begin{pmatrix} e_1(\lambda) & & & & & 0 \\ & e_2(\lambda) & & & & \\ & & \ddots & & & \\ & & & e_s(\lambda) & & \\ & & & & 0 & \\ 0 & & & & & \ddots \;\; 0 \end{pmatrix},$$ waarbij de $e_i(\lambda)$ veeltermen in $\lambda$ voorstellen ; $i \in \{1, \ldots, s\}$ ,

$s \leq r$ ; $e_1(\lambda)$ een deler is van $e_2(\lambda)$ , $e_2(\lambda)$ van $e_3(\lambda)$, dus in

het algemeen $e_i(\lambda)$ van $e_{i+1}(\lambda)$.
De coëfficiënt van de hoogste macht van $\lambda$ is in elke $e(\lambda)$ gelijk 1.

Zij nu $d_k(\lambda)$ de grootste gemene deler van alle onderdeterminanten van de orde $k$ van $A(\lambda)$. De coëfficiënt van de hoogste macht van $\lambda$ in $d_k(\lambda)$ noemen wij echter 1 en wij definiëren verder $d_k(\lambda)=0$ als alle onderdeterminanten van de orde $k$ gelijk 0 zijn, identiek in $\lambda$. Dan geldt de:

Stelling 2: $d_k(\lambda)$ is invariant bij elementaire transformatie voor elke $k \in \{1,\dots,r\}$.
Hieruit volgt, dat wij de $d_k(\lambda)$'s van $A(\lambda)$ kunnen berekenen door uit te gaan van de met $A(\lambda)$ aequivalente matrix (2).
Duidelijk blijkt zo:

$$d_k(\lambda) = e_1(\lambda)\,e_2(\lambda)\ldots\ldots e_k(\lambda) \quad \text{voor } k\leqq s \quad \text{en } d_k(\lambda)=0 \text{ voor } k>s.$$

Daar ook $d_{k-1}(\lambda) = e_1(\lambda)\ldots\ldots\ldots e_{k-1}(\lambda)$ is dus:

$$(3) \qquad e_k(\lambda) = \frac{d_k(\lambda)}{d_{k-1}(\lambda)}, \text{ waarbij } k \in \{1,\dots,r\} \text{ en } d_0(\lambda) \stackrel{\text{def}}{=} 1.$$

Men noemt de $e_k(\lambda)$'s de <u>elementaire</u> delers van de matrix $A(\lambda)$.
Dus geldt:

Stelling 3: Twee matrices met veeltermen als elementen zijn dan en slechts dan aequivalent als ze dezelfde elementaire delers hebben.

Als nu $M$ en $N$ twee "gewone" matrices zijn (dus met elementen, die niet van een variabele $\lambda$ afhangen, dan geldt de belangrijke

Stelling 4: Twee matrices $M$ en $N$ zijn dan en slechts dan door middel van een lineaire transformatie in elkaar over te voeren als de matrices $M-\lambda I$ en $N-\lambda I$ dezelfde elementaire delers hebben, dus als $M-\lambda I$ en $N-\lambda I$ aequivalent zijn.

Daar $|M-\lambda I| = (-1)^r d_r(\lambda) = (-1)^r \prod_\rho (\lambda-\lambda_\rho)^{n_\rho}$ en $d_r(\lambda) = e_1(\lambda)\ldots e_r(\lambda)$ zijn de elementaire delers van $(M-\lambda I)$ van de vorm:

$$(4) \qquad e_k(\lambda) = \prod_\rho (\lambda-\lambda_\rho)^{v_{k\rho}} \text{ met } \sum_1^r{}^k v_{k\rho} = n_\rho \text{ en } v_{1\rho} \leqq v_{2\rho} \leqq \ldots \leqq v_{s\rho}.$$

Is de matrix $M$ van de vorm (8.6.4), dus opgebouwd uit vierkanten van de vorm (8.6.5), dan zijn de elementaire delers van $M-\lambda I$ eenvoudig uit te rekenen en eveneens van de vorm (4). Uit de voorwaarde $\sum_1^r{}^k v_{k\rho} = n_\rho$ volgt, dat als er in $M$ (van de vorm (8.6.4)) een vierkant is met $\lambda_\rho$ op de hoofddiagonaal en lengt- $\ell_\rho$, dat dan $(\lambda-\lambda_\rho)^{\ell_\rho}$ als factor in precies één elementaire

deler van $M-\lambda I$ voorkomt.
Houden wij verder rekening met de in (4) genoemde voorwaarde $v_{1p} \leqq v_{2p} \leqq \ldots\ldots \leqq v_{sp}$, dan zijn de $e(\lambda)$'s in elk concreet geval van een matrix van de vorm (8.6.4) eenvoudig te berekenen.
Stel bijvoorbeeld:

$$M-\lambda I = \left(\begin{array}{cccccccccc}
\lambda_1-\lambda & 1 & 0 & & & & & & & \\
0 & \lambda_1-\lambda & 1 & & & & & & & \\
0 & 0 & \lambda_1-\lambda & & & & & & & \\
 & & & \lambda_1-\lambda & 1 & & & & & \\
 & & & 0 & \lambda_1-\lambda & & & & & \\
 & & & & & \lambda_2-\lambda & 1 & & & \\
 & & & & & 0 & \lambda_2-\lambda & & & \\
 & & & & & & & \lambda_2-\lambda & 1 & \\
 & & & & & & & 0 & \lambda_2-\lambda & \\
 & & & & & & & & & \lambda_2-\lambda
\end{array}\right)$$

De vierkanten leveren de factoren (op het teken na):

$$(\lambda-\lambda_1)^3, (\lambda-\lambda_1)^2, (\lambda-\lambda_2)^2, (\lambda-\lambda_2)^2 \text{ en } (\lambda-\lambda_2).$$

Dus $e_s(\lambda) = (\lambda-\lambda_1)^3(\lambda-\lambda_2)^2$, $e_{s-1}(\lambda) = (\lambda-\lambda_1)^2(\lambda-\lambda_2)^2$, $e_{s-2}(\lambda) = (\lambda-\lambda_2)$,
$e_k(\lambda) = 1$ *als* $k < s-2$.
Verder is $s = r = 10$, daar $e_{10}(\lambda) = \dfrac{d_{10}(\lambda)}{d_9(\lambda)} \neq 0$, daar $d_{10}(\lambda) = |M-\lambda I| \neq 0$.

$$d_9(\lambda) = \frac{d_{10}(\lambda)}{e_{10}(\lambda)} = (\lambda-\lambda_1)^2(\lambda-\lambda_2)^3.$$ Verifieer dit rechtstreeks!

Is $M$ een willekeurige matrix, dan is er altijd een matrix van de vorm (8.6.4) te construeren, die dezelfde elementaire delers heeft als $M$. De lengte van de benodigde vierkanten volgt immers direct uit (4). Uit stelling 4 volgt nu direct, dat iedere matrix door een lineaire transformatie in de vorm (8.6.4) te brengen is.
Hiermede is dus een tweede bewijs aangegeven van de stelling, dat een matrix door een lineaire transformatie in zijn normaalvorm te brengen is.
Verder geldt:

<u>Stelling 5</u>: De laatste elementaire deler $\neq 0$, van $(M-\lambda I)$, $e_s(\lambda)$, is gelijk aan de minimumveelterm $\psi(\lambda)$ van $M$.
In ons voorbeeld is de minimumveelterm dus $(\lambda-\lambda_1)^3(\lambda-\lambda_2)^2$.
In het algemeen vinden wij als $|M-\lambda I| \neq 0$ volgens (3) en stelling 5 de minimumveelterm door de g.g.d. van alle onderc

terminanten van de orde $r-1$ te delen op $|M-\lambda I|$.

Stelling 5 is direct duidelijk als wij bedenken, dat zowel in $\psi(\lambda)$ als in $e_s(\lambda)$ de multipliciteit van $\lambda_\rho$ gelijk is aan de lengte van het grootste vierkantje van (8.6.4), waarin $\lambda_\rho$ op de hoofddiagonaal staat.

Wat $e_s(\lambda)$ betreft volgt deze bewering uit het voorgaande, wat $\psi(\lambda)$ betreft uit de wijze, waarop wij in 8.6 de normaalvorm construeerden. (Wij krijgen immers eerst een positief aantal vierkanten met $\lambda_\rho$ op de hoofddiagonaal en lengte $m_\rho$, daarna een aantal met lengte $m_\rho - 1$ enz., dus de lengte van de grootste is $m_\rho$, d.i. de multipliciteit van $\lambda_\rho$ in $\psi(\lambda)$.)

## 8.8 Norm van vectoren

In dit caput zullen wij covectoren steeds door een accent van gewone (contravariante) vectoren onderscheiden.
Wij definiëren de volgende normen:
Voor contravariante vectoren (kolomvectoren):

(1) $\quad \|f\| \overset{def}{=} \max\limits_{i\in J} |f^i| \;, \quad J=\{1,\dots,r\}$.

Voor covectoren (covariante vectoren (rijvectoren)):

(2) $\quad \|g'\| \overset{def}{=} \sum^i |g_i|$.

Dan geldt:

(3) $\quad \forall_c^R \; \|cf\| = |c|.\|f\|$

(4) $\quad \|f+g\| \leqq \|f\|+\|g\|$ (driehoeksongelijkheid)

$\quad$ want $\|f+g\| = \max|f^i+g^i| \leqq \max(|f^i|+|g^i|) \leqq \|f\|+\|g\|$

(5) $\quad \forall_c^R \; \|cg'\| \leqq |c|.\|g'\|$

(6) $\quad \|g'+h'\| \leqq \|g'\|+\|h'\|$

(7) $\quad |g_i f^i| \leqq \|g'\|.\|f\|$.

Immers $|g_i f^i| = |\sum^i g_i f^i| \leqq \sum^i |g_i||f^i| \leqq \sum |g_i|\|f\| = \|g'\|\,\|f\|$.
Het gelijkheidsteken in (7) geldt als alle $f^i$ gelijk en alle $g_i$ positief zijn.

Deze normdefinities zijn een bijzonder geval van de algemene normen van Hölder, die luiden:

$$\|f\|_q \overset{def}{=} \left(\sum^i |f^i|^q\right)^{\frac{1}{q}} \text{ en } \|g'\|_p \overset{def}{=} \left(\sum^i |g_i|^p\right)^{\frac{1}{p}}.$$

Men kan nu bewijzen:

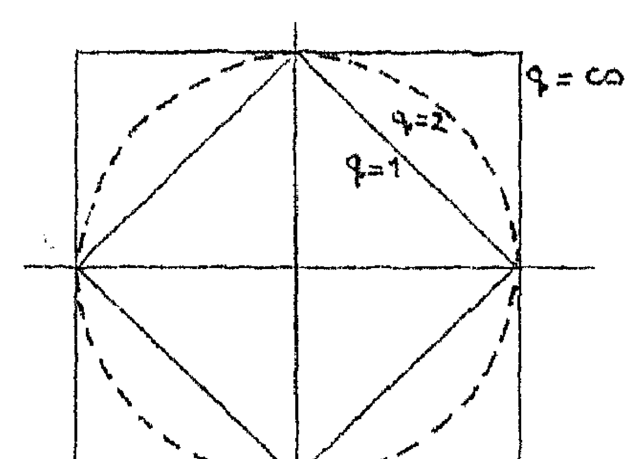
$|g_i f^i| \leqq \|g\|_p \|f\|_q$ mits $\frac{1}{p}+\frac{1}{q}=1$, de ongelijkheid van Hölder

Als $p=q=2$ krijgen wij de gewone Pythagoraanse normen ( $\|f\|_2$ is de lengte van de vector $f$ ). De ongelijkheid wordt dan die van Schwartz.
De door ons ingevoerde normen (1) en (2) zijn nu een bijzonder geval van de algemene normdefinitie, namelijk het geval $p=1$, $q=\infty$. Dat $p=1$ de definitie (2) oplevert, is direct duidelijk, dat (1) het geval $q=\infty$ is, zien wij als volgt:

Stel er zijn $k$ waarden van $i$, waarvoor $|f^i|=\|f\|$ en dus $r-k$ waarden van $i$, waarvoor $|f^i|<\|f\|$.
Voor deze $r-k$ $i$'s geldt: $|f^i|=\theta_i\|f\|$ met $0\leqq\theta_i<1$. Noem de grootste van deze $r-k$ $\theta_i$'s $\theta$, dan is dus $0\leqq\theta<1$ (Bedenk namelijk, dat $\mathfrak{I}$ eindig ondersteld is, dan is ook $k\geqq 1$ en is $r-k$ eindig).

$$\left(\sum^i |f^i|^q\right)^{\frac{1}{q}}=\left\{k\|f\|^q+\sum^i \theta_i^q\|f\|^q\right\}^{\frac{1}{q}}=\|f\|\,k^{\frac{1}{q}}\left\{1+\frac{\sum^i\theta_i^q}{k}\right\}^{\frac{1}{q}}.$$

Nu is $\dfrac{\sum^i\theta_i^q}{k}\leqq\dfrac{r-k}{k}\theta^q$, dus nadert tot $0$ als $q\to\infty$ $(0\leqq\theta<1)$.
Verder nadert $k^{\frac{1}{q}}$ tot $1$ als $q\to\infty$ zodat $\lim\limits_{q\to\infty}\left(\sum^i|f^i|^q\right)^{\frac{1}{q}}=\|f\|$, q.e.d.
De ongelijkheid (7) is dus een bijzonder geval van de ongelijkheid van Hölder.
Als $r=2$, dus als de vectoren twee componenten hebben, kunnen wij een eenvoudige meetkundige illustratie geven van de verschillende mogelijkheden voor $q$.
In onderstaande figuur is namelijk de meetkundige plaats getekend van alle vectoren met dezelfde norm voor verschillende waarden van $q$, waarbij het beginpunt der vectoren steeds in de oorsprong is.




Voor $q=1$ vinden wij het binnenste vierkant, voor $q=2$ de cirkel, voor $q=\infty$ het buitenste vierkant.

## 9. Gegeneraliseerde matrices

### 9.1 $\sigma$-velden en $\delta$-velden

Gegeven is een willekeurige abstracte vz $\mathfrak{I}$[1] en een stelsel deelvzn $\sigma_{\mathfrak{I}}$ van $\mathfrak{I}$ met de eigenschappen:

-------------

1) Zie voor de eenvoudigste eigenschappen van vzn en voor de notatie 1.2.

1. Als $X \in \sigma_{\mathfrak{J}}$ en $Y \in \sigma_{\mathfrak{J}}$, dan $X \cup Y \in \sigma_{\mathfrak{J}}$
2. Als $X \in \sigma_{\mathfrak{J}}$ en $Y \in \sigma_{\mathfrak{J}}$, dan $X -- Y \in \sigma_{\mathfrak{J}}$ en $Y -- X \in \sigma_{\mathfrak{J}}$ [1]

$3_\sigma$. Als $\forall_k^N X_k \in \sigma_{\mathfrak{J}}$, dan $\cup X_k \in \sigma_{\mathfrak{J}}$.

Een stelsel vzn, dat aan 1. en 2. voldoet, noemen wij een veld. De vereniging van ieder eindig aantal tot een veld behorende vzn behoort weer tot het veld. Dit volgt direct uit 1. Hetzelfde geldt voor de doorsnede van ieder eindig aantal tot het veld behorende vzn. Immers $X \cap Y = X -- (X -- Y)$, dus behoort tot het veld. Geldt ook $3_\sigma$, dan krijgen wij een $\sigma$-veld, ook wel Borel-veld of Borel-ring genaamd. De doorsnede van ieder aftelbaar aantal tot het $\sigma$-veld behorende vzn behoort weer tot het $\sigma$-veld. Immers stel $X = \cap X_n$, $\forall_n^N X_n \in \sigma_{\mathfrak{J}}$, dan is $X = \bigcup_{n=1}^{\infty} X_n -- \bigcup_{k=1}^{\infty}\left(\bigcup_{n=1}^{\infty} X_n -- X_k\right)$, dus $\in \sigma_{\mathfrak{J}}$. Ook geldt, dat $o \in \sigma_{\mathfrak{J}}$ [2] (volgt uit 2. als $X -- Y = o$, dus geldt reeds bij een veld).

Wij zullen voortaan de notatie $\sigma_{\mathfrak{J}}$ alleen gebruiken voor $\sigma$-velden van deelvzn van $\mathfrak{J}$, waarvoor geldt:

4. $\mathfrak{J} \in \sigma_{\mathfrak{J}}$,

dus voor $\sigma$-velden, die de hele vz bevatten.

Wij noemen een $\sigma$-veld <u>atomistisch</u> als bovendien voldaan is aan:

5. $\forall_x^{\mathfrak{J}} \{x\} \in \sigma_{\mathfrak{J}}$ [3]

Wij zullen ook $\delta$-velden beschouwen, dat zijn stelsels deelvzn van $\mathfrak{J}$ met de volgende eigenschappen:

1. Als $X \in \delta_{\mathfrak{J}}$ en $Y \in \delta_{\mathfrak{J}}$, dan $X \cup Y \in \delta_{\mathfrak{J}}$
2. Als $X \in \delta_{\mathfrak{J}}$ en $Y \in \delta_{\mathfrak{J}}$, dan $X -- Y \in \delta_{\mathfrak{J}}$ en $Y -- X \in \delta_{\mathfrak{J}}$

$3_\delta$. Als $\forall_k^N X_k \in \delta_{\mathfrak{J}}$, dan $\cap X_k \in \delta_{\mathfrak{J}}$.

Elk $\delta$-veld is een veld, elk $\sigma$-veld is een $\delta$-veld. Dus $3_\delta$ volgt uit 1,2 en $3_\sigma$.

---------

1) $A -- B \overset{def}{=} \mathcal{E}ns\{x \in A \mid x \notin B\}$. Meestal schrijft men hiervoor $A - B$ en spreekt van de verschilvz van $A$ en $B$. Wij schrijven een dubbel minteken, omdat wij aan $A - B$ een andere betekenis willen toekennen, namelijk $A - B \overset{def}{=} \mathcal{E}ns\{x - y \mid x \in A, y \in B\}$. Dit kan dus alleen als $\mathfrak{J}$ een additief geschreven groep is (b.v. de vz van alle reële getallen of van alle vectoren in een n-dimensionale ruimte).

2) Met $o$ (ronde nul) geven wij de lege vz aan. De lege vz is deelvz van elke vz (dan geldt namelijk zonder uitzondering, dat uit $A \cap B = C$ volgt, dat $C \subset A$ en $C \subset B$).

3) $\{x\}$ is de vz, alleen bestaande uit het element $x$.

Wanneer $\forall_n^N X_n \in \delta_{\mathfrak{J}}$ , dan zal $X = \overset{\infty}{\underset{0}{\cup}} X_n$ dan en slechts dan tot het $\delta$-veld behoren, wanneer $\exists_Y^{\delta_{\mathfrak{J}}} \forall_n^N X_n \subset Y$ is. In dat geval immers zal ook $X \subset Y$ zijn en gelden: $X = Y -- \overset{\infty}{\underset{1}{\cap}} (Y -- X_n)$, dus $\in \delta_{\mathfrak{J}}$, daar $Y \in \delta_{\mathfrak{J}}$.
De doorsnede [1] van een willekeurig aantal velden, resp. $\sigma$ - of $\delta$-velden is weer een veld, resp. $\sigma$- of $\delta$-veld. Bij elk willekeurig stelsel deelvzn van $\mathfrak{J}$ , $\lambda$ , bestaat een $\delta$-en een $\sigma$-veld, dat $\lambda$ omvat, waarmee wij bedoelen, dat elke vz van $\lambda$ tevens vz van het $\lambda$ omvattende $\delta$-, resp. $\sigma$-veld is. Immers het stelsel van alle deelvzn van $\mathfrak{J}$ omvat $\lambda$ en voldoet aan de eis. Het door $\lambda$ voortgebrachte $\delta$-, resp. $\sigma$-veld is het kleinste $\delta$-, resp. $\sigma$-veld, dat $\lambda$ omvat, d.w.z. is de doorsnede van alle $\delta$-, resp. $\sigma$-velden, die $\lambda$ omvatten. Wij schrijven hiervoor $\delta(\lambda)$, resp. $\sigma(\lambda)$.
Het door $\delta_{\mathfrak{J}}$ voortgebrachte $\sigma$-veld $\sigma(\delta_{\mathfrak{J}})$ krijgen we door aan $\delta_{\mathfrak{J}}$ toe te voegen alle vzn $X$, die geschreven kunnen worden in de vorm $X = \cup X_n$ met $\forall_n^N X_n \in \delta_{\mathfrak{J}}$ . Verifiëer, dat het op die manier uitgebreide $\delta$-veld inderdaad een $\sigma$-veld is.
Wij zullen voortaan de notatie $\delta_{\mathfrak{J}}$ alleen gebruiken voor $\delta$-velden van deelvzn van $\mathfrak{J}$ ,waarvoor geldt:

4. $\sigma(\delta_{\mathfrak{J}}) = \sigma_{\mathfrak{J}}$ ,

dus het door $\delta_{\mathfrak{J}}$ voortgebrachte $\sigma$-veld bevat de hele vz $\mathfrak{J}$ .
Gezien het bovenstaande is deze eis aequivalent met de eis:

4' $\mathfrak{J} = \cup \mathfrak{J}_n$ met $\forall_n^N \mathfrak{J}_n \in \delta_{\mathfrak{J}}$ , dus $\mathfrak{J}$ is de vereniging van een rij tot $\delta_{\mathfrak{J}}$ behorende vzn.
Door hierin $\mathfrak{J}_n$ te vervangen door $\overset{n}{\underset{1}{\cup}} \mathfrak{J}_k$ zien wij in, dat we zonder beperking kunnen onderstellen, dat $\forall_n^N \mathfrak{J}_n \subset \mathfrak{J}_{n+1}$.

Uit het bovenstaande (alinea onder $3_\delta$) is duidelijk, dat $\delta_{\mathfrak{J}}$ dan en slechts dan een $\sigma$-veld is als $\mathfrak{J} \in \delta_{\mathfrak{J}}$
Wij noemen een $\delta$-veld <u>atomistisch</u> als bovendien voldaan is aan:

5. $\forall_x^{\mathfrak{J}} \{x\} \in \delta_{\mathfrak{J}}$.

Onder een dissectie van een vz $X \in \delta_{\mathfrak{J}}$ verstaan wij een stelsel deelvzn $X_i$ van $X$ met de eigenschappen:

1. $\forall_i^N X_i \in \delta_{\mathfrak{J}}$.
2. $X_i \cap X_j = 0$ als $i \neq j$
3. $\cup X_i = X$.

-------------

1) Onder de doorsnede van twee stelsels vzn verstaan we het stelsel, bestaande uit die vzn, die zowel tot het ene als tot het andere stelsel behoren. N.B. De doorsnede van twee stelsels $\lambda_1$ en $\lambda_2$ van vzn is dus niet gelijk aan het stelsel van alle vzn $A \cap B$ met $A \in \lambda_1$ en $B \in \lambda_2$.

Een dergelijke dissectie geven we aan met $\{X_i\}$ . Het stelsel van alle dissecties van $X$ noemen we $\mathcal{D}(X)$. De eigenschappen 1,2 en 3 tezamen worden dus aangegeven met $\{X_i\} \in \mathcal{D}(X)$.

9.2 Functies van de eerste resp. tweede soort.

Een functie van de eerste soort is een reële fct[1] $f$ , gedefiniëerd voor alle $x \in \mathfrak{I}$ , die begrensd en $\delta$-meetbaar is op elke vz $X \in \delta_{\mathfrak{I}}$. Dat wil zeggen, dat geldt:

$$(1) \qquad \forall_X^{\delta_{\mathfrak{I}}} \sup_{x \in X} |f^x| < \infty$$

$$(2) \qquad \forall_a^R \forall_X^{\delta_{\mathfrak{I}}} \mathcal{E}ns\,\{x \in X \mid f^x \leqq a\} \in \delta_{\mathfrak{I}}$$

waarbij wij dus de waarde, die $f$ in $x$ aanneemt met $f^x$ aanduiden.

Een functie van de tweede soort is een reële fct $F$ , gedefiniëerd en aftelbaar additief[2] op $\delta_{\mathfrak{I}}$ . Dit is dus een verzamelingsfct, een wet, waardoor reële getallen worden toegevoegd aan die deelvzn van $\mathfrak{I}$ , die tot het gegeven $\delta$-veld $\delta_{\mathfrak{I}}$ behoren. Voor een dergelijke fct geldt als $X \in \delta_{\mathfrak{I}}$:

$$(3) \qquad \forall_{\{X_i\}}^{\mathcal{D}(X)} F_X = \sum^i F_{X_i}.$$

Opmerkingen: In caput II van [15] is de notatie $f_x$ , resp. $F^X$ gebruikt in plaats van $f^x$, resp. $F_X$. In de literatuur schrijft men meestal $f(x)$ en $F(X)$ . De voordelen van onze notatie zullen in de volgende paragrafen wel blijken.

Meestal worden aftelbaar additieve vz-fcts gedefiniëerd op $\sigma$-velden. Door ze op $\delta$-velden te definiëren kunnen we functiewaarden $\pm\infty$ vermijden, namelijk door die vzn van het $\sigma$-veld weg te laten, waarop de fct oneindig wordt. Leggen we bovendien aan de fcts van de eerste soort de begrensdheidsvoorwaarde (1) op, dan zijn enkele stellingen zonder beperking geldig (b.v. de in 4. van 9.3 te noemen stelling). Een nadeel is echter, dat bij een andere vz-fct wellicht andere vzn van het $\sigma$-veld weggelaten moeten worden, waardoor deze fct gedefiniëerd wordt op een ander $\delta$-veld. Beide fcts zijn dan echter in ieder geval gedefiniëerd op de doorsnede van hun $\delta$-velden. Zouden we alleen begrensde fcts van de eerste en tweede soort beschouwen, dan konden we gerust met $\sigma$-velden werken. Dit is echter met het

-------------

1) Het is niet van belang, dat een fct van de eerste soort reële getallen toevoegt aan de elementen van $\mathfrak{I}$ .Dit kunnen ook elementen van een andere vz $\Omega$ zijn. In deze vz $\Omega$ moet echter ook een $\delta$-veld gegeven zijn. Inplaats van (2) krijgen we nu:

$$(2') \qquad \forall_S^{\delta_\Omega} \forall_X^{\delta_{\mathfrak{I}}} \mathcal{E}ns\,\{x \in X \mid f^x \in S\} \in \delta_{\mathfrak{I}}$$

dus $f$ beeldt de meetbare vzn van $\Omega$ en $\mathfrak{I}$ op elkaar af.

2) Men spreekt ook wel van absoluut, totaal, volledig of $\sigma$-additief.

oog op de toepassing van de te geven theorie in hoofdstuk 6 (speciaal 6.4) onmogelijk.

We voeren de volgende normen in:

(4) $$\|f\|^X \overset{def}{=} \sup_{x \in X} |f^x| \ , \ \|f\| \overset{def}{=} \|f\|^{\mathfrak{J}} .$$

Het supremum behoeft hier niet te worden bereikt. $\|f\|^X$ is geen fct van de tweede soort, is namelijk behalve in triviale gevallen niet additief.

(5) $$\|F\|_X \overset{def}{=} |F|_X \overset{def}{=} \sup_{\{X_i\} \in \mathcal{D}(X)} \sum^i |F_{X_i}| \ , \ \|F\| \overset{def}{=} \|F\|_{\mathfrak{J}} .$$

Het supremum wordt hier bereikt, is dus een maximum.[1] Ook $|F|_X$ is aftelbaar additief, dus een fct. van de tweede soort op $\delta_{\mathfrak{J}}$. Het is bij de normdefinities niet nodig, dat $X \in \delta_{\mathfrak{J}}$. Wel moet bij de tweede definitie er minstens een dissectie $\{X_i\}$ van $X$ zijn met $X_i \in \delta_{\mathfrak{J}}$, d.w.z. $X \in \sigma(\delta_{\mathfrak{J}}) = \sigma_{\mathfrak{J}}$

Dus geldt:

(6) $$\forall_X^{\delta_{\mathfrak{J}}} \ |F|_X < \infty \ , \ \|f\|^X < \infty .$$

Verder voor alle $X \subset \mathfrak{J}$ (met alleen gezien (2) de beperking, dat de norm gedefiniëerd kan worden):

(7) $$\begin{cases} 0 \leq \|f\|^X \leq \|f\| \\ \forall_\lambda^R \ \|\lambda f\|^X = |\lambda| . \|f\|^X \\ \|f_1 + f_2\|^X \leq \|f_1\|^X + \|f_2\|^X \end{cases}$$

(8) $$\begin{cases} 0 \leq |F|_X \leq \|F\| \\ \forall_\lambda^R \ \|\lambda F\|_X = |\lambda| . \|F\|_X \\ \|F_1 + F_2\|_X \leq \|F_1\|_X + \|F_2\|_X . \end{cases}$$

Tenslotte:

(9) $$\forall_x^{\delta_{\mathfrak{J}}} |F_X| \leq |F|_X .$$

Het geval kan zich voordoen, dat $|F|_X$ en $\|f\|^X$, dus ook $F_X$ en $f^x$ begrensd zijn op $\delta_{\mathfrak{J}}$. Dan zijn zij ook begrensd op $\sigma(\delta_{\mathfrak{J}}) = \sigma_{\mathfrak{J}}$. Het definitiegebied van $F$ kunnen we dan uitbreiden tot $\sigma_{\mathfrak{J}}$ door als $X \in \sigma_{\mathfrak{J}}$ en $X \notin \delta_{\mathfrak{J}}$ te definiëren $F_X = \sum_1^{\infty i} F_{X_i}$ als $\{X_i\} \in \mathcal{D}(X)$ is.

De fct blijft hierdoor met begrensde fcts te maken hebben, zullen we aannemen, dat het definitiegebied is uitgebreid.

-------------

1) Voorzover geen bewijzen gegeven worden, zijn die te vinden in [15] ,caput II.

Van een fct. van de eerste soort en een van de tweede soort kunnen we een integraal vormen:

$$(11)\qquad (Ff)_X \stackrel{\text{def}}{=} \int_X F_{dx} f^x \stackrel{\text{def}}{=}$$
$$\stackrel{\text{def}}{=} \lim_{\varepsilon \downarrow 0} \left\{ \Sigma^i F_{X_i} f^{x_i} \,\middle|\, \{X_i\} \in \mathcal{D}(X),\ \forall_i^N \left(\sup_{x\in X_i} f^x - \inf_{x \in X_i} f^x\right) \leq \varepsilon,\ \forall_i^N x_i \in X_i \right\}$$

De definitie is alleen maar mogelijk als $X \in \sigma_\mathfrak{J}$, daar natuurlijk $\forall_i\ X_i \in \delta_\mathfrak{J}$.
De limiet bestaat in ieder geval en is onafhankelijk van de keuze der dissecties en der punten $x_i \in X_i$ als $X \in \delta_\mathfrak{J}$. ($\|f\|^X < \infty$).
Verder bestaat de integraal voor $X \in \sigma_\mathfrak{J}$ als $f$ en $F$ begrensd zijn op $X$, dus als $|F|_X$ en $\|f\|^X < \infty$ zijn, òf als er een dissectie $\{X_i^{(0)}\} \in \mathcal{D}(X)$ te vinden is, zodanig dat $\Sigma\, |F|_{X_i^{(0)}}\, |f^{x_i^{(0)}}| < \infty$ $(x_i^{(0)} \in X_i^{(0)})$.
Zie [15], blz. 55 en 56, waar tevens bewezen wordt, dat $(Ff)_X$ weer een fct. van de tweede soort is. Als $(Ff)_\mathfrak{J}$ bestaat, schrijven we daarvoor ook wel $(Ff)$.
Wanneer wij de variatie van $f$ op een vz $X$ definiëren als:

$var_X f \stackrel{\text{def}}{=} \sup_{x\in X} f^x - \inf_{x \in X} f^x$ en verder invoeren $var_{\{X_n\}} f \stackrel{\text{def}}{=} \sup_n (var_{X_n} f)$

dan geldt als $\{X_n\} \in \mathcal{D}(X)$ en $\forall_n\ x_n \in X_n$ de ongelijkheid:

$$(12)\qquad \left|(Ff)_X - \Sigma^n F_{X_n} f^{x_n}\right| \leq |F|_X\, var_{\{X_n\}} f.$$

Dit zien wij in door te definiëren: $g^x \stackrel{\text{def}}{=} f^{x_n}$ als $x \in X_n$.
Dan is $\Sigma^n F_{X_n} f^{x_n} = (Fg)_X$ en

$$|(Ff)_X - (Fg)_X| = |(F(f-g))_X| \leq |F|_X \cdot \|f-g\|^X \leq |F|_X\, var_{\{X_n\}} f.$$

Ook geldt:

$$(13)\qquad \forall_X^{\sigma_\mathfrak{J}}\ |(Ff)|_X \leq |F|_X \cdot \|f\|^X$$

Immers $\quad |(Ff)|_X = \sup_{\{X_i\}\in\mathcal{D}(X)} \Sigma^i |(Ff)_{X_i}| \leq \sup_{\{X_i\}\in\mathcal{D}(X)} \Sigma^i |F|_{X_i} \|f\|^X = |F|_X \cdot \|f\|^X$,

daar $\quad |(Ff)_{X_i}| = \left|\int_{X_i} F_{dx} f^x\right| = \lim_{\varepsilon\downarrow 0} \left\{ |\Sigma\, F_{X_{ij}} f^{x_{ij}}| \,\middle|\, \{X_{ij}\} \in \mathcal{D}(X_i) \right\} \leq$
$\leq \lim_{\varepsilon \downarrow 0} \Sigma^i |F_{X_{ij}} f^{x_{ij}}| \leq \|f\|^{X_i} \Sigma^i |F_{X_{ij}}| \leq |F|_{X_i} \|f\|^X.$

Is $\mathfrak{J}$ aftelbaar en $\delta_\mathfrak{J}$ atomistisch, dan is dus iedere $X \subset \mathfrak{J}$ aftelbaar en geldt:

$$F_X = \sum_{x\in X} F_x\ ;\ (Ff)_X = \int_X F_{dx} f^x = \sum_{x \in X} F_x f^x\ ;\ (Ff) = \sum_{x\in\mathfrak{J}} F_x f^x\ ;$$
$\|f\|^X = \sup_{x\in X} |f^x|$ (bij eindig aantal de absoluut grootste $f$),

alsmede $|F|_x = \sum_{x \in X} |F_x|$.

De ongelijkheid (13) wordt nu:

$$\Big|\sum_{x \in X} F_x f^x\Big| \leqq \sum_{x \in X} |F_x| \cdot \sup_{x \in X} |f^x| .$$

Wij krijgen dus precies het in 8.8 beschouwde geval $p = \infty$, $q = 1$ van de algemene normdefinitie van Hölder. De ongelijkheid (13) is een generalisatie van de ongelijkheid van Hölder in het speciale geval $p = \infty$, $q = 1$ (Zie 8.8).

De andere mogelijkheden voor $p$ en $q$ (zoals $p = 2, q = 2$, de Pythagoraanse normen) zijn niet eenvoudig te generaliseren. Meestal treedt dan geen convergentie op of is de uitkomst altijd 0.

Tevens blijkt uit het voorgaande, dat de integraal $(Ff)$ een generalisatie is van het vectorproduct (product van covariante en contravariante vector).

## 9.3 Functies van tweeërlei soort

Gegeven zijn twee willekeurige vzn $\mathfrak{J}$ en $\mathfrak{J}'$ met bijbehorende $\delta$-velden $\delta_{\mathfrak{J}}$ en $\delta_{\mathfrak{J}'}$. Wij beschouwen nu een fct. van tweeërlei soort $M_y^x$ met $x \in \mathfrak{J}$ en $\in \delta_{\mathfrak{J}}$

Hiervoor moet gelden:

1. Voor iedere vaste vz $y \in \delta_{\mathfrak{J}}$ is $M_y^x$ een fct. van de eerste soort van $x$ op $\mathfrak{J}$, dus:

(1) $$\forall_y^{\delta_{\mathfrak{J}}} \forall_\lambda^R \forall_X^{\delta_{\mathfrak{J}}} \mathcal{E}ns \{x \in X \mid M_y^x \leqq \lambda\} \in \delta_{\mathfrak{J}} ,$$

(2) $$\forall_y^{\delta_{\mathfrak{J}'}} \forall_X^{\delta_{\mathfrak{J}}} \sup_{x \in X} |M_y^x| < \infty$$

2. Voor iedere vaste $x \in \mathfrak{J}$ is $M_y^x$ een fct. van de tweede soort van $y$ op $\delta_{\mathfrak{J}'}$, dus:

(3) $$\forall_x^{\mathfrak{J}} \forall_{\{y_i\}}^{\mathcal{D}(y)} M_y^x = \sum^i M_{y_i}^x .$$

Een dergelijke fct. is te beschouwen als een gegeneraliseerde matrix en noemen wij een $(\mathfrak{J}, \mathfrak{J}')$ matrix.

Als namelijk $\mathfrak{J}$ en $\mathfrak{J}'$ aftelbaar zijn, dan is $M_Y^x = \sum_{y \in Y} M_y^x$ en wordt dus $M_Y^x$ bepaald door de gewone eindige of aftelbaar oneindige matrix met elementen $M_y^x$

Als norm definiëren wij:

(4) $$\|M\|_Y^X \overset{def}{=} \sup_{x \in X} \|M^x\|_Y \overset{def}{=} \sup_{x \in X} \sup_{\{y_i\} \in \mathcal{D}(Y)} \sum^i |M_{y_i}^x| \text{ en } \|M\| \overset{de}{=}$$

Uit (2) volgt, dat als $X \in \delta_{\mathcal{Y}}$ en $Y \in \delta_{\mathcal{Y}'}$ geldt:

(5) $$\|M\|_Y^X < \infty .$$

Als $\|M\| < \infty$ wordt de matrix $M$ begrensd genoemd. In dat geval kan $M_Y^x$ voor alle $Y \in \sigma_{\mathcal{Y}'}$ worden gedefiniëerd.
Zij nu gegeven een $(\mathcal{Y},\mathcal{Y}')$ matrix $M_Y^x$, een fct. van de eerste soort $f^y$ op $\mathcal{Y}'$ en een fct. van de tweede soort $F_X$ op $\delta_{\mathcal{Y}}$, dan definiëren wij:

(6) $$g_Y^x \overset{\text{def}}{=} \int_Y M_{dy}^x f^y \qquad \text{(Bestaalt altijd als } Y \in \delta_{\mathcal{Y}'}\text{)}$$

(7) $$G_{X,Y} \overset{\text{def}}{=} \int_X F_{dx} M_Y^x \qquad \text{(Bestaat altijd als } X \in \delta_{\mathcal{Y}} \text{ en } Y \in \delta_{\mathcal{Y}'}\text{)}$$

Er geldt:

1. $g_Y^x$ is bij vaste $x \in \mathcal{Y}$ aftelbaar additief op $\delta_{\mathcal{Y}'}$.

2. $G_{X,Y}$ is bij vaste $Y \in \delta_{\mathcal{Y}'}$ aftelbaar additief op $\delta_{\mathcal{Y}}$

3. $g_Y^x$ is bij vaste $Y \in \delta_{\mathcal{Y}'}$ een fct. van de eerste soort van $x$ op $\mathcal{Y}$.

Immers $g_Y^x = \lim\limits_{\{Y_i\} \in \mathcal{D}(Y)} \sum^i M_{Y_i}^x f^{y_i}$ . Nu is $M_{Y_i}^x$ bij vaste $Y_i \in \delta_{\mathcal{Y}'}$ een meetbare fct. van $x$, dus ook $M_{Y_i}^x f^{y_i}$ (vermenigvuldiging met constant getal), dus ook $\sum^i M_{Y_i}^x f^{y_i}$, dus ook $g_Y^x$.
Hierbij hebben wij gebruik gemaakt van lemma 1, blz.72 van [15].
Uit (6),(4) en (9.2.12) volgt verder:

$$\forall_x^{\mathcal{Y}} \|g^x\|_Y \leqq \|M^x\|_Y . \|f\|^Y \leqq \|M\|_Y^X \|f\|^Y \quad \text{als } X \in \delta_{\mathcal{Y}} \text{ en } x \in X,$$

dus $$\forall_X^{\delta_{\mathcal{Y}}} \forall_Y^{\delta_{\mathcal{Y}'}} \|g\|_Y^X \leqq \|M\|_Y^X . \|f\|^Y < \infty , \text{ q.e.d.}$$

Uit 1. en 3 volgt, dat $g_Y^x$ weer een $(\mathcal{Y},\mathcal{Y}')$ matrix is.

4. $G_{X,Y}$ is bij vaste $X \in \delta_{\mathcal{Y}}$ een aftelbaar additieve fct. van $Y$ op $\delta_{\mathcal{Y}'}$, dus een fct. van de tweede soort zowel op $\delta_{\mathcal{Y}}$ als op $\delta_{\mathcal{Y}'}$.

Bewijs: Duidelijk is, dat $G_{X,Y}$ additief is in $Y$ bij vaste $X \in \delta_{\mathcal{Y}}$. De aftelbare additiviteit is dus aangetoond, wanneer wij kunnen bewijzen, dat $\lim\limits_{n\to\infty} G_{X,Y_n} = 0$ voor elke rij $Y_n$, waarvoor geldt:

$$\forall_n\ Y_n \in \delta_{\mathcal{Y}'},\ Y_n \supset Y_{n+1} \text{ en } \bigcap_1^\infty Y_n = 0.$$ 1)

Wij beschouwen dus een willekeurige rij $Y_n$ met genoemde eigenschappen.
Kies $\varepsilon > 0$ en definieer:

$$B_n \overset{\text{def}}{=} \mathcal{E}ns\{x \in X \mid |M^x|_{Y_n} \geqq \varepsilon\}$$

$B_n \in \delta_{\mathcal{Y}}$, daar $M^x_{Y_n}$ en dus ook $|M^x|_{Y_n}$ meetbaar is voor iedere $Y_n$. $|M^x|$ is aftelbaar additief en $\geqq 0$, dus $|M^x|_A \leqq |M^x|_B$ als $A \subset B$. Hieruit en uit $Y_n \supset Y_{n+1}$ volgt: $B_{n+1} \subset B_n$. Bovendien zal $\cap B_n = 0$ zijn, want stel $x \in \cap B_n$, dan zal voor deze $x$ gelden, dat $\forall_n |M^x|_{Y_n} \geqq \varepsilon$ is, hetgeen in strijd is met de aftelbare additiviteit van $|M^x|$ (zie voetnoot). Daar ook $|F|$ aftelbaar additief is zal $\lim_{n\to\infty} |F|_{B_n} = 0$, dus voor voldoend grote n $|F|_{B_n} \leqq \varepsilon$.

Als $C_n \overset{\text{def}}{=} X -- B_n$, dan zal

$$\mathcal{G}_{X,Y_n} = \mathcal{G}_{B_n,Y_n} + \mathcal{G}_{C_n,Y_n}.$$

Maar

$$\left|\mathcal{G}_{B_n,Y_n}\right| \leqq \int_{B_n} |F|_{dx}\, |M^x|_{Y_n} \leqq |F|_{B_n} . \|M\|^{B_n}_{Y_n} \leqq \varepsilon C \ \text{als}\ C \overset{\text{def}}{=} \|M\|^{B_n}_{Y_n} < \infty \ (B_n \in \delta_{\mathcal{Y}}).$$

$$\left|\mathcal{G}_{C_n,Y_n}\right| \leqq \int_{C_n} |F|_{dx}\, |M^x|_{Y_n} \leqq |F|_{C_n} \|M\|^{C_n}_{Y_n} \leqq \varepsilon C' \ \text{als}\ C' \overset{\text{def}}{=} |F|_X \ \text{daar}\ C_n \subset X.$$

en $\|M\|^{C_n}_{Y_n} < \varepsilon$ op grond van de definitie van $B_n$ en $C_n$.

Dus $\left|\mathcal{G}_{X,Y_n}\right| \leqq \varepsilon(C + C')$ als n maar voldoende groot is, dus $\lim_{n\to\infty} \mathcal{G}_{X,Y_n} = 0$, q.e.d.

5. Wanneer $\mathcal{G}_Y \overset{\text{def}}{=} \mathcal{G}_{\mathcal{Y},Y} = \int_{\mathcal{Y}} F_{dx}\, M^x_Y$ bestaat, dan is een voldoende voorwaarde voor de aftelbare additiviteit van $\mathcal{G}_Y$, dat

$$\int_{\mathcal{Y}} |F|_{dx}\, |M^x|_Y < \infty.$$

---

1) Immers laat $F_X$ additief zijn op een $\delta$-veld $\delta_{\mathcal{Y}}$ en laat $\lim_{n\to\infty} F_{X_n} = 0$ voor elke rij $X_n$, waarvoor geldt $\forall_n X_n \in \delta_{\mathcal{Y}}, X_n \supset X_{n+1}$ en $\cap X_n = 0$, dan is F aftelbaar additief op $\delta_{\mathcal{Y}}$. Stel namelijk $X = \cup X_n$ met $\forall_n X_n \in \delta_{\mathcal{Y}}, X \in \delta_{\mathcal{Y}}$ en $\{X_n\} \in \mathcal{D}(X)$. Definieer dan $X'_n = \bigcup_{i=1}^n X_i$ en $X''_n = X -- X'_n$. Dan zal $\forall_n X'_n \in \delta_{\mathcal{Y}}, X''_n \in \delta_{\mathcal{Y}}, \cup X'_n = X$, dus $\cap X''_n = 0$, verder $\forall_n X'_n \subset X'_{n+1}$, dus $X''_n \supset X''_{n+1}$, dus $\lim F_{X''_n} = 0$.
Dus $\lim F_{X -- X_n} = 0$, $F_X = \lim_{n\to\infty} F_{X'_n} = \lim_{n\to\infty} \sum_1^n F_{X_i} = \sum_1^\infty F_{X_i}$, q.e.d.
(Tweemaal van de eindige additiviteit gebruik gemaakt). Stel omgekeerd F aftelbaar additief en beschouw een willekeurige rij $X_n$ met eigenschappen $\forall_n X_n \in \delta_{\mathcal{Y}}, X_n \supset X_{n+1}$ en $\cap X_n = 0$. Definieer $X'_n = X_1 -- X_n$ en $X''_n = X'_n -- X'_{n-1}$ $(X''_1 = X'_1)$, dan $\forall_n X'_n \in \delta_{\mathcal{Y}}, X''_n \in \delta_{\mathcal{Y}}, X'_n \subset X'_{n+1}$, $\cup X''_n = \cup X'_n = X_1$, $X''_i \cap X''_j = 0$ als $i \neq j$, dus $F_{X_1} = F_{X''_n} = \sum(F_{X'_n} - F_{X'_{n-1}}) = \lim_{n\to\infty} F_{X'_n} = F_{X_1} - \lim_{n\to\infty} F_{X_n}$, dus $\lim_{n\to\infty} F_{X_n} = 0$.

Immers stel $\{y_i\} \in \mathcal{D}(y)$ . Uit 9.1 volgt, dat er een dissectie $\{\mathcal{J}_n\} \in \mathcal{D}(\mathcal{J})$ is, waarbij $\forall_n\ \mathcal{J}_n \in \delta_{\mathcal{J}}$ . De voorwaarde $\int_{\mathcal{J}} |F|_{dx} |M^x|_y < \infty$ maakt het mogelijk ons te beperken tot niet negatieve $F$ en $M$. Dan is ook $G$ niet negatief en geldt:

$$G_y = G_{\mathcal{J},y} = \sum_1^{\infty n} G_{\mathcal{J}_n,y} = \sum_1^{\infty n} \sum_1^{\infty i} G_{\mathcal{J}_n,y_i} = \sum_1^{\infty i} \sum_1^{\infty n} G_{\mathcal{J}_n,y_i} = \sum_1^{\infty i} G_{\mathcal{J},y_i}, \text{ q.e.d.}$$

Hierbij is gebruik gemaakt van de aftelbare additiviteit van $G_{\mathcal{J}_n,y}$ $(\mathcal{J}_n \in \delta_{\mathcal{J}})$.

6. Ook is onmiddellijk in te zien, dat, wanneer

$g_x \overset{\text{def}}{=} g^x_{\mathcal{J}'} = \int_{\mathcal{J}'} M^x_{dy} f^y$ bestaat voor alle $x \in \mathcal{J}$ , dat dan $g^x$ ook meetbaar is.

Beschouw immers een dissectie $\{\mathcal{J}_n'\} \in \mathcal{D}(\mathcal{J}')$ . Dan is $g^x_{\mathcal{J}_n'}$ meetbaar volgens 3, dus $g^x$ ook, daar $g^x = \sum_1^{\infty n} g^x_{\mathcal{J}_n'}$ en de som van meetbare fcts. weer meetbaar is.

Wanneer bovendien $g^x = (Mf)^x$ begrensd is op iedere $X \in \delta_{\mathcal{J}}$ , dan is $g^x$ eveneens een fct. van de eerste soort. Dit is bijvoorbeeld zo als $M$ en $f$ begrensd zijn.

Uit 5. en 6. volgt als bijzonder geval nu direct de

Stelling:

Een begrensde $(\mathcal{J},\mathcal{J}')$ matrix $M$ voert de begrensde fcts. $f$ van de eerste soort op $\mathcal{J}'$ over in begrensde fcts. van de eerste soort op $\mathcal{J}$ en de begrensde fcts. van de tweede soort op $\sigma_{\mathcal{J}}$ in begrensde fcts. van de tweede soort op $\sigma_{\mathcal{J}'}$ , waarbij bovendien geldt:

(9) $\quad \|(Mf)\| \leqq \|M\| . \|f\|,$

(10) $\quad |(FM)| \leqq |F| . \|M\| .$

Wanneer $\mathcal{J}, \mathcal{J}'$ en $\mathcal{J}''$ willekeurige vzn zijn met $\delta$-velden $\delta_{\mathcal{J}}, \delta_{\mathcal{J}'}$ en $\delta_{\mathcal{J}''}$ , wanneer verder $M$ een $(\mathcal{J},\mathcal{J}')$ en $N$ een $(\mathcal{J}',\mathcal{J}'')$ matrix is, dan definiëren wij:

(11) $\quad (MN)^x_{Y,z} = \int_Y M^x_{dy} N^y_z .$

Deze integraal bestaat voor alle $x \in \mathcal{J}$, $Y \in \delta_{\mathcal{J}'}$, $Z \in \delta_{\mathcal{J}''}$ en bepaalt voor iedere vaste $Y \in \delta_{\mathcal{J}'}$ een $(\mathcal{J},\mathcal{J}'')$ matrix.

Als $M$ en $N$ begrensd zijn, kunnen wij $Y = \mathcal{J}'$ nemen. Wij krijgen dan een gegeneraliseerd matrixproduct:

(12) $\quad (MN)^x_z = \int_{\mathcal{J}'} M^x_{dy} N^y_z .$

Wanneer $M$ en $N$ niet begrensd is en de integraal in (12) toch bestaat, dan zal volgens 5. en 6. het linkerlid van (12) eveneens een matrix zijn, dus voldoen aan (1),(2) en (3), wanneer

(13) $\quad \int_{\mathcal{Y}'} |M^x|_{dy} |N^y|_z$ bestaat voor alle $x \in \mathcal{Y}$ en $Z \in \delta_{\mathcal{Y}''}$ en begrensd is als functie van $x$ op iedere $X \in \delta_{\mathcal{Y}}$ bij vaste $Z \in \delta_{\mathcal{Y}''}$.

Alleen wanneer het linkerlid van (12) weer een matrix is, dus bijvoorbeeld wanneer $M$ en $N$ begrensd zijn of algemener wanneer zij aan (13) voldoen, is het zinvol van een gegeneraliseerde matrixvermenigvuldiging te spreken. Bovendien zal deze matrixvermenigvuldiging associatief moeten zijn (hetgeen, zoals in 9.6 zal worden aangetoond, onder zeer algemene voorwaarden inderdaad het geval is), terwijl de gewone matrixvermenigvuldiging een bijzonder geval van de hier ingevoerde gegeneraliseerde vermenigvuldiging moet zijn.

Dit laatste is inderdaad het geval. Laat $\mathcal{Y}, \mathcal{Y}'$ en $\mathcal{Y}''$ immers aftelbaar zijn. Dan is

$$(MN)^x_Z = \sum_{x \in Z} (MN)^x_z \quad en \quad (MN)^x_z = \int M^x_{dy} N^y_z = \sum_{y \in \mathcal{Y}} M^x_y N^y_z ,$$

precies de gewone matrixvermenigvuldiging.

Tenslotte merken wij op, dat uit (12) volgt, dat voor de normen de volgende ongelijkheid geldt:

(14) $\quad \|MN\| \leqq \|M\| . \|N\| .$

## 9.4 Eenheidsmatrices

Wij nemen aan, dat de vzn $\mathcal{Y}$ en $\mathcal{Y}'$ en tegelijkertijd de $\delta$-velden $\delta_{\mathcal{Y}}$ en $\delta_{\mathcal{Y}'}$ samenvallen.

(1) $\quad /^x_X \overset{def}{=} \begin{cases} 1 & als \ x \in X \\ 0 & als \ x \notin X \end{cases}$

Hiervoor geldt:

1. $/_X$ is een fct. van de eerste soort.

Want $\quad \mathcal{E}ns\{x \mid /^x_X \leqq \lambda\} = \begin{cases} \mathcal{Y} & als \ \lambda \geqq 1 \\ \mathcal{Y} -- X & als \ 0 \leqq \lambda < 1 \\ 0 & als \ \lambda < 0 \end{cases}$ , dus $/_X$ is meetbaar

$/_X$ is de karakteristieke fct van $X$ , d.w.z. de fct. die =1 is, als $x \in X$ en anders =0.

2. $/^x$ is een fct. van de tweede soort en geeft de totale massa in $X$ aan (0 als $x \notin X$ , 1 als $x \in X$).

Immers neem $\{X_i\} \in \mathcal{D}(X)$.

Als $x \in X$, dan behoort $x$ tot precies één $X_i$. Voor die $i$ is $/^x_{X_i} = 1$, voor de andere 0, dus $\sum^i /^x_{X_i} = 1 = /^x_X$.
Als $x \notin X$ is $\forall_i\, /^x_{X_i} = 0$, dus ook $\sum^i /^x_{X_i} = 0 = /^x_X$.

3. $$(M/)^x_X = \int M^x_{dy} /^y_X = \int_X M^x_{dy} /^y_X + \int_{\mathfrak{J}--X} M^x_{dy} /^y_X =$$
$$= \lim\left(\sum M^x_{Y_i} /^{y_i}_X \,\middle|\, \{Y_i\} \in \mathcal{D}(X), y_i \in Y_i\right) + \lim\left\{\sum M^x_{Y_i} /^{y_i}_X \,\middle|\, \{Y_i\} \in \mathcal{D}(\mathfrak{J}--X), y_i \in Y_i\right\} =$$
$$= \lim\left(\sum M^x_{Y_i} \,\middle|\, \{Y_i\} \in \mathcal{D}(X)\right) + 0 = M^x_X.$$

Dus:

(2) $(M/)^x_X = M^x_X$, $M/ = M$ voor alle $M$.

De matrix $/$ is dus een rechtereenheidsmatrix.

4. $$(/M)^x_X = \int /^x_{dy} M^y = \lim\left(\sum /^x_{Y_i} M^{y_i}_X \,\middle|\, \{Y_i\} \in \mathcal{D}(X), y_i \in Y_i\right) =$$
$$= \lim_{\varepsilon \downarrow 0} \left\{ M^{y_i}_X \,\middle|\, y_i \in Y_i, x \in Y_i \right\} = M^x_X, \text{ daar } |M^{y_i}_X - M^x_X| \leq \varepsilon.$$

Dus:

(3) $/M = M$ voor alle $M$.

$/$ is dus ook een linkereenheidsmatrix, dus er is slechts één rechter en slechts één linker eenheidsmatrix, d.w.z. $/$ is de eenheidsmatrix (Want is b.v. $E$ een linker eenheidsmatrix, d.w.z. $EM = M$ voor alle $M$, dan is ook $E/ = /$. Maar $E/ = E$ daar $/$ rechter eenheidsmatrix is, dus $E = /$ ).

5. Als $\mathfrak{J}$ aftelbaar, dan is $/^x_X = \sum_{y \in X} /^x_y$ met $/^x_y = \begin{cases} 1 \text{ als } x = y \\ 0 \text{ als } x \neq y \end{cases}$,

dus dan is $/$ de matrix:

$$\begin{pmatrix} 1 & & & 0 \\ & 1 & & \\ & & 1 & \\ 0 & & & \ddots \end{pmatrix}$$

6. $$(/f)^x = f^x$$
$$(F/)_X = F_X$$

7. $(f/)^x_X \stackrel{\text{def}}{=} f^x /^x_X$, is dus een matrix (vergelijk 6.: $/f$ is een fct., $f/$ een matrix!).

Als $\mathfrak{J}$ aftelbaar is, krijgen wij een diagonaalmatrix:

$$\begin{pmatrix} f^1 & & & 0 \\ & f^2 & & \\ & & f^3 & \\ 0 & & & \ddots \end{pmatrix}$$ , dus definitie 7. geeft een generalisatie van

de diagonaalmatrices. Deze hebben de prettige eigenschap, dat het product van twee zulke matrices commutatief is (zie 8.1). Dit blijft gelden:

$$((f/)(g/))^x_X = \int f^x/^x_{dy}\, g^y/^y_X = f^x g^x/^x_X = ((g/)(f/))^x_X, \text{ dus}$$
$$f/.g/ = g/.f/.$$

8. $\quad /^x_X \,.\, /^x_Y = /^x_{X\cap Y}$.

Dikwijls is $Y$ gegeven ($Y = A$). Dan definiëren wij $(/_A)^x_X \overset{def}{=} /^x_{A\cap X}$. In het aftelbare geval wordt dit een matrix, waarbij buiten de hoofddiagonaal alleen nullen staan en op de hoofddiagonaal 1 als $x\in A$ is, anders 0.
Ook geldt:

(4) $\quad /_A /_B = /_{A\cap B}$

(5) $\quad (/_A f)^x = \begin{cases} f^x & \text{als } x\in A \\ 0 & \text{als } x\notin A \end{cases}$, dus $/_A$ knot de fct. $f$ af.

(6) $\quad (F/_A)_X = F_{A\cap X}$ dus $F/_A$ heeft in $X$ evenveel "lading" als $F$ in $A\cap X$ ; $/_A$ annuleert alle lading buiten $A$

9. $\quad \int_X F_{dx} f^x = \int_{\mathcal{J}} F_{dx} f^x /^x_X = (Ff/)_X$.

Wij kunnen dus in plaats van over de deelvz $X$ van $\mathcal{J}$ te integreren ook over geheel $\mathcal{J}$ integreren, waarbij wij dán als integrand de matrix $f/_X$ gebruiken in plaats van de fct. $f$ van de eerste soort.

## 9.5 Omkeerbaarheid van de integratievolgorde

Stelling 1: Als $f$ en $F$ fcts. van de eerste resp. tweede soort zijn, respectievelijk op $\mathcal{J}$ en $\delta_{\mathcal{J}}$ en als $M$ een $(\mathcal{J}', \mathcal{J})$ matrix is, dan geldt:

(1) $$\forall^{\delta_{\mathcal{J}}}_X \forall^{\delta_{\mathcal{J}'}}_Y \int_X \left\{ \int_Y F_{dy} M^y_{dx} \right\} f^x = \int_Y F_{dy} \int_X M^y_{dx} f^x,$$

dus de integratievolgorde mag omgekeerd worden.

Bewijs:
Wij stellen voor willekeurige $U\in\delta_{\mathcal{J}}$ en $V\in\delta_{\mathcal{J}'}$ :

$$F'_V \overset{def}{=} F_{Y\cap V},\quad f'^x \overset{def}{=} f^x/^x_X \quad \text{en} \quad M'^y_U \overset{def}{=} /^y_Y M^y_{X\cap U}.$$

Dan zijn $F'$, $M'$ en $f'$ begrensd en is (1) aequivalent met:

(2) $$\int_{\mathcal{J}} \left\{ \int_{\mathcal{J}'} F'_{dy} M'^y_{dx} \right\} f'^x = \int_{\mathcal{J}'} F'_{dy} \int_{\mathcal{J}} M'^y_{dx} f'^x.$$

Dus is het voldoende de stelling te bewijzen voor <u>begrensde</u> fcts. en matrices met integratie over de hele vz $\mathcal{Y}$ ,resp. $\mathcal{Y}'$
Wij laten de indices weg en schrijven (2) in de vorm:

(3) $$(FM)f = F(Mf).$$

Verder stellen wij:

(4) $$g \overset{\text{def}}{=} Mf, \text{ dus } g^y = \int_{\mathcal{Y}} M^y_{dx} f^x$$

(5) $$\mathcal{G} \overset{\text{def}}{=} FM, \text{ dus } \mathcal{G}_X = \int_{\mathcal{Y}'} F_{dy} M^y_X .$$

Wij moeten bewijzen:

(6) $$\mathcal{G}f = Fg, \text{ dus } \int_{\mathcal{Y}} \mathcal{G}_{dx} f^x = \int_{\mathcal{Y}'} F_{dy} g^y .$$

Kies $\varepsilon > 0$ . Daar $f$ begrensd is, kunnen wij een <u>eindige</u> dissectie $\{X_n\} \in \mathcal{D}(\mathcal{Y})$ vinden, waarvoor $var_{\{X_n\}} f \leq \varepsilon$ .
(Verdeel bijvoorbeeld het eindige interval $-|f|^{\mathcal{Y}}, +|f|^{\mathcal{Y}}$ in N onderling disjuncte deelintervallen $\mathcal{J}_n$ met lengten $\leq \varepsilon$ en neem $X_n \overset{\text{def}}{=} \mathcal{E}ns\{x \in \mathcal{Y} \,|\, f^x \in \mathcal{J}_n\}$.)
Dus als wij $x_n \in X_n$ willekeurig kiezen geldt volgens (9.2.12):

(7) $$\left| \int \mathcal{G}_{dx} f^x - \sum^n \mathcal{G}_{X_n} f^{x_n} \right| \leq |\mathcal{G}| . \varepsilon .$$

Daar alle N fcts. $M_{X_n}$ begrensd zijn, kunnen wij een <u>eindige</u> dissectie $\{Y_m\} \in \mathcal{D}(\mathcal{Y}')$ vinden waarvoor $var_{\{Y_m\}} M_{X_n} \leq \frac{\varepsilon}{N}$ voor alle $n \in \{1, \ldots, N\}$ .(Neem bijvoorbeeld de "doorsnede" van de dissesties, die bij iedere $M_{X_n}$ afzonderlijk behoren).
Wij kiezen $y_m \in Y_m$ willekeurig. Nu geldt weer volgens (9.2.12):

(8) $$\left| \mathcal{G}_{X_n} - \sum^m F_{Y_m} M^{y_m}_{X_n} \right| \leq |F| . \frac{\varepsilon}{N} .$$

Uit (7),(8) en de betrekking $|\mathcal{G}| \leq |F|.\|M\|$ volgt dan:

(9) $$\left| \int \mathcal{G}_{dx} f^x - \sum^n \sum^m F_{Y_m} M^{y_m}_{X_n} f^{x_n} \right| \leq C . \varepsilon ,$$

waarbij

(10) $$C \overset{\text{def}}{=} |F| \left( \|M\| + \|f\| \right).$$

Aan de andere kant geldt voor iedere $y \in \mathcal{Y}'$, dus ook voor $y = y_m$:

(11) $$\left| \int M^y_{dx} f^x - \sum^n M^y_{X_n} f^{x_n} \right| \leq \|M\| . \varepsilon .$$

Dus:

(12) $\quad \left| \int F_{dy}\, g^y - \sum^n \left( \int F_{dy}\, M^y_{X_n} \right) f^{x_n} \right| \leqq |F| \cdot \|M\|\, \varepsilon .$

Bovendien:

(13) $\quad \left| \int F_{dy}\, M^y_{X_n} - \sum^m F_{y_m}\, M^{y_m}_{X_n} \right| \leqq |F| \cdot \frac{\varepsilon}{N} ,$

daar $\quad var_{\{y_m\}}\, M_{X_n} \leqq \frac{\varepsilon}{N} .$

Uit (12) en (13) volgt nu:

(14) $\quad \left| \int F_{dy}\, g^y - \sum^n \sum^m F_{y_m}\, M^{y_m}_{X_n}\, f^{x_n} \right| \leqq C\varepsilon$

(waarbij C gedefinieerd is volgens (10)).

Combineren wij (9) en (14), dan krijgen wij:

(15) $\quad \left| \int G_{dx}\, f^x - \int F_{dy}\, g^y \right| \leqq 2\, C\, \varepsilon .$

Dit geldt voor elke $\varepsilon > 0$, dus:

(6) $\quad Gf = Fg$, q.e.d.

Essentieel bij dit bewijs is de begrensdheid van $f$, $F$ en $M$. Hierdoor konden wij de dissecties eindig nemen.

Stelling 1 is dus een gevolg, zoals wij zagen, van de

Stelling 2: Als $f$, $F$ en $M$ begrensd zijn, dan geldt:

(16) $\quad \int_{\mathcal{Y}} \left\{ F_{dy}\, M^y_{dx} \right\} f^x = \int_{\mathcal{Y}'} F_{dy} \int_{\mathcal{Y}} M^y_{dx}\, f^x .$

Ook als $f$, $F$ en $M$ niet begrensd zijn, is het mogelijk, dat de integratievolgorde omgekeerd mag worden. Een voldoende voorwaarde hiervoor geeft

Stelling 3: Als

(17) $\quad H_X = \int_{\mathcal{Y}'} |F|_{dy}\, |M^y|_X \quad$ en

(18) $\quad h^y = \int_{\mathcal{Y}} |M^y|_{dx}\, |f^x|$

bestaan voor alle $X \in \delta_{\mathcal{Y}}$ en $y \in \mathcal{Y}'$ en als

òf $\quad \int_{\mathcal{Y}} H_{dx}\, |f^x| < \infty ,$

òf $\quad \int_{\mathcal{Y}'} |F|_{dy}\, h^y < \infty ,$

dan geldt ook de andere ongelijkheid en tevens:

(19) $$\int_{\mathcal{Y}} \left\{ \int_{\mathcal{Y}'} F_{dy} M^{y}_{dx} \right\} f^{x} = \int_{\mathcal{Y}'} F_{dy} \int_{\mathcal{Y}} M^{y}_{dx} f^{x}.$$

Bewijs: Het is voldoende de stelling te bewijzen voor niet negatieve $F$, $M$ en $f$, daar zij dan door haar (hoogstens) 8 maal toe te passen ook het algemene geval geeft. In dat geval is $H = G$ en $h = g$.
Er bestaat volgens 9.1 een rij vzn $\mathcal{I}_n$ met de eigenschappen $\forall_n\ \mathcal{I}_n \in \delta_{\mathcal{Y}}$, $\mathcal{I}_n \subset \mathcal{I}_{n+1}$ en $\cup \mathcal{I}_n = \mathcal{I}$.
Evenzo een rij $\mathcal{I}'_n$ met eigenschappen $\forall_n\ \mathcal{I}'_n \in \delta_{\mathcal{Y}'}$, $\mathcal{I}'_n \subset \mathcal{I}'_{n+1}$ en $\cup \mathcal{I}'_n = \mathcal{I}'$.

Stel $\int_{\mathcal{Y}} H_{dx} |f^x| < \infty$ (dus $\int_{\mathcal{Y}} G_{dx} f^x < \infty$).

Er geldt: $\int_{\mathcal{Y}} G_{dx} f^x \geq \int_{\mathcal{Y}_k} G_{dx} f^x \geq \int_{\mathcal{Y}_k} \left\{ \int_{\mathcal{Y}'_\ell} F_{dy} M^{y}_{dx} \right\} f^x = \int_{\mathcal{Y}'_\ell} F_{dy} \int_{\mathcal{Y}_k} M^{y}_{dx} f^x.$

Hierbij is stelling 1 gebruikt ($\mathcal{Y}_k \in \delta_{\mathcal{Y}}$, $\mathcal{Y}'_\ell \in \delta_{\mathcal{Y}'}$).

Daar de laatste uitdrukking niet dalend en begrensd is als $k \to \infty$ en $\ell \to \infty$ ($\leq \int G_{dx} f^x$), heeft zij een limiet. Deze is gelijk aan het rechterlid van (19).

Dus $\int_{\mathcal{Y}'} F_{dy}\, g^y < \infty$ en $\int_{\mathcal{Y}} G_{dx} f^x \geq \int_{\mathcal{Y}'} F_{dy}\, g^y$.

Nu wij weten, dat $\int_{\mathcal{Y}'} F_{dy}\, g^y < \infty$ is, kunnen wij, hiervan gebruik makende, op dezelfde manier aantonen, dat $\int_{\mathcal{Y}'} F_{dy}\, g^y \geq \int_{\mathcal{Y}} G_{dx} f^x$ is, waarmee (19) bewezen is.

Beginnen wij met aan te nemen, dat $\int_{\mathcal{Y}'} F_{dy}\, g^y < \infty$ is, dan loopt het bewijs analoog.

9.6 <u>Matrixvermenigvuldiging</u>

In 9.3 definieerden wij:

(9.3.11) $$(MN)^{x}_{y,z} = \int_{y} M^{x}_{dy} N^{y}_{z} \text{ en}$$

(9.3.12) $$(MN)^{x}_{z} = \int_{\mathcal{Y}'} M^{x}_{dy} N^{y}_{z}.$$

Het linkerlid van (9.3.12) is weer een matrix, wanneer $M$ en $N$ begrensd zijn of wanneer zij voldoen aan (9.3.13). Alleen in die gevallen is de definitie dus zinvol. Ook is het noodzakelijk, dat deze matrixvermenigvuldiging associatief is, hetgeen onder zeer algemene voorwaarden inderdaad zo is, zoals uit de volgende stelling volgt:

Stelling: Wanneer $\mathcal{J}_1, \mathcal{J}_2, \mathcal{J}_3$ en $\mathcal{J}_4$ willekeurige vzn zijn, waarop gegeven zijn $\delta$-velden $\delta_{\mathcal{J}_1}$, respectievelijk $\delta_{\mathcal{J}_2}, \delta_{\mathcal{J}_3}, \delta_{\mathcal{J}_4}$, wanneer wij de elementen van $\mathcal{J}_1, \mathcal{J}_2, \mathcal{J}_3$ en $\mathcal{J}_4$ aangeven met $x, y, z$ en $w$ respectievelijk, wanneer wij de vzn van $\delta_{\mathcal{J}_1}, \delta_{\mathcal{J}_2}, \delta_{\mathcal{J}_3}$ en $\delta_{\mathcal{J}_4}$ aangeven met $X, Y, Z$ en $W$ respectievelijk en wanneer verder $L, M$ en $N$ $(\mathcal{J}_1, \mathcal{J}_2)$, respectievelijk $(\mathcal{J}_2, \mathcal{J}_3)$ en $(\mathcal{J}_3, \mathcal{J}_4)$ matrices zijn, dan geldt:

$$(1) \qquad \int_Z \left\{ \int_Y L^x_{dy} M^y_{dz} \right\} N^z_W = \int_Y L^x_{dy} \int_Z M^y_{dz} N^z_W .$$

Als in het bijzonder $L, M$ en $N$ begrensd zijn, of als voldaan is aan de eis

$\int_{\mathcal{J}_3} |M^y|_{dz} |N^z|_W = H^y_W$ bestaat voor alle $y \in \mathcal{J}_2$ en $W \in \delta_{\mathcal{J}_4}$;

$\int_{\mathcal{J}_2} |L^x|_{dy} |M^y|_Z = K^x_Z$ bestaat voor alle $x \in \mathcal{J}_1$ en $Z \in \delta_{\mathcal{J}_3}$;

(2) alsmede

$\int_{\mathcal{J}_2} |L^x|_{dy} H^y_W < \infty$ voor alle $x \in \mathcal{J}_1$ en $W \in \mathcal{J}_4$, òf

$\int_{\mathcal{J}_3} K^x_{dz} |N^z|_W < \infty$ voor alle $x \in \mathcal{J}_1$ en $W \in \delta_{\mathcal{J}_4}$,

dan geldt de andere ongelijkheid ook en tevens (1), d.w.z. dat de integratie kan worden uitgestrekt over de hele vz $\mathcal{J}_2$ en de hele vz $\mathcal{J}_3$.

Bewijs: Volgt onmiddellijk uit de stellingen van 9.5 met - bij constante $x$ en $W$ - $L^x_Y$, $M^y_Z$ en $N^z_W$ in de plaats van $F_Y$, respectievelijk $M^y_X$ en $f^x$.

Als de producten $LM$ en $MN$ eveneens matrices zijn, bijvoorbeeld als $L$, $M$ en $N$ begrensd zijn of als aan de voorwaarde (9.3.13) voldaan is door $LM$ en $MN$, dan is het in (9.3.12) gedefiniëerde matrixproduct associatief:

$$(3) \qquad (LM)N = L(MN),$$

d.w.z.

$$(4) \qquad \int_{\mathcal{J}_3} (LM)^x_{dz} N^z_W = \int_{\mathcal{J}_2} L^x_{dy} (MN)^y_W .$$

De gewone matrixrekening, met Lebesque-Stieltjes-Radon integratie in plaats van sommatie, kan dus toegepast worden zodra de matrices en fcts. van eerste of tweede soort begrensd

zijn, of algemener, zodra aan (2) en bij matrixvermenigvuldiging eveneens aan (9.3.13) voldaan is.
Dikwijls zullen de vzn $\mathfrak{J}_1, \mathfrak{J}_2, \mathfrak{J}_3$ enz. samenvallen. In dat geval onderstellen wij, dat de corresponderende $\delta$-velden ook samenvallen.
Dan kunnen wij dus ook machten van matrices definiëren:

(5) $\quad (M^2)^x_X = \int_{\mathfrak{J}} M^x_{dx} M^x_X \quad$ als de integraal bestaat, dus bijvoorbeeld als $\|M\| < \infty$

Uit

(9.3.14) $\quad \|MN\| \leqq \|M\| . \|N\|$

volgt dan

(6) $\quad \|M^n\| \leqq \|M\|^n .$

## 9.7 Machtreeksen in een matrix

Definitie: De reeks $\sum_0^\infty{}^n c_n M^n \stackrel{def}{=} \lim_{N\to\infty} \sum_0^N{}^n c_n M^n$ convergeert wanneer $\forall_x^{\mathfrak{J}} \forall_X^{\delta_y} \sum_0^\infty{}^n c_n (M^n)^x_X$ convergeert $(M^0 \stackrel{def}{=} I)$.

Stelling: Als $\sum_0^\infty{}^n c_n z^n$ convergeert voor $|z| < R$, dan convergeert ook $\sum_0^\infty{}^n c_n M^n$ als $\|M\| < R$.

Bewijs: $\left(\sum_0^N c_n M^n\right)^x_X = \sum_0^N c_n (M^n)^x_X .$

Verder is $|(M^n)^x_X| \leqq \|M^n\| \leqq \|M\|^n$

Dus de reeks $\left(\sum_0^\infty c_n M^n\right)^x_X$ wordt absoluut gemajoreerd door de convergente reeks $\sum_0^\infty |c_n| \, \|M\|^n . \quad (\|M\| < R.)$

Dus convergeert ook $\sum_0^\infty c_n M^n$, q.e.d.

Als $\forall_n \, c_n = 1$, dan is ook $R = 1$ en zullen wij naar analogie van $\sum_0^\infty{}^n z^n = \frac{1}{1-z}$ voor $|z| < 1$, voor $\|M\| < 1$ verwachten:
$\sum_0^\infty M^n = (I - M)^{-1}$, de reciproke matrix van $I - M$.

Dit is inderdaad juist:

$$\lim_{N\to\infty} \sum_0^N{}^n M^n (I - M) = \lim_{N\to\infty} \sum_0^N{}^n (M^n - M^{n+1}) = I - \lim_{N\to\infty} M^{N+1} = I ,$$

daar $\lim_{N\to\infty} \|M\|^N = 0$, dus $\forall_x^{\mathfrak{J}} \forall_X^{\delta_y} \lim_{N\to\infty} (M^N)^x_X = 0 .$

Literatuurlijst Stochastische processen

1 N. Arley, 1943, On the theory of stochastic processes and their application to the theory of cosmic radiation. Kopenhagen, proefschrift universiteit van Kopenhagen.

2 G.A. Barnard, 1946, Sequential tests in industrial statistics. Journal Royal Statistical Society. Supplement 8, 1-21.

3 M.S. Bartlett, 1944, Negative probability, Proc.Cambr.Phil.Soc. 41, 71-73.

4 M.S. Bartlett, 1949, Some evolutionary stochastic processes. ibid.XI (1949), 211-230.

5 M.S. Bartlett and D.G. Kendall, 1951, On the use of the characteristic functional in the analysis of some stochastic processes occurring in physics and biology, Proc.Cambr.Phil.Soc.47, 65-76.

6 D. Blackwell and M.A. Girshick, 1946, On functions of sequences of independant chance vectors with applications to the problem of random walk in k dimensions, Annals Math.Statistics 17, 310-317.

7 A. Blanc-Lapierre et R. Fortet, 1953, Théorie des Fonctions Aléatoires, Parijs, Masson et Cie.

8 G. Blom, 1949, A generalization of Wald's fundamental identity, Annals Math.Statistics 20, 439-444.

9 J.P. Burman, 1946, Sequential sampling formulae for a Binomial population. Journal Royal Statistical Soc.8, 98-103.

10 L. le Cam, 1947, Un instrument d'étude des fonctions aléatoires: la fonctionnelle caractéristique, Comptes Rendus 224, 710-711.

11 R.G. Cooke, 1950, Infinite matrices and sequence spaces, Londen, Macmillan and Co.

12 D. van Dantzig, 1935, La notion de derivée d'une fonctionelle, Comptes Rendus 201, 1008-1010.

13 D. van Dantzig, 1936, Ricci-calculus and functional analysis, Proc.Kon.Ned.Ak.v.Wet.39, 785-794.

14 D. van Dantzig, 1941, Mathematische en empirische grondslagen der waarschijnlijkheidsrekening, Ned.tijdschrift voor Natuurkunde, 8, 70-93.

15 1946, Capita Selecta uit de waarschijnlijkheidsrekening. Caput 1, Grondslagen der waarschijnlijkheidsrekening; Caput 2, Absoluut Additieve Verzamelingsfuncties, Momenten en Ongelijkheden, Mathematisch Centrum, Amsterdam, gestencild.

16 1947, Mathematische Statistiek, Hoofdstuk 2, Collectieve kenmerken, voortbrengende en karakteristieke functies, Math.Centrum, Amsterdam, gestencild.

17 1947, General procedures of empirical sciences, Synthese 5, 441-455.

18 1949, Sur la méthode des fonctions génératrices, Colloques internationaux du centre national de la recherche scientifique, 13, 29-45.

19 1953/54, Het wiskundig model in de Ervaringswetenschappen, Euclides 29, 35-41.

20 1954, Chaînes de Markof dans les ensembles abstraits et applications aux processus avec régions absorbantes et au problème des boucles 14 (1955), 145-199, Ann. Inst. Poincaré.

21 D. van Dantzig and M.C. Scheffer, 1954, On arbitrary hereditary time-discrete stochastic processes, considered as stationary Markof-chains, and the corresponding general form of Wald's fundamental identity, Proc.Kon.Ned.Ak.v.Wet. 57, 377-388.

22 W. Doeblin, 1940, Eléments d'une théorie générale des chains simples constantes de Markof, Annals Sc.Ec.Norm.Sup.(3), 57, 61-111.

23 J.L. Doob, 1938, Stochastic processes with an integral-value parameter, Transactions Am.Math.Soc., 44, 87-151.

24 J.L. Doob, 1940, Regularity properties of certain families of chance variables,
Transactions Am.Math.Soc.47, 455-486.

25 1942, Topics in the theory of Markof chains,
Transactions Am.Math.Soc. 52, 37-46.

26 1945, Markof chains, denumerable case,
Transactions Am.Math.Soc.,58, 455-473.

27 1948, Asymptotic properties of Markof transition probabilities,
Transactions Am.Math.Soc.63, 393-421.

28 1953, Stochastic processes, New York, J. Wiley and Sons, Inc.

29 W. Feller, 1949, On the theory of stochastic processes, with particular reference to applications,
Proc. Berkeley Symposion on Math.Stat. and Prob.,403-432.

30 1950, An introduction to probability theory and its applications, Vol.I, New York, J. Wiley and Sons, Inc.

31 R. Fortet, 1935-1938, Sur les probabilités en chaînes,
Comptes Rendus 201, 202 et 204.

32 M. Fréchet, 1913, Sur la notion de différentielle dans le calcul des fonctionnelles,
Comptes Rendus Congrès des soc.savantes (1912), 45-58.

33 1914, Sur la notion de différentielle d'une fonction de ligne,
Transactions Am.Math.Soc.,15, 135-161.

34 1950, Généralités sur les probabilités. Eléments aléatoires - Traité du calcul des probabilités et de ses applications. Tome I (Les principes de la théorie des probabilités), Fascicule III (Recherches théoriques modernes sur le calcul des probabilités),
Livre I Généralités sur les probabilités et variables aléatoires.
Livre II Méthode des fonctions arbitraires. Théorie des événements en chaîne dans le cas d'un nombre fini d'états possibles.

35 M.A. Girshick, 1946, Contributions to the theory of sequential analysis, part II, Annals Math.Statistics 17, 282-291.

36 J. Hadamard, 1910, Leçons sur le calcul des variations, Tome I, Parijs, A. Hermann.

37 H. Hahn and A. Rosenthal, 1948, Setfunctions, Albuquerque, Univ. of New Mexico Press.

38 P.R. Halmos, 1950, Measure theory, New York, D.van Nostrand Company.

39 G.H. Hardy, J.E. Littlewood, G. Polya, 1934, Inequalities, Cambridge University Press.

40 M.B. Hostinsky, 1931, Methodes générales du calcul des probabilités. Mém.des Sciences Math.7.

41 1937, Sur les probabilités rélatives aux variables aléatoires liées entre elles. Am.Inst. H. Poincaré 7, 69.

42 J.H.B. Kemperman, 1950, The general one-dimensional random walk with absorbing barriers, with applications to sequential analysis, (dissertatie, Amsterdam), Den Haag, Excelsiors Foto-Offset.

43 D.G. Kendall, 1948, On the generalized birth and death process. Ann.Math.Stat. 19, 1-15.

44 1949, Stochastic processes and population growth. Suppl. J. Royal Stat. Soc. XI, 230-264.

45 A. Khintchine, 1934, Korrelationstheorie der stationären stochastischen Prozesse. Math. Ann. 109, 604-615.

46 A. Kolmogorof, 1933, Grundbegriffe der Wahrscheinlichkeitsrechnung, Erg.d.Math.2 no.3.

47 1936, Anfangsgründe der Markof-ketten mit unendlich vielen möglichen Zuständen.

48 1937, Markof-chains with a constable number of possible states (Russisch). Bull.Math.Univ.Moskou 1 no.3.

49 A. Kolmogorof and N. Dmitrieff, 1947, Branching stochastic processes. Comptes Rendus Ac.Sci. URSS (N.S.) 56, 5-8.

50 P. Levy, 1948, Processus stochastiques et mouvement brownien, Parijs, Gauthier Villars.

51 C.C. MacDuffee, 1942, Vectors and matrices, The Mathematical Association of America.

52 A. Markof, 1906, Extension of the law of large numbers to dependent events, Bull.Soc.Phys.Math.Kazan 15, 135-156.

53 V. Romanofsky, 1936, Recherches sur les chaînes de Markof. Acta Math.66, 147.

54 O. Schreier und E. Sperner, 1935, Einführung in die analytische Geometrie und Algebra, B.G. Teubner.

55 Olga Taussky, 1949, A recurring theorem on determinants, Am.Math.Monthly 56, 672-676.

56 B.L. van der Waerden, 1930, Moderne algebra, Berlijn, Julius Springer.

57 A. Wald, 1945, Sequential analysis, New York, J. Wiley and Sons, Inc.

58 W. Wasow, 1951, On the duration of random walks, Annals of Math.Stat.22, 199-216.

59 J.H.M. Wedderburn, 1934, Lectures on matrices, Am. Math.Soc., Colloquium Publications.