STICHTING
MATHEMATISCH CENTRUM
2e BOERHAAVESTRAAT 49
AMSTERDAM

ZW 1964-004

Voordracht in de serie "Actualiteiten"

Prof.dr. J.Th. Runnenburg

26 februari 1964

Stochastische wandeling en de driehoek van Pascal.

1964

ZW 1964-004

STICHTING
MATHEMATISCH CENTRUM
2e BOERHAAVESTRAAT 49
AMSTERDAM

AFDELING MATHEMATISCHE STATISTIEK

Voordracht in de serie "Actualiteiten"

door

Prof.dr. J.Th. Runnenburg

26 februari 1964

Stochastische wandeling en de driehoek van Pascal

In onze beschouwingen wordt voortdurend gebruik gemaakt van "optelregels", zoals die bij de driehoek van Pascal gebruikt worden. Deze "driehoek" heeft het volgende uiterlijk:

$$
\begin{array}{ccccccccccccc}
 & & & & & & 1 & & & & & & \\
 & & & & & 1 & & 1 & & & & & \\
 & & & & 1 & & 2 & & 1 & & & & \\
 & & & 1 & & 3 & & 3 & & 1 & & & \\
 & & 1 & & 4 & & 6 & & 4 & & 1 & & \\
 & 1 & & 5 & & 10 & & 10 & & 5 & & 1 & \\
1 & & 6 & & 15 & & 20 & & 15 & & 6 & & 1
\end{array}
$$

enz.

Schrijven wij op de bovenste regel links en rechts van de 1 rijen nullen, dan wordt de tweede regel verkregen door midden onder twee getallen van de bovenste regel de som van die getallen te plaatsen. Op een dergelijke wijze ontstaat de derde regel uit de tweede, enz. Met een eenvoudige optelregel (die uit het voorgaande wel duidelijk is), wordt zo uit één rij getallen (de bovenste regel of 0^{de} rij) elke volgende rij getallen gevonden. In de n^{de} rij van de driehoek staat op de j^{de} plaats het getal $\binom{n}{j}$ voor $0 \leq j \leq n$. De optelregel wordt weergegeven door $\binom{n-1}{j-1}+\binom{n-1}{j}=\binom{n}{j}$.

In de waarschijnlijkheidsrekening worden sommen

$$\underline{s}_n=\underline{s}_0+\underline{x}_1+\ldots+\underline{x}_n \qquad (n \geq 0)$$

bestudeerd, waarbij $\underline{s}_0,\underline{x}_1,\underline{x}_2,\ldots,\underline{x}_n,\ldots$ onderling onafhankelijke stochastische variabelen zijn, die slechts gehele waarden aannemen en de $\underline{x}_1,\underline{x}_2,\ldots$ alle dezelfde kansverdeling

$$p_j \overset{def}{=} P\{\underline{x}_n=j\} \quad (j \text{ geheel})$$

bezitten. Dus $p_j \geq 0$ en $\sum_{j=-\infty}^{\infty} p_j=1$.

Hier veronderstellen we, dat slechts twéé der p_j's ongelijk nul zijn en wel

$$\begin{cases} p \overset{def}{=} p_{-u} > 0 & \text{voor een } u > 0, \\ q \overset{def}{=} p_v > 0 & \text{voor een } v > 0. \end{cases}$$

We vragen onder deze omstandigheden naar

$$\begin{cases} q_{ik}^{(n)} \overset{def}{=} P\{\underline{s}_1 < 0,\ \underline{s}_2 < 0,\ldots,\underline{s}_{n-1} < 0, \underline{s}_n=k\,|\,\underline{s}_0=i\} & \text{voor } i,k < 0, \\ q_{ik}^{(n)} \overset{def}{=} P\{\underline{s}_1 > 0,\ \underline{s}_2 > 0,\ldots,\underline{s}_{n-1} > 0, \underline{s}_n=k\,|\,\underline{s}_0=i\} & \text{voor } i,k > 0. \end{cases}$$

Deze kansen zijn op grond van onze gegevens als volgt gedefiniëerd voor $i,k < 0$: beschouw alle n-tallen $(x_1,x_2,\ldots,x_n)$ met $x_\nu=-u$ of $x_\nu=v$ voor $1 \leq \nu \leq n$ en $\sum_{\nu=1}^{n} x_\nu=k-i$. Er is óf geen enkel n-tal (en dan zeggen we: k kan niet in n stappen vanuit i bereikt worden) of er is wel minstens één n-tal (k kan in n stappen vanuit i bereikt worden).

Is er een n-tal, dan bestaat dit uit j keer $-u$ en $n-j$ keer v met

$$-ju+(n-j)v=k-i$$

of

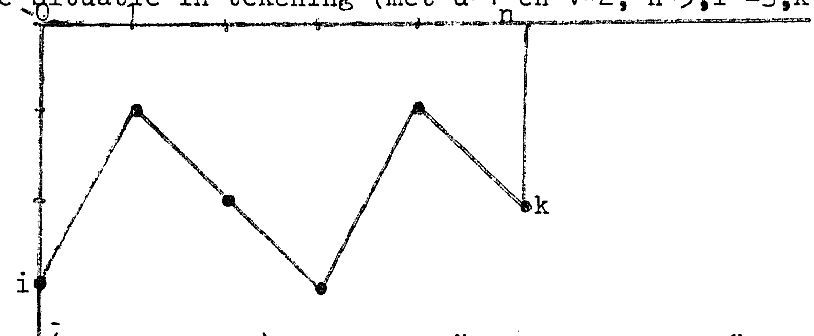
$$j=\frac{nv-k+i}{u+v}\quad .$$

Dus j kan slechts één waarde aannemen en er zijn $\binom{n}{j}$ verschillende n-tallen, die aan de condities voldoen. Een n-tal met j keer $-u$ en $n-j$ keer v heeft kans $p^j q^{n-j}$. Inmiddels weten we nu

$$p_{k-i}^{(n)} \overset{def}{=} P\{\underline{s}_n=k\,|\,\underline{s}_0=i\} = \binom{n}{j}\, p^j q^{n-j}.$$

Merk op, dat $p_{k-i}^{(n)}$ slechts van $k-i$ afhangt.

We brengen de situatie in tekening (met $u=1$ en $v=2$, $n=5, i=-3, k=-2$):



Elk n-tal $(x_1,x_2,\ldots,x_n)$ vormt een "pad ter lengte n" van i naar k. Sommige van deze paden blijven beneden de horizontale as (zie figuur), andere niet. Er geldt voor $i,k < 0$

$$q_{ik}^{(n)} = (\text{aantal paden ter lengte } n \text{ van } i \text{ naar } k \text{ beneden de as})\; p^j q^{n-j}.$$

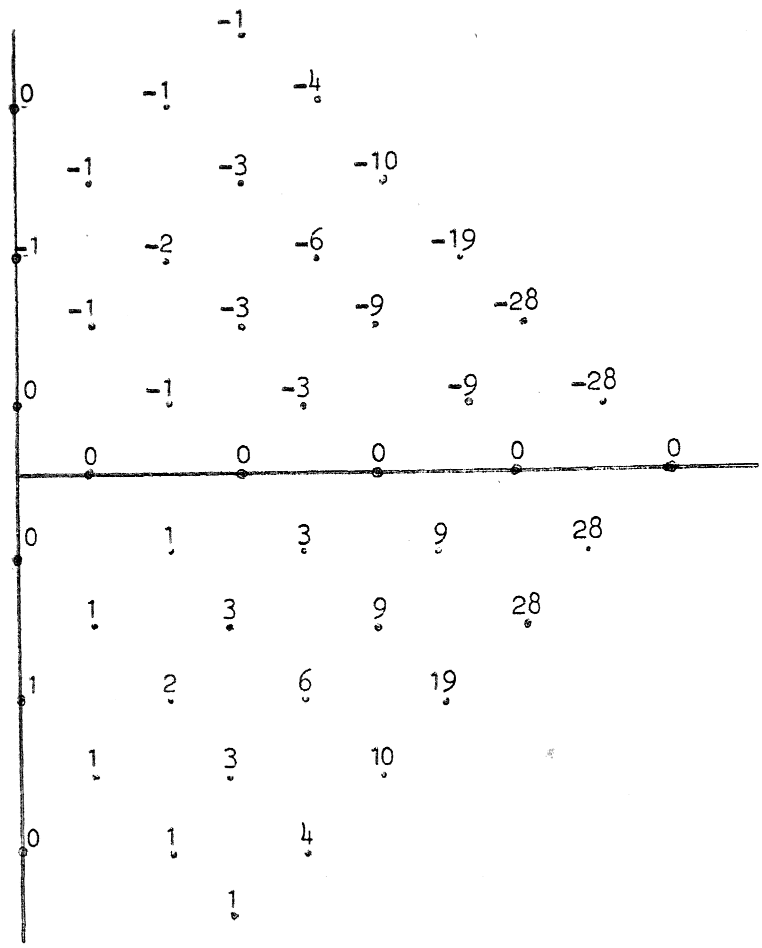
Als we dus afzien van de factor $p^j q^{n-j}$, dan wordt $p_{k-i}^{(n)}$ bepaald door op de juiste plaats in de driehoek van Pascal te kijken:

$$\begin{aligned}\binom{n}{j} &= \text{aantal paden ter lengte } n \text{ van } i \text{ naar } k = \\ &= \text{aantal paden ter lengte } n-1 \text{ van } i \text{ naar } k+u + \text{aantal} \\ &\qquad\quad \text{paden ter lengte } n-1 \text{ van } i \text{ naar } k-v = \\ &= \binom{n-1}{j-1}+\binom{n-1}{j}.\end{aligned}$$

Kan $q_{ik}^{(n)}$ nu óók met een optelregel verkregen worden (als we afzien van de factor $p^j q^{n-j}$)?

In het volgende geven we enige eenvoudige toepassingen van een idee van Keilson.

A Neem u=v=1 en i=-3. Dan worden de getallen onder de horizontale as gevraagd in onderstaande figuur.



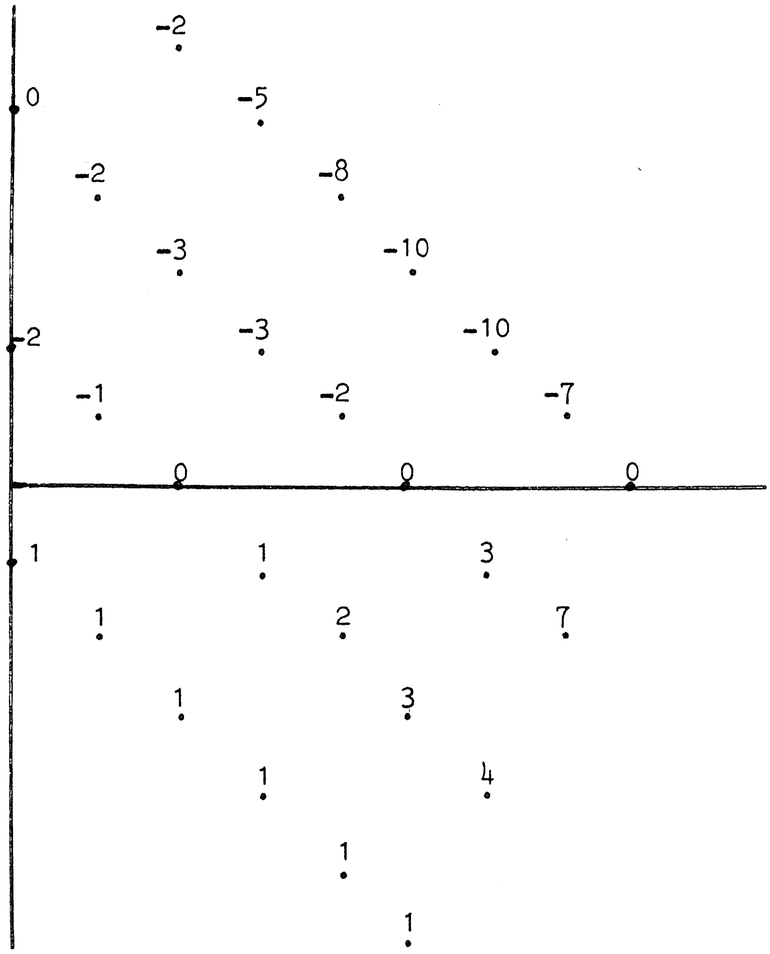
Ook hier kan met een <u>optelregel</u> gewerkt worden: in (0,i) op de verticale as plaatsen we het getal 1 en in de punten (0,i+h(u+v)) met h geheel, h≠0 en i+h(u+v) < 0 plaatsen we het getal 0. De volgende kolom wordt uit de vorige door optelling verkregen. Het rooster der uit i bereikbare punten van het vlak vullen we aan met bijpassende punten boven en op de horizontale as. In de roosterpunten op de horizontale as plaatsen we het getal 0 en door vast te houden aan de optelregel kennen we ook aan de roosterpunten boven de horizontale as een getal toe. Het getal, dat bij de plaats (n,k) wordt neergeschreven (met $k < 0$) is ≥ 0 en juist het tegengestelde van het getal, dat bij (n,-k) komt te staan. Op grond van de optelregel staat nu in het vanuit i bereikbare roosterpunt(n,i-n+2j) bij u=v=1 het getal

$$\binom{n}{j} - \binom{n}{j-i} ,$$

want de figuur kan beschouwd worden als ontstaan door optelling van twee driehoeken van Pascal, waarbij in één alle getallen negatief zijn genomen.

Door u=v=1 te kiezen, hebben we een heel eenvoudig antwoord voor $q_{ik}^{(n)}$ verkregen. Wat gebeurt er bij minder speciale u en v?

B Neem u=1 en v=2. Met de techniek van de optelregel van geval A vinden we voor i=-1 onderstaande figuur.



De getallen in de figuur worden verkregen, door vanuit (0,-1) de driehoek van Pascal met "gewicht 1" "uit te zenden" langs de roosterpunten en vanuit (0,2) met gewicht -2. In het vanuit i=-1 bereikbare roosterpunt (n,-1-n+3j) staat op grond van dit "zenden" het getal

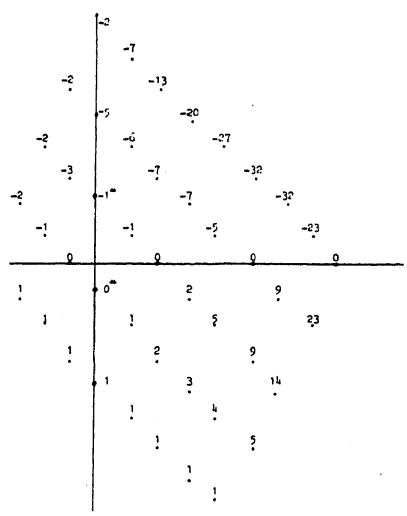
$$\binom{n}{j}-2\binom{n}{j-1} .$$

Dat op deze wijze het gewenste antwoord is verkregen, is bewezen, zodra we hebben aangetoond, dat het uitzenden van driehoeken van Pascal op de aangegeven wijze tot 0 in de roosterpunten op de horizontale as voert. We moeten dus controleren

$$\binom{3m+2}{m+1}-2\binom{3m+2}{m}=0 \quad \text{voor } m \geq 0.$$

Dit is inderdaad juist.

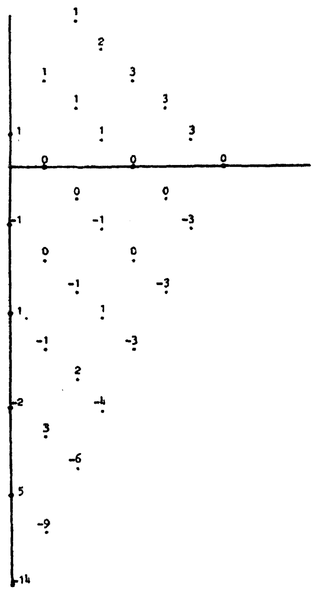
De figuur voor i=-2 vinden we uit de zojuist vervaardigde, door de verticale as over een afstand 1 naar rechts te verplaatsen: de optelregel blijft vervuld! Ook voor i=-3 vinden we op een dergelijke manier de juiste getallen. Voor i=-4 kunnen we de gezochte getallen vinden, door vanuit (-3,-1) met gewicht 1, vanuit (-3,2) met gewicht -2, vanuit (0,-1) met gewicht -1 en vanuit (0,2) met gewicht 2 de driehoek van Pascal uit te zenden. Onderstaande figuur geeft het resultaat.



Uit de figuur blijkt, dat het ook mogelijk is, de driehoek van Pascal vanuit (0,-4) met gewicht 1, vanuit (0,2) met gewicht -1, vanuit (0,5) met gewicht -5 en vanuit (0,8) met gewicht -2 uit te zenden.

Met u=1, v >2 vinden we analoge resultaten. Wat gebeurt er als we i > 0 nemen?

C Neem u=1 en v=2. Begin in (0,1), d.w.z. kies i=1. Als eerder bouwen we onze figuur op. Het resultaat is hieronder weergegeven. We blijken nu niet met het combineren van eindig veel Pascaldriehoeken te kunnen volstaan.



De rij 1,2,5,14,... blijkt te zijn $\frac{1}{m+1}\binom{2m}{m}$ voor m=1,2,... . Er komen vrij veel nullen in de figuur voor.

Het is duidelijk, dat we met i > 0 een veel lastiger probleem hebben aangesneden (bij u=1). Toch kan ook nu nog vrij eenvoudig $q_{ik}^{(n)}$ bepaald worden, voor i > 0 en k > 0. Volgens Kemperman geldt namelijk

$$q_{ik}^{(n)} = q_{-k,-i}^{(n)}$$

Dit kunnen we inzien, door de paden ter lengte n van i naar k ééneénduidig aan de paden ter lengte n van -k naar -i toe te voegen: als het eerste pad door $(x_1, x_2, \ldots, x_n)$ gegeven wordt, dan is het toegevoegde pad $(x_n, x_{n-1}, \ldots, x_1)$. Een pad boven de as gaat door deze toevoeging over in een pad onder de as. Als het eerste pad j keer $-u$ bevat, dan geldt dit ook voor het tweede pad. Beide hebben dus dezelfde kans $p^j q^{n-j}$. Op de horizontaal direct boven de as (in $(3m,1)$ met $m=0,1,2,\ldots$) vinden we hier de getallen $1,1,3,\ldots$. Deze komen inderdaad overeen met de getallen op de horizontaal direct onder de as (in $(3m,-1)$ met $m=0,1,2,\ldots$) in de figuur bij voorbeeld B.

Nemen we $u>1$ en $v>1$, dan is alle eenvoud uit de figuren verdwenen en heeft het blijkbaar geen zin meer deze volledige roosters (afgemaakt naar de verticale as) te beschouwen.

We dienen echter niet uit het oog te verziezen, dat alleen die getallen in de figuren voor ons van belang zijn, die aan dezelfde kant van de horizontale as liggen als i en k.

We behandelen de volgende methode aan de hand van een eenvoudig voorbeeld, hoewel pas bij ingewikkelde problemen wezenlijk nieuwe resultaten worden gevonden. Algemene theorie vindt men bij Kemperman en Keilson.

D Neem $u=1$ en $v>0$. Kies weer $i=1$. We zoeken als steeds $q_{1k}^{(n)}$. Bekend is $p_{k-1}^{(n)}$, de kans om zonder hinderlijke barrière van 1 naar k te gaan in n stappen:

$$p_{-n+j(v+1)}^{(n)} = \binom{n}{j} p^{n-j} q^j \quad \text{voor} \quad 0 \leq j \leq n.$$

Kunnen we $q_{1k}^{(n)}$ nu in de $p_{1k}^{(m)}$ uitdrukken? Er geldt voor $k>0$

$$p_{k-1}^{(n)} = q_{1k}^{(n)} + P\{\text{minstens één } \underline{s}_\nu \leq 0 \text{ voor } 1 \leq \nu \leq n \text{ en } \underline{s}_n = k \mid \underline{s}_0 = 1\} =$$

$$= q_{1k}^{(n)} + \sum_{\nu=1}^{n-1} P\{\underline{s}_1 > 0, \ldots, \underline{s}_{\nu-1} > 0,\ \underline{s}_\nu = 0,\ \underline{s}_n = k \mid \underline{s}_0 = 1\} =$$

$$= q_{1k}^{(n)} + \sum_{\nu=1}^{n-1} P\{\underline{s}_1 > 0, \ldots, \underline{s}_{\nu-1} > 0,\ \underline{s}_\nu = 0 \mid \underline{s}_0 = 1\}\ P\{\underline{s}_n = k \mid \underline{s}_\nu = 0\} =$$

$$= q_{1k}^{(n)} + \sum_{\nu=1}^{n-1} r_\nu p_k^{(n-\nu)}$$

of

$$(*) \qquad q_{1k}^{(n)} = p_{k-1}^{(n)} - \sum_{\nu=1}^{n-1} r_\nu p_k^{(n-\nu)} \ ,$$

waarbij

$$r_\nu \overset{\text{def}}{=} P\{\underline{s}_1 > 0,\ldots,\underline{s}_{\nu-1} > 0,\ \underline{s}_\nu=0 \mid \underline{s}_0=1\} \text{ voor } \nu \geq 1.$$

De relatie (*) kan aldus geïnterpreteerd worden: in (n,k) vinden we het gezochte aantal paden, door vanuit $(0,1)$ met gewicht 1 uit het onbeperkte proces te zenden en met gewicht $-r_\nu$ vanuit $(\nu,0)$ voor $\nu=1,2,\ldots$ het onbeperkte proces ook uit te zenden! Maar hoe groot is r_ν? We kunnen de laatste afleiding met $k=0$ herhalen. Er volgt dan

$$p_{-1}^{(n)} = \sum_{\nu=1}^{n} r_\nu p_0^{(n-\nu)} \quad \text{voor } n \geq 1$$

met

$$p_0^{(0)} \overset{\text{def}}{=} 1.$$

Alleen in roosterpunten kan $r_\nu \neq 0$ gelden. We dienen dus alleen $r_{1+(v+1)m}$ voor $m=0,1,2,\ldots$ te bepalen. Er geldt $r_{1+(v+1)m}=a_m p^{1+vm} q^m$ voor $m \geq 0$, waarbij a_m een aantal paden voorstelt. Nu is

$$p_{-1}^{(1+(v+1)m)} = \sum_{\nu=1}^{1+(v+1)m} r_\nu p_0^{(1+(v+1)m-\nu)}$$

of

$$p_{-1}^{(1+(v+1)m)} = \sum_{k=0}^{m} r_{1+(v+1)k} p_0^{((v+1)(m-k))}$$

of

$$\binom{1+(v+1)m}{m} p^{1+vm} q^m = \sum_{k=0}^{m} a_k p^{1+vk} q^k \binom{(v+1)(m-k)}{m-k} p^{v(m-k)} q^{m-k},$$

of

$$\binom{1+(v+1)m}{m} = \sum_{k=0}^{m} a_k \binom{(v+1)(m-k)}{m-k}.$$

De genererende functie

$$a(x) = \sum_{k=0}^{\infty} a_k x^k$$

voldoet dus aan

$$\sum_{k=0}^{\infty} \binom{1+(v+1)k}{k} x^k = a(x) \cdot \sum_{k=0}^{\infty} \binom{(v+1)k}{k} x^k.$$

Nu geldt volgens Pólya: als bij $\alpha \geq 0$ en $\beta > 0$ een ε met $0 < \varepsilon < 1$ gekozen wordt en x voldoet aan

$$|x| < \frac{\varepsilon}{(1+\varepsilon)^{\beta}} ,$$

dan is

$$\sum_{k=0}^{\infty} \binom{\alpha+\beta k}{k} x^k = \frac{(1+z_0)^{\alpha}}{1-\beta x(1+z_0)^{\beta-1}}$$

en (iets uitgebreid)

$$(1+z_0)^{\alpha} = \sum_{k=0}^{\infty} \frac{\alpha}{\alpha+\beta k} \binom{\alpha+\beta k}{k} x^k.$$

Hierbij is $z_0=z_0(x,\beta)$ de enige wortel van

$$z-x(1+z)^{\beta}=0,$$

die voldoet aan $|z_0|<\varepsilon$.

Maar dan is

$$a(x)= 1+z_0(x,v+1)= \sum_{k=0}^{\infty} \frac{1}{1+(v+1)k} \binom{1+(v+1)k}{k} x^k$$

of

$$a_k= \frac{1}{1+(v+1)k} \binom{1+(v+1)k}{k} \quad \text{voor } k \geq 0.$$

Dit antwoord kan met de eerst besproken methode sneller gevonden worden.

<u>Literatuur</u>

Keilson, J., Green's function methods for bounded processes, University of Birmingham (mimeographed), 1963.

Kemperman, J.H.B., The passage problem for a stationary Markov chain, University of Chicago Press, 1961.

Pólya, G., Sur les séries entières dont la somme est une fonction algébrique, l'Enseignement Mathématique <u>22</u> (1921) 38-47.

typ:HvD