ZW

stichting
mathematisch
centrum

MC

AFDELING ZUIVERE WISKUNDE ZC 80/71 FEBRUARI

J. VAN DER SLOT
CURSUS WISKUNDE

Topologie en Lineaire algebra

ZW

2e boerhaavestraat 49 amsterdam

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

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

INHOUD

Hoofdstuk 1

Inleiding tot de wiskunde

Deze inleiding heeft tot doel om een aantal begripsvormingen te behandelen die in alle delen van de wiskunde optreden.
Niet overal is naar volledigheid gestreefd; hier en daar is op een enigszins populaire wijze aansluiting gezocht aan de kennis van de schoolwiskunde.

De volgorde van de paragrafen is nogal willekeurig, en doet misschien vreemds aan. Dat komt doordat de stof zich tot een aantal af en toe los van elkaar staande zaken beperkt.

Er is geen serieuze poging gedaan om de grondslagen van de wiskunde te behandelen. Zo wordt niet gesproken over de eisen waaraan in de wiskunde de axioma's, definities, stellingen en bewijzen moeten voldoen. Evenmin wordt gepoogd een houdbare definitie van het begrip "verzameling" te geven.

§1. Logica.

Het is niet de bedoeling hier "de logica" te behandelen; dat is een vak op zichzelf. Het is ook niet nodig; de hoeveelheid logica die in de wiskunde dagelijks wordt gebruikt, is bij ieder weldenkend mens aanwezig. Wij wijzen hier slechts op enkele punten, en geven enkele belangrijke notaties.

Beweringen. We beschouwen allerlei beweringen, ook wel "volzinnen" of "uitspraken" genoemd. Gemakshalve stellen we vaak een bewering voor door één hoofdletter.

Zo'n bewering kan evengoed een juiste als een onjuiste bewering voorstellen. Voorbeeld: A = "$2+2=4$"; B = "$3 \times 6 < 7$"

Ontkenning. Is B een bewering, dan is $\neg B$ (spreek uit niet -B) de notatie voor de ontkenning van B. Steeds is of B of $\neg B$ waar.
Voorbeeld $\neg(2 \times 2=4)$ betekent $(2 \times 2 \neq 4)$.

Conjunctie. De uitspraak "A en B" (notatie $A \wedge B$) is slechts waar als A en B beide waar zijn.
v.b. $A : a \geq b \quad B : a \leq b \quad A \wedge B : a = b$.

Disjunctie. De uitspraak "A of B" (notatie $A \vee B$) betekent

$$\neg\{(\neg A) \wedge (\neg B)\},$$

en is dus slechts waar als niet beide onwaar zijn. In de dagelijkse spreektaal wordt "of" vaak (en dan meestal beklemtoond) in uitsluitende zin gebruikt: één van beide maar niet allebei. Als wij dat willen aanduiden, zullen we de zinswending "of A, óf B" gebruiken, maar we voeren daarvoor geen afzonderlijke notatie in.

Implicatie. De veel gebruikte formule $A \to B$ is een afkorting van de uitspraak $(\neg A) \vee B$. We kunnen hem onder woorden brengen met "als A waar is, dan is B ook waar of ook "Uit A volgt B".

$A \to B$ is waar in de volgende gevallen: A en B waar,
A onwaar en B waar, A onwaar en B onwaar

$A \to B$ is onwaar slechts als A waar en B onwaar.

Let op dat met deze beweringen niet wordt uitgesproken dat A waar is, en ook niet dat B waar is.

In de wiskunde wordt de implicatie veel gebruikt om bepaalde eigenschappen af te leiden. Stel men wil bijvoorbeeld van een getal g, waarvan zekere eigenschappen gegeven zijn, bewijzen dat het nul is. Men geeft een indirect bewijs uit de onderstelling dat $g \neq 0$ is, leidt men met behulp van de gegeven eigenschappen iets af dat kennelijk onjuist is, bijvoorbeeld $(g \neq 0) \to (2 + 2 = 2)$. Uit het feit dat deze implicatie juist is volgt evenzeer dat $g = 0$.

Equivalentie. De beweringen A en B heten equivalent als het niet waar is dat één van twee waar en de andere onwaar is. Notatie $A \longleftrightarrow B$. Men zegt vaak hiervoor "A geldt dan en slechts dan als B geldt. $A \longleftrightarrow B$ is dus alleen een ware bewering als $A \to B$ en $B \to A$ beide waar zijn.

Om in een bepaald geval $A \longleftrightarrow B$ te bewijzen, is het ook voldoende te laten zien dat

$$A \to B \text{ en } \neg A \to \neg B$$

Veelal maakt men de fout dat men $A \to B$ en $B \to \neg A$ bewijst en denkt dat men daarmee $A \leftrightarrow B$ heeft bewezen.

Contrapositie van een implicatie. Is $A \to B$ dan is ook $B \to \neg A$. De laatste implicatie heet de contrapositie van de eerste, en is hiermede gelijkwaardig. Op deze eigenschap berusten de bewijsmethoden "uit het omgerijmde": om te bewijzen $A \to B$ bewijst men $\neg B \to \neg A$. Bovenstaand bewijs met het getal g is hier ook een voorbeeld van. I.h.a. is de gang van zake bij zo'n bewijs als volgt: we moeten A bewijzen neem dus A aan. Stel nu dat B niet geldt. Leid uit A en $\neg B$ een tegenspraak af. De onderstelling dat B niet waar is was dus fout, zodat B wel geldt.

Omkering van een implicatie. $B \to A$ heet de omkering van $A \to B$. Soms is de omkering van een juiste implicatie ook nog waar en soms niet. Het is enigszins vaag om over de omkering van een stelling te spreken, want een stelling heeft niet altijd het karakter van een eenvoudige implicatie.

Nodige en voldoende voorwaarden. Als $A \to B$, dan zegt men vaak dat A een voldoende voorwaarde voor B is. Is $\neg A \to \neg B$ dan heet A een nodige voorwaarde voor B. Is A tegelijk nodig en voldoende voor B, dan zijn A en B equivalent. Voorbeelden

$g > 0$ is nodig op dat $g - 1 > 0$ is, het is echter niet voldoende
$g > 0$ is voldoende opdat $g + 1 > 0$ is, het is echter niet nodig.

Quantoren. Laat B(x) een uitdrukking zijn waarin op een of meer plaatsen de letter x optreedt. De letter x heet een variabele. Gemakshalve beperken we ons tot het substitueren van dingen van een nader afgesproken soort, bijvoorbeelde de reële getallen. B(x) is bijvoorbeeld een uitdrukking zoals $x > 3$, $x^2 = 4$, $x^2 + 2x + 1 = (x+1)^2$. In plaats van te zeggen dat voor een zekere x de uitspraak B(x) waar is, zeggen we ook wel, dat x aan B(x) voldoet. We gebruiken nu de formule

$$\forall x \quad (B(x))$$

om te beweren, dat voor elke x van de bedoelde soort B(x) waar is.

$\forall$ heet de <u>universele</u> <u>quantor</u> en ook wel het <u>al</u>-symbool. Even zo beweert

$$\exists x\ (B(x))$$

dat er minstens één x is waarvoor B(x) juist is.
$\exists$ heet de <u>existentiële</u> <u>quantor</u>, of het existentiesymbool.
Voorbeelden

$$\forall x\ (x^2+1) > 0$$
$$\forall x\ [x^2+2x+1 = (x+1)^2]$$
$$\exists x\ [x^2 = 4]$$
$$\exists x\ [(x^2 > 0) \rightarrow x = 1)]$$

Al deze uitspraken zijn juist.

De letter x had betrekking op dingen van een zekere soort. Als het zeker is dat er dingen van die soort bestaan, dan geldt de implicatie

$$\forall x\ (B(x)) \rightarrow \exists x\ (B(x)).$$

Veelal komen in de wiskunde beweringen voor met verschillende quantoren achter elkaar. Is B(x,y) een bewering die de beide letters x en y bevat, dan kan men bijvoorbeeld de bewering

$\forall x\ \exists y\ [B(x,y)]$ beschouwen. Deze beweert dat voor elke x er steeds een y bestaat zodanig dat B(x,y) geldt.

<u>§2 Verzamelingen</u>

Het begrip "verzameling" veronderstellen we bekend.-
Synoniemen voor het woord verzameling zijn o.a. familie, collectie, stelsel. De objecten waaruit een verzameling is opgebouwd heten de <u>elementen</u> van de verzameling. Het feit dat een object a een element is van een verzameling V wordt uitgedrukt door de formule $a \in V$. De ontkenning daarvan is $a \notin V$.

Verzamelingen kunnen op verschillende manieren worden gevormd.

1e. Door (als dat kan) de elementen in zekere volgorde op te noemen. Bijvoorbeeld de verzameling die bestaat uit de getallen 3, 8, 11. Deze verzameling geven we aan met $\{3, 8, 11\}$.
Men bedenke dat $\{3, 8, 11\}$ hetzelfde betekent als $\{8, 3, 11\}$.
2e. Door het noemen van een eigenschap: de verzameling is dan de verzameling van alle objecten die deze eigenschap hebben. Drukken we de eigenschap uit door $B(x)$ (dit is een bewering over x), en we zeggen dat x de eigenschap B heeft als $B(x)$ waar is), dan geven we de verzameling aan met

$$\{x \mid B(x)\}$$

<u>Voorbeelden</u>. 1. $\{x \mid x$ is een reëel getal en $x > 2\}$ stelt voor de verzameling van alle reële getallen > 2.
2. $\{x \in R \mid x^2-x = 0\} = \{0,1\}$.

Om zekere uitspraken een algemenere geldigheid te geven, voeren we ook de <u>lege</u> verzameling in, die geen enkel element bevat. Notatie $\emptyset$.

X heet een <u>deelverzameling van</u> Y, wanneer ieder element van X ook element van Y is. Notatie $X \subset Y$.
In onze logicanotaties kunnen we dit ook noteren met $(x \in X \to x \in Y)$

<u>Voorbeelden</u>. $\{1, 2, 5\} \subset \{1, 2, 3, 4, 5\}$; $N \subset Z$, $Z \subset Q$ etc. $\emptyset \subset X$.

Twee verzamelingen X en Y zijn gelijk (notatie $X = Y$) wanneer zij uit dezelfde elementen bestaan, dus als ieder element van X ook tot Y behoort en omgekeerd, m.a.w. als $X \subset Y$ en $Y \subset X$. De gelijkheid van twee verzamelingen wordt in de praktijk vaak bewezen door deze twee relaties te verifiëren. We moeten dus aantonen $x \in X \to x \in Y$ en bovendien $x \in Y \to x \in X$.

De <u>doorsnede</u> van twee verzamelingen X en Y is de verzameling die bestaat uit die elementen, die zowel tot X als tot Y behoren, notatie $X \cap Y$.
Als $X \cap Y = \emptyset$ dan is heten X en Y <u>disjunct</u>.
Dus $X \cap Y = \{x \mid (x \in X) \wedge (x \in Y)\}$.

De vereniging van twee verzamelingen X en Y is de verzameling die bestaat uit die elementen, die tot minstens één van de verzamelingen X en Y behoren.

Notatie $X \cup Y$. Dus $X \cup Y = \{x(x \in X) \vee (x \in Y)\}$.

Het verschil van de verzameling X met de verzameling Y is de verzameling van alle elementen van X, die niet tot Y behoren.

Notatie: $X \setminus Y$. Dus $X \setminus Y = \{x \in X \mid x \notin Y\}$.

Als $Y \subset X$ en X is vast, dan heet $X \setminus Y$ ook wel het complement van Y in X notatie Y^c.

Opgave: bewijs dat $(Y_1 \cup Y_2)^c = Y_1^c \cap Y_2^c$

$$(Y_1 \cap Y_2)^c = Y_1^c \cup Y_2^c.$$

Dit zijn de zogenaamde regels van de Morgan.

Families en Collectie's van verzamelingen worden vaak genoteerd met behulp van een indexverzameling. Indexverzamelingen worden aangegeven met Griekse hoofdletters en de elementen van een indexverzameling - de indices - met de corresponderende kleine Griekse letter - eventueel met boven- of benedenindex. Zo geven we met $\mathcal{E} = \{X_\alpha \mid \alpha \in A\}$ aan dat er bij ieder element $\alpha \in A$ een verzameling X_α is gegeven en dat $\mathcal{E}$ juist de verzameling is van al deze X_α's.

Zij $\mathcal{E} = \{X_\alpha \mid \alpha \in A\}$. Dan is per definitie

$\cup\mathcal{E} = \cup\{X_\alpha \mid \alpha \in A\} = \{x \mid (\exists\alpha \in A)(x \in X_\alpha)\}$ - de vereniging van de X_α's.

$\cap\mathcal{E} = \cap\{X_\alpha \mid \alpha \in A\} = \{x \mid (\forall\alpha \in A)(x \in X_\alpha)\}$ - de doorsnede van de X_α's.

Bewijs nu zelf dat de regels van de Morgan als volgt gegeneraliseerd kunnen worden.

Zij $\mathcal{E} = \{X_\alpha \mid \alpha \in A\}$ een geindiceerde collectie deelverzamelingen van een verzameling X. Dan is

1. $X \setminus \cup\{X_\alpha \mid \alpha \in A\} = \cap\{X \setminus X_\alpha \mid \alpha \in A\}$

2. $X \setminus \cap\{X_\alpha \mid \alpha \in A\} = \cup\{X \setminus X_\alpha \mid \alpha \in A\}$.

De verzameling der natuurlijke getallen. $\mathbb{N}$
Hiermede bedoelen we de getallen 1,2,3,4, ..., dus de positieve gehele getallen.
Een fundamentele eigenschap van de natuurlijke getallen is de volgende:
Als $V \subset \mathbb{N}$ de eigenschappen bezit $1 \in V$ en $h \in V \rightarrow h+1 \in V$ dan is $V = \mathbb{N}$ (immers $1 \in V$, dus $1 + 1 = 2 \in V$; omdat $2 \in V$ is $2 + 1 = 3 \in V$, dus $3 + 1 = 4 \in V$, dus $5 \in V$, dus $6 \in V$, ... dus alle natuurlijke getallen behoren tot V). Hierop berust het bewijsprincipe van volledige inductie: Wanneer we nu een uitspraak $A(n)$ moeten bewijzen, met als variabele een natuurlijk getal, kunnen we volstaan met

1e te verifieren dat $A(1)$ geldt

2e bewijzen dat geldt $A(n) \rightarrow A(n+1)$ (dit bewijs heet de inductiestap of de stap van n op n + 1; het aannemen van $A(n)$ om dan $A(n+1)$ te bewijzen heet de inductieonderstelling)

Voorbeeld: Te bewijzen dat voor elke $k \in N$ geldt:

$$1^3 + 2^3 + 3^3 + 1 \ldots + k^3 = \tfrac{1}{4}k^2(k+1)^2$$

Bewijs: Voor k = 1 is de formule juist. Nu de inductiestap:
We moeten bewijzen dat voor elke $n \in N$ geldt:

$$[1^3 + 2^3 + \ldots + n^3 = \tfrac{1}{4}n^2(n+1)^2] \rightarrow [1^3 + 2^3 + \ldots + (n+1)^3 = \\ = \tfrac{1}{4}(n+1)^2(n+2)^2].$$

Dit volgt uit het feit dat de linkerleden $(n+1)^3$ verschillen en de rechterleden eveneens:

$$\tfrac{1}{4}(n+1)^2(n+2)^2 - \tfrac{1}{4}n^2(n+1)^2 = (n+1)^2.\tfrac{1}{4}[(n+2)^2 - n^2] = (n+1)^3.$$

Cartesisch product, relaties en functies

We beschouwen twee verzamelingen A en B en vormen alle mogelijke paren (a,b) met $a \in A$ en $b \in B$. We letten hierbij op de volgorde, dus (a,b) is iets anders dan (b,a). (tenzij a=b).Van het paar (a,b) heet a de eerste en b de tweede coördinaat. Twee paren (a,b) en c,d) zijn alleen gelijk wanneer beide coördinaten gelijk zijn, dus als a=c en b=d.

<u>Definitie</u>: De verzameling van alle paren (a,b) met $a \in A$ en $b \in B$ heet het <u>cartesisch product</u> van A en B, notatie $A \times B$.

<u>Voorbeelden</u>

Als $A = \{1,2\}$ en $B = \{2,3,4\}$, dan is $A \times B = \{(1,2),(1,3),(1,4),(2,2),(2,3),(2,4)\}$.

Als $A = \mathbb{R}$ en $B = \mathbb{R}$ dan is $A \times B$ te zien als de verzameling punten in het platte vlak, aangegeven door een rechthoekig coördinatenstelsel.

In het laatste voorbeeld is A = B. Dit komt ook in andere belangrijke gevallen voor.

Soms gebeurt het, dat we een bepaalde betrekking kunnen vinden die geldt tussen sommige elementen uit een verzameling A en sommige uit een verzameling B.

<u>Voorbeeld</u>. Als $A = \mathbb{R}$ en $B = \mathbb{R}$, dan geldt voor sommige $x \in A$ en $y \in B$ dat $x^2 + y^2 = 1$; voor andere tweetallen (x,y) geldt dit niet. We zeggen: voor sommige paren (x,y) geldt de <u>relatie</u> $x^2 + y^2 = 1$. De relatie $x^2 + y^2 = 1$ bepaalt een deel verzameling van het cartesisch product $R \times \mathbb{R}$, n.l. de deelverzameling $\{(x,y) \mid x^2 + y^2 = 1\}$.

Wanneer we in het algemeen tussen elementen van A en elementen van B een relatie R hebben, zodat voor sommige $a \in A$ en $b \in B$ geldt, dat a in relatie R tot b staat (notatie aRb), dan hoort hier evenzo een deelverzameling $\mathbb{R}$ van het cartesisch product $A \times B$ bij, nl.

$$S = \{(a,b) \mid a \, R \, b\}$$

Omgekeerd behoort bij iedere deelverzameling R van $A \times B$ een relatie

R, gedefinieerd door: a R b wanneer $(a,b) \in R$.

De relatie R en de deelverzameling R bepalen elkaar dus geheel. De uitspraak a R b betekent geheel hetzelfde als $(a,b) \in R$; we kunnen beide schrijfwijzen dus als gelijkwaardig beschouwen. Het is nu duidelijk dat we formeel een relatie als volgt kunnen invoeren.

Definitie: Een relatie tussen elementen van A en elementen van B is een deelverzameling R van $A \times B$. Notatie: we geven de relatie $(a,b) \in R$ wel aan met a R b.

Opgave. Als $A = B = \{x \in \mathbb{R} \mid 0 \leq x \leq 1\}$, schets dan de deelverzameling van $A \times B$ die bij de relatie $\leq$ hoort.
Ook bij relaties ziet men vaak dat $A = B$ gebruikt wordt. Men spreekt kortweg van een relatie "tussen elementen van A".

Definitie: Een relatie R tussen elementen van A heet een equivalentierelatie als geldt

1. Voor iedere $a \in A$ is a R a (reflexiteit)
2. Voor iedere $a,b \in A$ met a R b geldt ook b R a (symmetrie)
3. Voor iedere $a,b,c \in A$ met a R b en b R c geldt ook a R c (transitiviteit)

De verzameling van alle x met a R x (a vast) heet een equivalentie klasse. Twee equivalentieklassen zijn paarsgewijs disjunct of vallen samen. Dit laatste zullen we nu bewijzen.

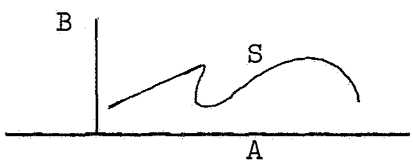
Stel A en B twee equivalentieklassen en onderstel $A \cap B \neq \emptyset$. Zij $A = \{x \mid (a,x) \in R\}$ $B = \{x \mid (b,x) \in R\}$ en $c \in A \cap B$. Dan is $(a,c) \in R$ en $(b,c) \in R$.
Op grond van de symmetrie en de transitiviteit volgt nu achtereenvolgens $(b,c) \in R \rightarrow (c,b) \in R$ en met $(a,c) \in \mathbb{R}$ geeft dit $(a,b) \in R$. Als nu $d \in A$, dan is $(a,d) \in R$. Uit de transitiviteit en symmetrie van R volgt nu uit $(a,b) \in R$ en $(a,d) \in R$ dat $(b,d) \in R$ dus $d \in B$ en dus volgt uit de willekeurigheid van d $A \subset B$. Analoog bewijst met $B \subset A$. Dus $A = B$.

Voorbeeld. Als $A = \mathbb{Z}$, dan is de relatie S bepaald door $(a,b) \in S$, wanneer $a - b$ deelbaar is door p (p vast) natuurlijk getal, een equivalentierelatie.
Verder zijn bekende voorbeelden van equivalentierelaties:

1. Congruentie-relatie in vlakke meetkunde en steriometrie
2. Gelijkvormigheidsrelatie in de meetkunde
3. Laat A de verzameling van rechte lijnen in de ruimte zijn. Voor l en $m \in A$ definieren we $(l,m) \in R$ als l evenwijdig is met m of als l samenvalt met m. De R equivalentieklassen worden richtingen genoemd.

Een belangrijke toepassing van het begrip equivalentierelatie is de definitie van kardinaalgetal wat in de volgende paragraaf behandeld wordt.

Laten nu A en B willekeurige verzamelingen zijn. Beschouw een relatie $R \subset A \times B$ met de eigenschap dat bij ieder element $a \in A$ precies één $b \in B$ voorkomt met $(a,b) \in R$.
Ieder element $a \in A$ staat dus in relatie met precies één element $b \in B$ (dat mag varieren met a).



Bij een dergelijke relatie wordt dus aan iedere $a \in A$ een bijbehorend element in B aangewezen. We kunnen zo'n relatie ook zien als een voorschrift, dat bij ieder element $a \in A$ precies één element $b \in B$ aangeeft; dit element b heet wel het beeld van a. Essentieel hierbij is, dat iedere $a \in A$ juist één beeld heeft.

Zodoende komen we tot de volgende definitie.

Definitie. Een afbeelding (functie) van een verzameling A naar een verzameling B is een relatie $F \subset A \times B$ met de eigenschap dat bij iedere $a \in A$ precies één element $b \in B$ bestaat, waarvoor geldt $(a,b) \in F$. b heet dan het beeld van a. Notatie: In plaats van $F \subset A \times B$ schrijven we meestal $f: A \to B$, $(a,b) \in F$ wordt dan $f(a) = b$.

Op grond van de beschouwing op de vorige bladzijde kunnen we een functie ook zien als volgt: Een functie of afbeelding van A naar B (notatie $f: A \to B$) is een voorschrift, dat bij ieder element a van A precies één element $b \in B$ aangeeft. Dit element heet het beeld van a (onder f). Notatie f(a).

Nog enkele notaties. De collectie $\{f(a) | a \in A\}$ heet het beeld van A en wordt genoteerd met f(A). Indien $f(A) = B$, dan heet f een op-afbeelding (surjectief). Indien uit $f(a_1) = f(a_2)$ steeds volgt $a_1 = a_2$ dan heet f ééneénduidig (injectief). Zij nu $f: A \to B$ en $g: B \to C$ afbeeldingen; we definieren nu een nieuwe afbeelding $h: A \to C$ als volgt $h(x) = g(f(x))$, $x \in A$. De afbeelding h heet de samenstelling van f en g en wordt aangeduid met $g \circ f$.

Als $f: A \to B$ zowel injectief als surjectief is, dan bestaat voor iedere $y \in B$ precies één $x \in A$ met $f(x) = y$; noem die x   g(y). De functie g heeft de eigenschap $(g \circ f)(x) = x$ en $(f \circ g)(x) = x$ voor elke in aanmerking komende x. g heet de inverse van f notatie f^{-1} en is eenduidig bij f bepaald. Tenslotte, als $f: A \to B$ een afbeelding is en $C \subset B$ dan noteert men $f^{-1}(C) = \{x \in A | f(x) \in C\}$ onafhankelijk van het feit of de inverse f^{-1} van f wel of niet bestaat.

Bij een functie $f: \mathbb{R} \to \mathbb{R}$ behoort een grafiek. Dit is een verzameling punten in het platte vlak met daarin een loodrecht assenstelsel n.l. die verzameling punten met eerste coördinaat x en tweede coördinaat f(x), waarbij we x alle reële getallen laten doorlopen. In formule

$$\{(x,y) \in \mathbb{R} \times \mathbb{R} \mid y = f(x)\}.$$

Algemeen definieren we nu: De grafiek van een afbeelding $f: A \to B$ is de deelverzameling $F \subset A \times B$, die bij de relatie f hoort. Dus

$$F = \{(a,b) \in A \times B \mid b = f(a)\}.$$

Machtigheid en kardinaalgetallen.

In de klasse van alle verzamelingen voeren we de volgende equivalentierelatie in die gelijkmachtigheid heet. We zeggen dat de verzamelingen V en W gelijkmachtig zijn (notatie $V \sim W$) als er een éénéénduidige afbeelding van V op W bestaat.

Een klasse van onderling gelijkmachtige verzamelingen heet een kardinaalgetal of machtigheid.

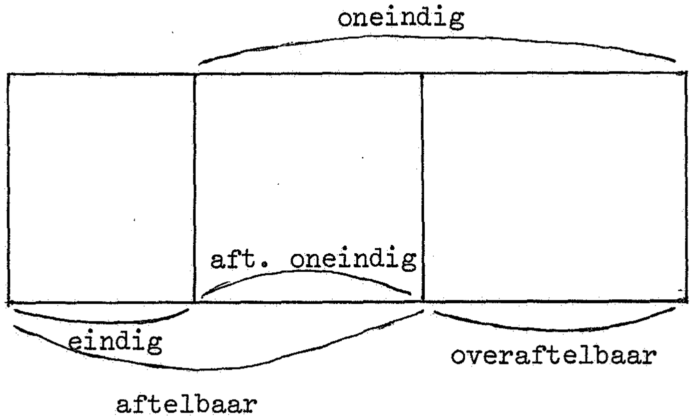
Zeer belangrijk is het begrip "eindige verzameling".
We beginnen met op te merken dat, als n en m natuurlijke getallen zijn, de verzamelingen

$$\{1,2,\ldots,n\} \quad \text{en} \quad \{1,2,\ldots,m\}$$

slechts dan gelijkmachtig zijn als $n = m$. (hoofdstelling van het tellen).
Een verzameling V heet nu eindig als er een natuurlijk getal n bestaat zodat $V \sim \{1,2,\ldots,n\}$ en de machtigheid van V wordt dan ook met de letter n aangeduid. Bovendien wordt ook de lege verzameling eindig genoemd; het betreffende kardinaalgetal wordt met 0 aangeduid.

Is V niet eindig, dan heet V oneindig. We zullen echter zien dat de oneindige verzamelingen niet allen onderling gelijkmachtig zijn. Een verzameling V is dan en slechts dan oneindig als V éénéénduidig op een echt deel van V kan worden afgebeeld. Een oneindige verzameling bevat steeds een deelverzameling die gelijkmachtig is met $\mathbb{N}$. ($\mathbb{N}$ = verzameling der natuurlijke getallen).

Een verzameling die gelijkmachtig is met $\mathbb{N}$ noemen we aftelbaar oneindig. Een verzameling die oneindig is en niet aftelbaar oneindig heet overaftelbaar.
Schematisch kunnen we de zaak als volgt weergeven



We zullen nu een niet aftelbare oneindige verzameling aangeven. We beschouwen afbeeldingen F van $\mathbb{N}$ in de verzameling $\{0,1\}$. Zo'n afbeelding is ook te schrijven door een oneindige rij van nullen en enen b.v.

$$(0,\ 0,\ 1,\ 1,\ 0,\ 1,\ 0,\ 0,\ 0,\ \ldots),$$

die onstaat door achtereenvolgens F(1), F(2), F(3).. op te schrijven. De verzameling van alle mogelijke rijen van dit type noemen we V. Duidelijk is dat V niet eindig is. Neem nu aan dat V aftelbaar is. Er is dan een ééneenduidige afbeelding van V op $\mathbb{N}$, bijvoorbeeld

$$\begin{aligned} 1 &\rightarrow (\underline{0},\ 0,\ 1,\ 1,\ 0,\ 1,\ 0,\ 0,\ \ldots\quad) = F_1 \\ 2 &\rightarrow (0,\ \underline{1},\ 0,\ 1,\ 1,\ 0,\ 1,\ 0,\ \ldots\quad) = F_2 \\ 3 &\rightarrow (1,\ 0,\ \underline{1},\ 1,\ 0,\ 0,\ 1,\ 1,\ \ldots\quad) = F_3 \end{aligned}$$

Construeer nu een nieuwe rij als volgt: Beschouw de rij

$$(o,\ 1,\ 1,\ \ldots\quad)$$

die onstaat door de onderstreepte cijfers (diagonaal) achter elkaar te zetten. (het is de rij $F_1(1)$, $F_2(2)$, $F_3(3)$, ...)
Maak uit deze een andere, door overal 0 door 1 en 1 door 0 te vervangen. Noem deze nieuwe rij F dan is F niet gelijk aan een der F_n's (F wijkt op de n-de plaats van F_n af.) Het was dus geen afbeelding van N op maar <u>in</u>, in strijd met de onderstelling.

Een belangrijke stelling is nog: De vereniging van aftelbaar vele aftelbare verzamelingen is weer aftelbaar.

Toepassing: De verzameling van alle rationale getallen is aftelbaar.

Hoofdstuk 2.

De verzameling der reële getallen $\mathbb{R}$, limietbegrip, reële functies.

1. Axiomatiek van de reële getallen.

Hieronder sommen we de grondeigenschappen op van de reële getallen $\mathbb{R}$.

a. Er bestaat een zogenaamde optelling in $\mathbb{R}$, d.w.z. voor elke $a,b \in \mathbb{R}$ bestaat de som $a + b$ die uniek bepaald is en de volgende eigenschappen bezit

L1. $(a+b)+c = a+(b+c)$ associatieve wet

L2. $a+b = b+a$ commutatieve wet

L3. Er bestaat een getal $0 \in \mathbb{R}$, uniek bepaald met de eigenschap $a+0 = 0+a = a$ voor alle $a \in \mathbb{R}$.

L4. Voor $\forall a \in \mathbb{R}$ $\exists a^* = -a \in \mathbb{R}$, uniek bepaald zodat $a+(-a) = (-a)+a = 0$.

b. Verder is er binnen $\mathbb{R}$ een z.g. vermenigvuldiging gedefinieerd. Het z.g. product ab van a en $b \in \mathbb{R}$ dat uniek bij a en b bepaald is heeft de volgende eigenschappen.

L5. $ab = ba$ commutativiteit van de vermenigvuldiging

L6. $(a+b)c = ac+bc$ distibutiviteit

L7. $\exists$ een getal $1 \in \mathbb{R}$, uniek bepaald, zodat $1.a = a.1 = a$ voor alle $a \in \mathbb{R}$.

L8. Voor iedere $a \in \mathbb{R}$, $a \neq 0$ bestaat $a^* = \frac{1}{a}$ zodat $a.\frac{1}{a} = \frac{1}{a}.a = 1$.

L9. $1 \neq 0$.

De eigenschappen L1-L9 vertellen ons dat hieruit het volgende afgeleid kan worden:
(we noteren $a.\frac{1}{a}$ met $\frac{a}{b}$).

$(-a)(-b) = ab$ $\qquad$ $(-a).b = -ab = a.(-b)$

$\frac{-a}{-b} = \frac{a}{b} = (a.\frac{1}{b})$ $\qquad$ $\frac{-a}{b} = \frac{a}{-b} = -\frac{a}{b}$.

N.B. We schrijven voor $a + (-b)$ altijd $a - b$.

Bovenstaande eigenschappen zijn de bekende rekenregels van de middelbare school.

Ordeningseigenschappen van de reële getallen.

O_1 Een deel der getallen van $\mathbb{R}$ heet positief. De verzameling der positieve getallen duiden we aan met $\mathbb{R}^+$. Als $a \in \mathbb{R}$ dan $a \in \mathbb{R}^+$, $a = 0$ of $-a \in \mathbb{R}^+$

O_2 $a \in \mathbb{R}^+,\ b \in \mathbb{R}^+ \Rightarrow ab \in \mathbb{R}^+$

O_3 $a \in \mathbb{R}^+,\ b \in \mathbb{R}^+ \Rightarrow a+b \in \mathbb{R}^+$

We spreken af dat $a > b$ betekent $a - b \in \mathbb{R}^+$, $b < a$ betekent $a > b$.

Ga na dat de volgende stellingen afgeleid kunnen worden.

Stelling 1 $a,b \in \mathbb{R} \Rightarrow a = b \quad a < b$ of $a > b$ (lineairiteit van de ordening)

Stelling 2 $a > b,\ b > c \Rightarrow a > c$ (transitiviteit)

Stelling 3 $a > 0 \Leftrightarrow -a < 0$

Bovendien gelden de volgende rekenregels die we uit O_1, O_2 en O_3 ook kunnen afleiden

$a \in \mathbb{R} \Rightarrow a^2 > 0$ of $a^2 = 0$

$1 > 0$

$a > b \Rightarrow a + c > b + c$ voor alle $c \in \mathbb{R}$

$a > b;\ c > 0 \Rightarrow ac > bc$

$a > b;\ c < 0 \Rightarrow ac < bc$

$a > 0 \Rightarrow \frac{1}{a} > 0$

$ab > 0 \Leftrightarrow$ a en b hebben hetzelfde teken

$ab < 0 \Leftrightarrow$ a en b hebben verschillende tekens.

De natuurlijke getallen $\mathbb{N}$ vormen een deelcollectie van de verzameling der reële getallen.

$\mathbb{N} = 1, 1+1, 1+1+1, 1+1+1+1, \ldots$

$\quad 1, \quad 2, \qquad 3, \qquad\quad 4,$ enz.

Hierboven merkten we reeds op dat het volgende geldt:

Als een deelcollectie van $\mathbb{R}$ het getal 1 bevat en met elk getal n ook het getal n + 1 dan omvat deze collectie juidt de verzameling der natuurlijke getallen.

Op deze eigenschap berustte het principe der volledige inductie.

Bewijs zelf de volgende vraagstukken.

1. $(a+1)^n > 1+na \quad (a > 0)$ (ongelijkheid van Bernoully)
2. $(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$ (binomiaalformule van Newton)

Hierin is $n! = 1.2.3.\ldots.n$ en $\binom{n}{k} = \frac{n!}{k!(n-k)!} \qquad 0! = 1.$

3. $\left(\sum_{i=1}^{n} a_i b_i\right)^2 \leq (\Sigma a_i^2)(\Sigma b_i^2)$ (ongelijkheid van Cauchy Swarz).

Stelling 4 $\quad m \in \mathbb{N} \quad n \in \mathbb{N} \Longrightarrow m+n \in \mathbb{N}$

$\qquad\qquad m \in \mathbb{N} \quad n \in \mathbb{N} \Longrightarrow m.n \in \mathbb{N}$

Uit bovenstaande stelling volgt dat de natuurlijke getallen een z.g. halfgroep vormen.

Stelling 5 $\quad m > n \quad m,n \in \mathbb{N} \Longrightarrow m-n \in \mathbb{N}$

Opgave. Bewijs met volledige inductie dat iedere niet lege deelverzameling der natuurlijke getallen een kleinste element bezit.

Toepassing: delen met rest, decimaalontwikkeling.

De gehele en de rationale getallen.

Definitie. De deelverzameling van $\mathbb{R}$ bestaande uit alle getallen 0, n, -n met $n \in \mathbb{N}$ noemt men de verzameling der gehele getallen. (aangeduid met $\mathbb{Z}$).

We gaan gemakkelijk na dat som, verschil, product van gehele getallen weer geheel is. Bovendien bewijst men eenvoudig dat tussen p en p+1 (p geheel) geen andere gehele getallen liggen.

Definitie. De deelverzameling van $\mathbb{R}$ bestaande uit alle getallen $\frac{p}{q}$, p, q, geheel is de verzameling der rationale getallen.

Ga na dat tussen elke twee rationale getallen r en s andere rationale getallen liggen (Bijv. $r_1 = \frac{1}{2}(r+s)$ $r_2 = \frac{1}{2}(r+r_1)$ enz. Verder is de som, verschil, product en quotiënt van twee rationale getallen weer rationaal.

Stelling 6. De verzameling der rationale getallen Q is aftelbaar.

Bewijs. We schrijven de verzameling der rationale getallen als vereniging van aftelbaar veel deelverzamelingen Q_i, i = 1, 2, ----. Hierbij is $Q_i = \frac{1}{i}, -\frac{1}{i}, \frac{2}{i}, \frac{-2}{i}, \ldots$ en schrijven de verzameling der rationale getallen als volgt op.

$$Q_1 = \frac{1}{1}, -\frac{1}{1}, +\frac{2}{1}, -\frac{2}{1}, +\frac{3}{1}, -\frac{3}{1}, \ldots$$

$$Q_2 = \frac{1}{2}, -\frac{1}{2}, +\frac{2}{2}, -\frac{2}{2}, +\frac{3}{2}, -\frac{3}{2}, \ldots$$

$$Q_3 = \frac{1}{3}, -\frac{1}{3}, +\frac{2}{3}, -\frac{2}{3}, +\frac{3}{3}, -\frac{3}{3}, \ldots$$

.

.

.

.

We lezen dit schema volgens nevendiagonalen:

$1, -\frac{1}{1}, \frac{1}{2}, \frac{2}{1}, -\frac{1}{2}, \frac{1}{3}, -\frac{2}{1}, +\frac{2}{2}, -\frac{1}{3}$

Dit is kennelijk een aftelling voor de rationale getallen Q.

Uit één van de laatste beweringen volgt nu dat de rationale getallen van $\mathbb{R}$ voldoen aan alle eigenschappen $L_1, .., L_8$ en $O_1, .., O_3$. Er rijst dus de vraag of misschien wel $Q = \mathbb{R}$ i.e. of de reële getallen wel voldoende gekarakteriseerd zijn door bovenstaande eigenschappen. Dit behandelen we in het onderstaande.

<u>Irrationale getallen; de volledigheidseigenschap van $\mathbb{R}$.</u>

We laten eerst aan de hand van een voorbeeld zien dat de rationale getallen nog een belangrijke eigenschap missen.

<u>Steeling 7</u>. Er bestaat geen rationaal getal r met $r^2 = 2$. M. a. w. niet elke kwadratische vergelijking is binnen Q oplosbaar.

<u>Bewijs</u>. Zij $r = \frac{p}{q}$ en p en q geheel. Veronderstel dat toch r aan de vergelijking $r^2 = 2$ voldoet. We mogen onderstellen $p,q > 0$; q minimaal. We hebben dus $p^2 = 2q^2$. Dus $p > q$ en $p < 2q$. Dan is

$$\left(\frac{2q-p}{p-q}\right)^2 = \frac{4q^2-4pq+p^2}{p^2-2pq+q^2} = \frac{6q^2-4pq}{3q^2-2pq} = 2.$$ Dit is een tegenspraak

want $(2q-p) > 0$; $0 < p-q < q$.

We weten dat binnen de reële getallen dergelijke kwadratische vergelijkingen oplosbaar zijn. Dit is een direkt gevolg van het feit dat de reële getallen nog aan de volgende grondeigenschap voldoen.

Onder een <u>bovengrens</u> van $A \subset \mathbb{R}$ verstaan we een reëel getal $a \in \mathbb{R}$ met $a \geq x$ voor alle $x \in A$. Analoog heet $b \in \mathbb{R}$ een <u>benedengrens</u> van A als $b \leq x$ voor alle $x \in A$. In het geval dat A werkelijk een bovengrens bezit, dan heet A <u>naar boven begrensd</u>, is er sprake van een benedengrens, dan heet A <u>naar beneden begrensd</u>. In het geval dat A zowel naar boven als naar beneden begrensd is, heet A <u>begrensd</u>.

Als de verzameling bovengrenzen van A een kleinste element bezit dan heet deze bovengrens het <u>supremum</u> van A; notatie a = sup A (merk op dat a éénduidig bepaald is). Analoog heet b het <u>infimum</u> van A (notatie b = inf A) als b de grootste benedengrens is van A.

<u>Grondeigenschap van $\mathbb{R}$</u>. Voor iedere naar boven begrensde verzameling $A \subset \mathbb{R}$ bestaat a = sup A; voor iedere naar beneden begrensde verzameling $A \subset \mathbb{R}$ bestaat b = inf A.

Met behulp van deze grondeigenschap van $\mathbb{R}$ is het mogelijk om te bewijzen dat de vergelijking $r^2 = 2$ binnen $\mathbb{R}$ een oplossing bezit.

Immers, beschouw $A = \{r \in \mathbb{R} \mid r^2 < 2\}$.
Blijkbaar bestaat sup A = a want Ais naar boven begrensd. (Ga na).
Stel nu eerst $a^2 < 2$. Zij $b = 2.\frac{a+1}{a+2}$.

Dan $b^2-2 = \frac{4(a^2+2a+1)}{a^2+4a+4} - 2 = \frac{2(a^2-2)}{(a+2)^2} < 0$ dus $b^2 < 2$

en $b-a = 2.\frac{a+1}{a+2} - a = \frac{2-a^2}{a+2} > 0$, dus $b > a$. Dus $b \in A$ en $b > a$.
Dit is in tegenspraak met het feit dat a bovengrens is van A.
Onderstel nu $a^2 > 2$. Blijkbaar geldt nu $b^2 > 2$ en $b < a$ i.e. b is een kleinere bovengrens van A dan a. Dit is ook onmogelijk, dus $a^2 = 2$, m.a.w. a^2 is een oplossing van de vergelijking $r^2 = 2$.

De volgende stelling is intuïtief reeds duidelijk.
<u>Stelling 8</u>. (Eudoxus, Archimedes). Als $b, c \in \mathbb{R}$ $b > 0$, dan bestaat er een natuurlijk getal n met $nb > C$.

<u>Bewijs</u>. Onderstel dat voor alle n = 1, 2, .. $nb \leq C$. Laat $A = \{nb \mid n = 1, 2, ..\}$. A is naar boven begrensd dus a = sup A bestaat; Er is dus een natuurlijk getal m zodat $a-b < mb$(anders zou gelden $a-b \geq nb$ voor alle n = 1, 2, .. en dan zou a-b een kleinere bovengrens zijn van A dan a omdat $b > 0$). Nu volgt $a < (m+1)b$; echter in tegenspraak met onze onderstelling $a \geq nb$ voor alle n = 1, 2, .. .

§2. Rijen

Modulus: de modulus (of absolute waarde) van $a \in \mathbb{R}$ is per definitie a indien $a \geq 0$ en $-a$ indien $a \leq 0$. Notatie $|a|$.

Eigenschappen:

1) $|a| \geq 0$ en $= 0 \iff a = 0$
2) $|ab| = |a|.|b|$
3) $|a+b| \leq |a|+|b|$ driehoeksongelijkheid)

Bewijs van 3) $a \leq |a|$ $b \leq |b|$ dus $(a+b) \leq |a|+|b|$
evenzo $-(a+b) \leq |a|+|b|$.

Bewijs zelf:

$$|a-b| \geq \big||a|-|b|\big| \qquad |-x| = |x|$$

$|x| \leq a$ is gelijkwaardig met $-a \leq x \leq a$.

Interval. Laten a,b getallen zijn met $a < b$.

Dan (open) interval (a,b): verz. der x met $a < x < b$
(gesloten) interval $[a,b]$: verz. der x met $a \leq x \leq b$
linksgesloten interval $[a,b)$: verz. der x met $a \leq x < b$
rechtsgesloten interval $(a,b]$: verz. der x met $a < x \leq b$
omgeving van a : open interval (c,d) met $a \in (c,d)$

Rij getallen: een afbeelding van de verzameling der natuurlijke getallen $\mathbb{N}$ naar de reële getallen. Schrijfwijze $\{a_n\}_{n=1}^{\infty}$

Een rij reële getallen is dus gekarakteriseerd als een zeker voorschrift die aan elk natuurlijk getal n een reeël getal a_n toevoegt.

Limiet van een rij.

Definitie. Zij $\{a_n\}_{n=1}^{\infty}$ een rij getallen. Dan heet a de limiet van de rij $\{a_n\}_{n=1}^{\infty}$, schrijfwijze $a = \lim\limits_{n\to\infty} a_n$ als bij elk getal $\varepsilon > 0$ een index n_0 bestaat zodat geldt: als $n > n_0$ dan $|a_n - a| < \varepsilon$.

Deze definitie zegt zo ongeveer: hoe we ook een omgeving van a kiezen, op den duur ligt a_n erin. De index n_0 is i.h.a. afhankelijk van de keuze van ε.
Verder merken we op dat $\lim\limits_{n\to\infty} a_n$ eenduidig bepaald is (als deze bestaat).

Bepaalde rijen hebben limieten. Deze rijen heten <u>convergent</u>.
Beschouw de rij $\{\frac{1}{n}\}_{n=1}^{\infty}$, deze heeft limiet 0. De rij $\{\frac{1}{n^2}\}$ heeft ook limiet 0 en ook de rij $\{\frac{1}{2^n}\}_{n=1}^{\infty}$. Dit laatste zullen we nu streng bewijzen.

Zij $\varepsilon > 0$ en beschouw n_0 als zijnde een geheel getal groter dan $\frac{1}{\varepsilon}$ (stelling van Exodus!) dan geldt voor elke $n > n_0$ $\frac{1}{2^n} < \frac{1}{n}$(gebruik de ongelijkheid van Bernouilli)$< \frac{1}{n_0} < \varepsilon$.
Dus aan de definitie van limiet is voldoen voor $a = 0$.
Hieruit volgt $\lim\limits_{n\to\infty} \frac{1}{2^n} = 0$.

Voor limieten van rijen gelden de volgende stellingen.

1) $\lim\limits_{n\to\infty} a_n \cdot \lim\limits_{n\to\infty} b_n = \lim\limits_{n\to\infty} a_n b_n$

2) $\lim\limits_{n\to\infty} a_n + \lim\limits_{n\to\infty} b_n = \lim\limits_{n\to\infty} (a_n+b_n)$

3) Als $c \in \mathbb{R}$ dan $c\,(\lim\limits_{n\to\infty} a_n) = \lim\limits_{n\to\infty} ca_n$.

4) Als $\lim\limits_{n\to\infty} a_n \neq 0$ dan $\lim\limits_{n\to\infty} \frac{1}{a_n} = \frac{1}{\lim\limits_{n\to\infty} a_n}$.

We zullen 1) bewijzen. Doe 2), 3) en 4) zelf. Stel $\lim\limits_{n\to\infty} a_n = a$ $\lim\limits_{n\to\infty} b_n = b$. Zij ε een willekeurig getal >0. Neem n_1, zodat $|b_n-b| < 1$ voor $n > n_1$ (definitie toegepast met $\varepsilon=1$). Dan is $|b_n| = |b+b_n-b| < |b|+1$ voor $n > n_1$, dus $|a_nb_n-ab| = |a_nb_n-ab_n+ab_n-ab| \leq |a_n-a|.|b_n|+|a|.|b_n-b| \leq K\{|a_n-a|+|b_n-b|\}$voor $n > n_1$ als we stellen $K = \max\,(|b|+1,|a|)$. neem nu n_2, n_3 met $|a_n-a| < \frac{\varepsilon}{2K}$ voor $n > n_2$, $|b_n-b| < \frac{\varepsilon}{2}$ voor $n > n_3$.

Voor $n > n_0 = \max(n_1, n_2, n_3)$ is dan

$$|a_n b_n - ab| < K.\left(\frac{\varepsilon}{2K} + \frac{\varepsilon}{2K}\right) = \varepsilon.$$

Hieruit volgt het gestelde.

Opgave: Bewijs ook dat $\lim\limits_{n\to\infty} \frac{a_n}{b_n} = \frac{\lim\limits_{n\to\infty} a_n}{\lim\limits_{n\to\infty} b_n}$ $(b_n \neq 0)$

<u>Definitie</u>. Een rij $\{a_n\}_{n=1}^{\infty}$ heet <u>monotoon stijgend</u> (<u>monotoon niet dalend</u>) indien $a_1 < a_2 < a_3 < \ldots < a_n < \ldots$ $(a_1 \leq a_2 \leq a_3 \leq \ldots \leq a_n \leq \ldots)$ en <u>monotoon dalend</u> (<u>monotoon niet stijgend</u>) indien $a_1 > a_2 > a_3 > \ldots > a_n > ..$ $(a_1 \geq a_2 \geq a_3 \geq \ldots a_n \geq \ldots)$.

Voor monotone rijen geldt de volgende stelling.

<u>Stelling 1</u>.

<u>a</u>. Iedere monotoon stijgende (monotoon niet dalende) naar boven begrensde rij $\{a_n\}_{n=1}^{\infty}$ heeft een limiet.

<u>b</u>. Iedere monotoon dalende (monotoon niet stijgende) naar beneden begrensde rij heeft een limiet.

(merk op dat de rij $\{a_n\}_{n=1}^{\infty}$ naar boven begrensd heet indien de verzameling $\{a_n | n = 1,2,\ldots\}$ van reële getallen naar boven begrensd is, analoog naar beneden begrensd).

Bewijs van <u>a</u>. Blijkbaar bestaat $\sup\{a_n | n=1,2,\ldots\} = a$.
We zullen aantonen dat $a = \lim\limits_{n\to\infty} a_n$. Stel $\varepsilon > 0$. Omdat a kleinste bovengrens is van $\{a_n|n=1,2,\ldots\}$ is $a-\varepsilon/2$ geen bovengrens van deze verzameling; d.w.z. voor zekere n_0 geldt $a-\varepsilon/2 \leq a_{n_0}$. Als n een willekeurige index is met $n > n_0$ dan geldt $|a-a_n| = a-a_n \leq a-a_{n_0} \leq \varepsilon/2 < \varepsilon$ (want de rij is monotoon). Dus geldt inderdaad volgens de definitie van limiet $\lim\limits_{n\to\infty} a_n = a$.

<u>Toepassing</u>. We onderzoeken nu de volgende rij

$$e_n = \left(1 + \frac{1}{n}\right)^n \qquad n = 1,2,\ldots$$

We zullen aantonen dat $\lim\limits_{n\to\infty} e_n$ bestaat. Per definitie noemen we de limiet dan e.

Als n = 1,2,3,4 dan $e_1 = 2$, $e_2 = 2{,}25$, $e_3 = 2{,}37$, $e_4 = 2{,}45$ enz.

Het valt op dat de termen steeds groter worden en onder de drie blijven.

Algemeen tonen we aan

1) $e_n < e_{n+1}$ $n = 1,2,\ldots$ 2) $e_n < 3$ $n = 1,2,\ldots$

Uit de vorige stelling volgt dan dat $\lim\limits_{n\to\infty} e_n$ bestaat.

We moeten bewijzen

$$(1 + \frac{1}{n})^n < (1 + \frac{1}{n+1})^{n+1}$$

Ofwel $$1 < \frac{(1 + \frac{1}{n+1})^{n+1}}{(1 + \frac{1}{n})^n} = \frac{(\frac{n+2}{n+1})^{n+1}}{(\frac{n+1}{n})^n} = (\frac{n+2}{n+1}) \frac{(\frac{n+2}{n+1})^n}{(\frac{n+1}{n})^n} =$$

$$= \frac{n+2}{n+1}. \quad \left[\frac{n(n+2)}{(n+1)^2}\right]^n$$

Blijft dus te bewijzen $$\left[\frac{n(n+2)}{(n+1)^2}\right]^n > \frac{n+1}{n+2}$$

Nu geldt voor $h > -1$ dat $(1+h)^n \geq 1+nh$ (ongelijkheid van Bernouilli).

Pas dit toe op

$$\left[\frac{n(n+2)}{(n+1)^2}\right]^n = \left[\frac{n^2+2n+1-1}{(n+1)^2}\right]^n = \left[1 + \frac{-1}{(n+1)^2}\right]^n = (1 + h)^n$$

met $h = -\frac{1}{(n+1)^2}$. Dit is $\geq 1 + nh = 1 - \frac{n}{(n+1)^2} = \frac{n^2+n+1}{n^2+2n+1}$.

Dus aan te tonen $$\frac{n^2+n+1}{n^2+2n+1} > \frac{n+1}{n+2}.$$

Hetgeen neerkomt op $(n^2+n+1)(n+2) > (n+1)(n^2+2n+1)$

Hetgeen equivalent is met

$$n^3+3n^2+3n+2 > n^3+3n^2+3n+1.$$

Hieraan is altijd voldaan, <u>dus bewezen is dat</u> $e_{n+1} > e_n$.

We zullen nu aantonen dat $e_n < 3$ voor alle n = 1,2,...

We passen het binomium van Newton toe

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^{n-k} b^k$$

Neem $a = 1$, $b = \frac{1}{n}$, dan blijkt dat

$$(1+\frac{1}{n})^n = \sum_{k=0}^{n} \binom{n}{k} 1^{n-k}(\frac{1}{n})^k = \sum_{k=0}^{n} \binom{n}{k} \frac{1}{n^k} = 1 + \binom{n}{1}\frac{1}{n} + \binom{n}{2}\frac{1}{n^2} +$$

$$\ldots\ldots \binom{n}{n}\frac{1}{n^n} = 1 + 1 + \frac{n(n-1)}{2!\, n^2} + \frac{n(n-1)(n-2)}{3!\, n^3} + \ldots \frac{n(n-1)\ldots 3.2.1}{n(n-1)\ldots 3.2.1}\,\frac{1}{n^n}$$

$$= 1 + 1 + \frac{1}{2!}(1-\frac{1}{n}) + \frac{1}{3!}(1-\frac{1}{n})(1-\frac{2}{n}) + \ldots \frac{1}{n!}(1-\frac{1}{n})(1-\frac{2}{n})\ldots(1-\frac{n-1}{n}).$$

We kunnen dus zeggen

$$(1+\frac{1}{n})^n < 1 + 1 + \frac{1}{2!} + \frac{1}{3!} + \frac{1}{4!} + \ldots + \frac{1}{n!} =$$

$$= 1 + 1 + \frac{1}{1.2} + \frac{1}{1.2.3} + \frac{1}{1.2.3.4} + \ldots \frac{1}{1.2.\ldots.n}$$

$$< 1 + 1 + \frac{1}{2} + \frac{1}{2.3} + \frac{1}{3.4} + \frac{1}{4.5} + \ldots + \frac{1}{(n-1)n} =$$

$$= 1 + 1 + \frac{1}{2} + (\frac{1}{2}-\frac{1}{3}) + (\frac{1}{3}-\frac{1}{4}) + \ldots + (\frac{1}{n-1}-\frac{1}{n}) = 3 - \frac{1}{n} < 3.$$

Hiermede is bewezen dat $\{e_n\}_{n=1}^{\infty}$ begrensd is en $\{e_n\}_{n=1}^{\infty}$ is monotoon.
Dus $\lim_{n\to\infty} e_n = e$ bestaat. Bij benadering geldt

$$e = 2{,}718281828459045\ \ldots$$

<u>We beschouwen nu de volgende rij</u>

$$a_1 = 1, \quad a_{n+1} = \frac{1}{2}(a_n + \frac{c}{a_n}) \qquad c > 0$$

We merken op dat als $a, b > 0$ dan geldt $\frac{a+b}{2} \geq \sqrt{ab}$
Dit toegepast met $a = a_n$ $b = \frac{c}{a_n}$ geeft

$$a_{n+1} \geq \sqrt{a_n \frac{c}{a_n}} = \sqrt{c} \qquad n > 1$$

Dus $a_1 = 1$ en $a_n \geq \sqrt{c}$ voor $n > 1$

De rij $\{a_n\}_{n=1}^{\infty}$ is dus naar beneden begrensd

$$\frac{a_{n+1}}{a_n} = \frac{1}{2}\left(1 + \frac{c}{a_n^2}\right) \leq \frac{1}{2}\left(1 + \frac{c}{(\sqrt{c})^2}\right) \neq 1$$

Dus $\{a_n\}$ is monotoon niet stijgend d.w.z. $\lim\limits_{n\to\infty} a_n$ bestaat.

Zij $L = \lim\limits_{n\to\infty} a_n$. Kennelijk geldt nu $L = \frac{1}{2}(L + \frac{c}{L})$ d.w.z.

$L^2 - c = 0$. Dus $L = \;\; c$.

Dit geeft ons een methode om wortels expliciet te berekenen.

<u>Stelling 2</u>. (Cauchy).Een rij $\{a_n\}_{n=1}^{\infty}$ van reële getallen heeft dan en slechts dan een limiet wanneer voor ieder positief getal ε een index N bestaat zodat voor $n,m > N$ steeds $|a_n - a_m| < \varepsilon$ geldt. (Een rij met deze eigenschap heet een fundamentaal rij).

<u>Bewijs</u>. Onderstel dat de rij $\{a_n\}_{n=1}^{\infty}$ aan bovenstaande voorwaarden voldoet. Beschouw voor $n = 1,2,..$ $b_n = \sup \{a_k | k > n\}$. Blijkbaar is $\{b_n\}_{n=1}^{\infty}$ een nieuwe rij getallen en deze is monotoon niet stijgend. De rij $\{b_n\}_{n=1}^{\infty}$ is ook naar beneden begrensd, want anders was voor iedere $k \in \mathbb{N}$ een index $n(k)$ te vinden zodat $b_{n(k)} < b_{n(k-1)} - 1$. Het is duidelijk dat dan nooit aan bovenstaande voorwaarden voldaan kan zijn. Blijkbaar bestaat nu volgens de voorgaande stelling $a = \lim\limits_{n\to\infty} b_n$. We zullen aantonen dat geldt $\lim\limits_{n\to\infty} a_n$ m.a.w. dat de rij $\{a_n\}_{n=1}^{\infty}$ convergeert. Zij ε een positief getal. Er bestaat nu een index n zodat $|a-b_n| < \varepsilon/3$ voor $n > n_1$. Kies ook voor elke n een index $k(n) > n$ zodat $b_n < a_{k(n)} + \varepsilon/3$ d.w.z. $|b_n - a_{k(n)}| < \varepsilon/3$. Kies tenslotte een index N zodat $|a_n - a_m| < \varepsilon/3$ voor $n,m > N$. Zij $n_0 = \max(n_1, N)$; dan geldt voor $n>n_0$
$|a - a_n| \leq |a - b_n| + |b_n - a_{k(n)}| + |a_{k(n)} - a_n| < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon$.
Omgekeerd is het vrijwel direct duidelijk dat als een rij $\{a_n\}_{n=1}^{\infty}$ convergent is dat dan aan bovenstaande voorwaarden voldaan is.

Inderdaad, als $\varepsilon > 0$ en n_0 is zo gekozen dat $|a-a_n| < \varepsilon/2$ voor $n > n_0$, dan geldt voor $n,m > n_0$ $|a_n-a_m| < |a_n-a| + |a_m-a| \le \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$.

We geven hieronder een paar voorbeelden die in feite toepassingen zijn van bovenstaande stellingen.

Zij X een deelverzameling van de reële getallen $\mathbb{R}$ dan heet $a \in \mathbb{R}$ een verdichtingspunt van X wanneer iedere omgeving (d.w.z. ieder interval $(a-\varepsilon, a+\varepsilon)$ met $\varepsilon > 0$) punten van X verschillend van a bevat.

Voorbeelden. 1

1 0 is v.d.p. van de verzameling $\{x | 0 \le x \le 1\}$ en ook elk getal α met $0 < \alpha \le 1$ is v.d.p. van deze verzameling.

2 0 is het enige verdichtingspunt van de verzameling $\{\frac{1}{n} \mid n = 1,2, \ldots\}$.

Voor het gevoel is een v.d.p. van een verzameling een punt van $\mathbb{R}$ dat door punten uit de verzameling willekeurig goed benaderd kan worden.

Complexe getallen.

Voor verschillende doeleinden voeren we het systeem der complexe getallen in.

Zij $\mathbb{C} = \mathbb{R} \times \mathbb{R}$, het cartesisch product van $\mathbb{R}$ met zichzelf. $\mathbb{C}$ stelt dus de punten van het platte vlak voor. Door op een geschikte manier een optelling en vermenigvuldiging op $\mathbb{C}$ te definieren (uitgaande van de operaties die op $\mathbb{R}$ gelden) verkrijgen we de z.g. complexe getallen.

We definieren de som van twee paren (a,b) en (c,d) als volgt

$$(a,b) + (c,d) = (a+c,\ b+d).$$

De vermenigvuldiging wordt aldus gedefinieerd

$$(a,b) \times (c,d) = (ac-bd,\ bc+ad)$$

Dan blijkt dat de volgende eigenschappen gelden:
Stel z_1, z_2, z_3 willekeurige elementen van $\mathbb{C}$ en stel $0 = (0,0)$
$(1,0) = 1$. Dan geldt:

1) $z_1 + z_2 = z_2 + z_1$ commutativiteit

2) $z_1 + (z_2+z_3) = (z_1+z_2) + z_3$

3) $0 + z = z + 0 = z_1$ voor alle $z \in \mathbb{C}$

4) Voor alle $z \in \mathbb{C}$ bestaat $z^* \in \mathbb{C}$ met $z + z^* = z^* + z = 0$
(Als $z = (a,b)$ neem dan voor $z^* = (-a,-b)$. I.p.v. z^* schrijven we $-z$.

5) $z_1 . z_2 = z_2 . z_1$

6) $z_1(z_2.z_3) = (z_1.z_2)z_3$

7) $1.z_1 = z$ voor alle $z \in \mathbb{C}$

8) Als $z \neq 0$ dan bestaat $z' \in \mathbb{C}$ zodat $zz' = 1$ (als $z = (a,b)$ neem dan voor $z' = (\frac{a}{a^2+b^2}, \frac{-b}{a^2+b^2}))$. I.p.v. z' schrijven we z^{-1}.

Met deze definities voldoet het systeem $\mathbb{C}$ aan alle algebraïsche eigenschappen die ook reeds voor $\mathbb{R}$ golden.

We noteren de elementen $(a,0) \in \mathbb{C}$ gewoon met a. Dus hierdoor wordt $\mathbb{R}$ een deelverzameling van $\mathbb{C}$.
Merk op dat optelling en vermenigvuldiging van elementen uit de collectie $\{(a,0)|a \in \mathbb{R}\} \subset \mathbb{C}$ overeenkomt met optelling en vermenigvuldiging met elementen uit $\mathbb{R}$.

Als $z = (a,b)$ dan $z = (a,0) + (0,b) = (a,0) + (b,0).(0,1)$; dan volgt $z = (a,b) = a + bi$ met a,b reëel. Ieder complex getal is dus op deze wijze voor te stellen.

b heet het imaginaire gedeelte van z notatie Imz
a heet het reële gedeelte van z. notatie Rez

$\sqrt{a^2 + b^2}$ heet de modulus van z. Notatie $|z|$. Als $z \in \mathbb{C}$ en $z = (a,b)$ dan definieren we $\bar{z}$ als $(a,-b)$.

<u>Propositie</u> $z\bar{z} = |z|^2$ voor $z \in \mathbb{C}$

<u>Propositie</u> $z + \bar{z} = 2$ Rez; $z - \bar{z} = 2$ Imz.

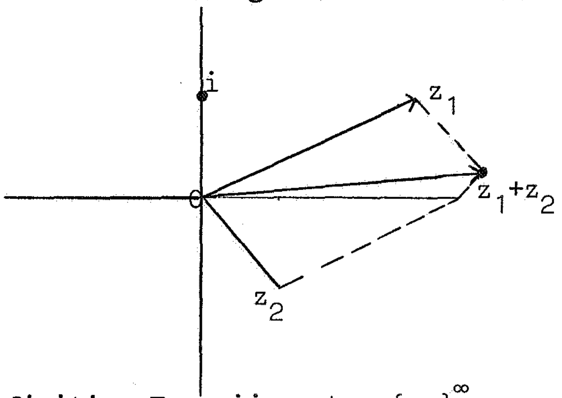
<u>Propositie</u> $|\ |$ voldoet aan de volgende eigenschappen

<u>a</u>. $|z| \geq 0 \quad \forall z \in \mathbb{C}$ en $= 0 \Longleftrightarrow z = 0$

<u>b</u>. $|z_1+z_2| \leq |z_1| + |z_2|$ voor $\forall z_1, z_2 \in \mathbb{C}$

<u>c</u>. $|z_1 z_2| = |z_1|\ |z_2|$.

Schematisch verkrijgen we de optelling van elementen uit $\mathbb{C}$ door samenstelling van vectoren. De lengte van de "vector" z is precies $|z|$. Hieruit volgt planimetrisch de vorige propositie. We definieren het <u>argument</u> van z als bg tg $\frac{b}{a}$.



<u>Definitie</u>. Een rij punten $\{z_n\}_{n=1}^{\infty}$ van het complexe vlak heet een <u>fundamentaalrij</u> indien $|z_n - z_m| < \varepsilon$ voor $n,m > N_0$.

Nu geldt de volgende stelling die een generalisatie is van de stelling van Cauchy voor de reële rechte.

<u>Stelling.</u> Elke fundamentaalrij in $\mathbb{C}$ convergeert in $\mathbb{C}$. M.a.w.: Als $\{z_n\}_{n=1}^{\infty}$ een f. rij is in $\mathbb{C}$ dan bestaat $z \in \mathbb{C}$ (eenduidig bepaald) zodat $|z - z_n| < \varepsilon$ voor $n > n_0$.

De overgang van $\mathbb{R}$ naar $\mathbb{C}$ heeft enige consequenties waar we overigens nu niet op ingaan.
I.h.b. blijkt het nu <u>niet</u> mogelijk een ordening op $\mathbb{C}$ in te voeren die analooge eigenschappen bezit als de ordening die op $\mathbb{R}$ bestond.
Dit heeft zekere consequenties.

§3. Reële functies, continuiteit.

Definitie. Zij X een verzameling van reële getallen. Dan is een reële functie op X niets anders dan een afbeelding f van X naar $\mathbb{R}$. Slordig noteren we i.p.v. f ook $f(x)$.

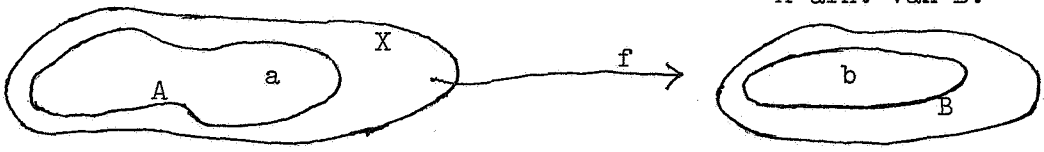
Beschouw $f(x) = \sqrt{1-x}$ (later zullen we het begrip vierkantswortel exact formuleren).
Dan geldt dat er alleen een voorschrift is voor $x \leq 1$, want voor $x > 1$ is er geen reële waarde waar $f(x)$ gedefinieerd is. Dus deze functie is gedefinieerd op $X = \{x \in \mathbb{R} \mid x \leq 1\}$. Wil men f definieren op de gehele reële rechte dan dient men nog te definieren welke waarden $f(x)$ aanneemt voor $x > 1$.

Een rij is nu niets anders dan een functie gedefinieerd op de verzameling der natuurlijke getallen, immers voor elk natuurlijk getal n is er een voorschrift dat aan n een reeël getal toevoegt. Afgezien van dit geval zal X in de regel een interval zijn.

Als f en g functies zijn die op $X \subset \mathbb{R}$ gedefinieerd zijn, dan noemen we $f + g$ die functie op X die gedefinieerd is door $(f+g)(x) = f(x)+g(x)$. Analoog definieren we het product $f.g.$ van f en g door $(f.g)(x) = f(x)g(x)$. Ook kan men als $g(x) \neq 0$ voor $x \in X$ definieren, het quotiënt van f en g n.l. $\frac{f}{g}$, door $\frac{f}{g}(x) = \frac{f(x)}{g(x)}$.
Als f gedefinieerd is op X en g gedefinieerd is op $Y \supset f(X)$ dan bestaat ook de samenstelling $g \circ f$ van f en g (zie vroeger).

Zij nu een functie $f(x)$ gedefinieerd op $X \subset \mathbb{R}$ en zij a een verdichtingspunt van X. Dan zeggen we dat $f(x)$ in a de limiet b heeft en we schrijven $\lim_{x \to a} f(x) = b$ indien geldt: Bij elke omgeving B van b bestaat een omgeving A van a zodat uit $x \in A \cap X$, $x \neq a$ steeds volgt $f(x) \in B$. Schematisch is voor elke B de situatie als volgt

A afh. van B.



Anders gezegd: $\lim\limits_{x\to a} f(x) = b$ indien voor elk positief getal ε een positief getal δ bestaat zodat $0 < |x-a| < \delta$, $x \in X$, impliceert $|f(x)-b| < \varepsilon$.

Als $\lim\limits_{x\to a} f(x)$ bestaat, dan is deze eenduidig bepaald.

Nog enkele begrippen.

$\lim\limits_{x\downarrow a} f(x) = b$ (d.w.z. $f(x)$ heeft een rechterlimiet b in a)

indien voor $\forall \varepsilon>0 \quad \exists \delta>0$ zodat $a < x < a + \delta$ impliceert $|f(x)-b| < \varepsilon$

$\lim\limits_{x\uparrow a} f(x) = b$ (d.w.z. $f(x)$ heeft een linkerlimiet b in a)

indien voor $\forall \varepsilon>0 \quad \exists \delta>0$ zodat $a\delta < x < a$ impliceert $|f(x)-b| < \varepsilon$

$\lim\limits_{x\to\infty} f(x) = b$ (d.w.z. $f(x)$ heeft voor $x \to \infty$ de limiet b)

indien $\forall \varepsilon>0 \, \exists M > 0$ zodat $|x| > M$ impliceert $|f(x)-b| < \varepsilon$.

Bewijs zelf de volgende belangrijke stelling.

Stelling 1. Zij $X \subset R$ en zij a een verdichtingspunt van X. Dan geldt de volgende equivalentie

(i) $\lim\limits_{x\to a} f(x) = b$

(ii) Voor alle rijen $\{x_i\}_{i=1}^{\infty}$ met $\lim\limits_{i\to\infty} x_i = x$, $x_i \neq x$ geldt $\lim\limits_{i\to\infty} f(x_i) = b$

We hebben voor limieten van functies de volgende rekenregels:

Stelling 2. Laten twee functies $f(x)$ en $g(x)$ beide gedefinieerd zijn op een verzameling X, en a een verdichtingspunt van X, en laten f(x) en g(x) in a de eindige limieten b resp. c hebben. Dan geldt

I. $\lim_{x\to a} \{f(x) + g(x)\} = b + c$

II. $\lim_{x\to a} f(x) \,.\, g(x) = b.c$

III. $\lim_{x\to a} \frac{1}{f(x)} = \frac{1}{b}$, als $b \neq 0$.

Bewijs van II. Zij $\varepsilon > 0$ willekeurig en zij $K = \max(|b| + 1, |c|)$
We hebben

$$|f(x)g(x) - bc| = |f(x)g(x) - f(x).c + f(x).c - bc|$$
$$\leq |f(x)| \,.\, |g(x) - c| + |c| \,.\, |f(x) - b|.$$

Omdat $\lim_{x\to a} f(x) = b$ en $\lim_{x\to a} g(x) = c$ kunnen we (door de doorsnede van enige omgevingen te nemen) een omgeving A van a vinden met de volgende eigenschappen:
als $x \in A$, $x \in X$, en $x \neq a$.

$$|f(x) - b| < 1, \text{ dus } |f(x)| \leq |b| + |f(x) - b| < |b| + 1$$

en verder $|f(x) - b| < \frac{\varepsilon}{2K}$ $|g(x) - c| < \frac{\varepsilon}{2K}$.

Voor zulke x is dan ook

$$|f(x)g(x) - bc| \leq K|g(x) - c| + K|f(x) - b| < K \,.\, \frac{\varepsilon}{2K} + K\,\frac{\varepsilon}{2K} = \varepsilon.$$

Hieruit volgt het gestelde.

Definitie. Stel f een functie gedefinieerd op $X \subset \mathbb{R}$ en $a \in X$. Dan heet f continu in het punt $a \in X$ indien

$$\forall \varepsilon > 0 \quad \exists \delta > 0 \text{ zodat } |x - a| < \delta \Rightarrow |f(x) - f(a)| < \varepsilon.$$

Indien f continu is in elk punt $x \in X$ dan heet f continu op X.

Opmerking. Indien a een v.d.p. is van X dan is blijkbaar continuiteit in a precies hetzelfde als de eigenschap $\lim_{x\to a} f(x) = f(a)$.

Stelling 3. De som, het product en het quotiënt van twee continue functies gedefinieerd op $X \subset \mathbb{R}$ is weer continu op X. De samenstelling van twee continue functies is weer continu.

Bewijs. Het eerste gedeelte volgt direct uit de vorige stelling. Het bewijs van het tweede gedeelte wordt aan de lezer overgelaten.

Toepassingen 1. De functie $f(x) = x$ is gedefinieerd op $\mathbb{R}$ en overal continu. Het is de identieke afbeelding van $\mathbb{R}$ op $\mathbb{R}$. Ook volgt dat i.h.a. $f(x) = x^n (n = 1,2,..)$ continu is. Immers deze functie is het product van de continue functies $\underbrace{x, x, x, \ldots}_{nx}$.

Stel nu $f(x) = \dfrac{a_m x^m + a_{m-1} x^{m-1} + \ldots + a_1 x + a_0}{b_m x^m + b_{n-1} x^{m-1} + \ldots + b_1 x + b_0}$ $(a_m, b_m \neq 0;\ x \neq 0)$

Dan is f blijkbaar ook continu op $\mathbb{R} \setminus \{0\}$. Definieert men $f(0) = \dfrac{a_m}{b_m}$ dan is f continu op $\mathbb{R}$. We hebben dus het definitiegebied van f uitgebreid met 0(verdichtingspunt van $X = \mathbb{R} \setminus \{0\}$). In feite hebben we dus eigenlijk een nieuwe functie gedefinieerd.

2, Stel $f(x) = c$ een vaste constante en wel voor $\forall x \in \mathbb{R}$. Dan is f continu.

Definitie. Zij $f: X \subset \mathbb{R} \to \mathbb{R}$ een reële functie, f heet monotoon stijgend (monotoon niet dalend) op X indien $x_1 > x_2$, $x_1, x_2 \in X$ impliceert $f(x_1) > f(x_2)$ $(f(x_1) \geq f(x_2))$ f heet monotoon dalend (monotoon niet stijgend) indien $x_1 > x_2$ $x_1, x_2 \in X$ impliceert $f(x_1) < f(x_2)$ $(f(x_1) \leq f(x_2))$.

Voorbeeld. De functie $f(x) = x^n$ is monotoon stijgend op $x = [0,\infty)$

Hoofdstuk 3. Topologie

§1 Metrische ruimten, het begrip topologie.

Zij X een verzameling, en zij $\mathbb{R}^+$ de verzameling der niet negatieve getallen.
Een functie

$$\rho : X \times X \to R^+$$

heet een metriek voor X, indien voor alle $x,y,z \in X$ geldt dat

(i) $\rho(x,y) = 0 \Leftrightarrow x = y$

(ii) $\rho(x,y) = \rho(y,x)$

(iii) $\rho(x,z) \leq \rho(x,y) + \rho(y,z)$ (driehoeksongelijkheid).

Het paar (X,ρ) heet dan een metrische ruimte. $\rho(x,y)$ heet de afstand tussen x en y.
De verzameling $B_r(p) = \{x \in X \mid \rho(x,p) < r\}$ $(r > 0)$ heet de (open) bol, met straal r en middelpunt p

Voorbeelden.

I. Voor een willekeurige verzameling X kunnen we definieren

$$\begin{cases} \rho(x,y) = 0 & \text{als} \quad x = y \\ \rho(x,y) = 1 & \text{als} \quad x \neq y \end{cases} \qquad (x,y \in X)$$

Dan is (X,ρ) een metrische ruimte.

II. Zij $\mathbb{R}$ de verzameling der reële getallen. Indien $\rho(x,y) = |x-y|$ $(x,y \in \mathbb{R})$ dan is (R,ρ) een metrische ruimte.

III. (Generalisatie van II), Zij

$$R^n = \{(x_1, \ldots, x_n) \mid x_i \in \mathbb{R} \text{ voor } i = 1,2,\ldots, n\}.$$

Indien $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$ en $y = (y_1, \ldots, y_n) \in \mathbb{R}^n$ dan definieren we

$$\rho(x,y) = \{\sum_{i=1}^{n} (x_i - y_i)^2\}^{\frac{1}{2}} \qquad (*)$$

We zullen bewijzen dat ρ een metriek is voor $\mathbb{R}^n$. Eerst tonen we echter aan de volgende

<u>Hulpstelling</u>. Indien p_i, $q_i \in \mathbb{R}$ voor $i = 1,2, \ldots, n$ dan geldt

$$\left(\sum_{i=1}^{n} p_i q_i\right)^2 \leq \left(\sum_{i=1}^{n} p_i^2\right) \left(\sum_{i=1}^{n} q_i^2\right).$$ (ongelijkheid van Cauchy Schwarz)

<u>Bewijs</u>.

Voor alle reële t geldt

$$0 \leq \sum_{i=1}^{n} (p_i t + q_i)^2 = t^2 \sum_{i=1}^{n} p_i^2 + 2t \sum_{i=1}^{n} p_i q_i + \sum_{i=1}^{n} q_i^2.$$

Maar dan is de discriminant van de kwadratische vorm in het rechterlid ≤ 0 d.w.z.

$$4\left(\sum_{i=1}^{n} p_i q_i\right)^2 - 4\left(\sum_{i=1}^{n} p_i^2\right)\left(\sum_{i=1}^{n} q_i^2\right) \leq 0$$

of $$\left(\sum_{i=1}^{n} p_i q_i\right)^2 \leq \left(\sum_{i=1}^{n} p_i^2\right)\left(\sum_{i=1}^{n} q_i^2\right).$$

Bewering (*) is een metriek voor $\mathbb{R}^n$

<u>Bewijs</u>.

Van de drie eisen, die aan een metriek zijn gesteld, zijn de eerste twee evident. We behoeven dus slechts de driehoeksongelijkheid te bewijzen. D.W.Z. we moeten aantonen dat

$$\left\{\sum_{i=1}^{n} (x_i-z_i)^2\right\}^{\frac{1}{2}} \leq \left\{\sum_{i=1}^{n} (x_i-y_i)^2\right\}^{\frac{1}{2}} + \left\{\sum_{i=1}^{n} (y_i-z_i)^2\right\}^{\frac{1}{2}}$$

Ofwel als men stelt $x_i-y_i = p_i$, $y_i-z_i = q_i$ (zodat $x_i-z_i = p_i+q_i$),

$$\sum_{i=1}^{n} (p_i+q_i)^2 \leq \sum_{i=1}^{n} p_i^2 + \sum_{i=1}^{n} q_i^2 + 2(\Sigma p_i^2)^{\frac{1}{2}} (\Sigma q_i^2)^{\frac{1}{2}}$$

i.e.

$$2 \sum_{i=1}^{n} p_i q_i \leq 2\left(\sum_{i=1}^{n} p_i^2\right)^{\frac{1}{2}} \left(\sum_{i=1}^{n} q_i^2\right)^{\frac{1}{2}}$$

maar dit volgt onmiddellijk uit de hulpstelling.

(R^n,ρ) heet de <u>n-dimensionale</u> <u>Enclidische</u> <u>ruimte</u> (voor $n = 1$ is dit de rechte lijn, voor $n = 2$ het platte vlak).

IV. Zij $H = \{(x_1, x_2, \ldots) \mid x_i \in R$ voor $i = 1,2, \ldots; \sum_{i=1}^{\infty} x_i^2 < \infty$.
Indien voor alle $x = (x_1,x_2, \ldots) \in H$ en $y = (y_1,y_2, \ldots) \in H$
$\rho(x,y)$ wordt gedefinieerd door

$$\rho(x,y) = \{\sum_{i=1}^{\infty} (x_i-y_i)^2\}^{\frac{1}{2}}$$

dan kunnen we weer aantonen dat ρ een metriek is voor $H.(H,\rho)$ heet de (separabele) Hilbertruimte. Notatie ℓ_2

Zij (M,ρ) een metrische ruimte. Zij $p \in M$. Een verzameling $G \subset M$ heet een omgeving <u>van</u> p, indien er een bol $B_r(p)$ bestaat zodanig dat $B_r(p) \subset G$.
In het bizonder is $B_r(p)$ een omgeving van p en ook van elk punt $q \in B_r(p)$. Immers, als $q \in B_r(p)$, laat $s = r-\rho(p,q)$, dan $x \in B_s(q) \Rightarrow$ $\Rightarrow \rho(x,q) < s \Rightarrow \rho(x,p) \leq \rho(x,q) + \rho(p,q) < s + \rho(p,q) = r \Rightarrow x \in B_r(p)$.
Dus $q \in B_s(q) \subset B_r(p)$.

Een niet lege verzameling $0 \subset M$ heet <u>open</u> indien 0 omgeving is van elke $p \in 0$ of equivalent wanneer bij elke $p \in 0$ een $r > 0$ bestaat zodat $p \in B_r(p) \subset 0$. I.h.b. is elke bol $B_r(p)$ voor $p \in M, r > 0$ een open verzameling van M; M zelf is een open verzameling van M. De lege verzameling is per definitie open.
Ga na dat een verzameling $G \subset M$ juist dan een omgeving is van $p \in M$ indien er een open verzameling 0 van M bestaat zodat $p \in 0 \subset G$.
We bewijzen nu twee fundamentele eigenschappen van open verzamelingen in een metrische ruimte. Deze eigenschappen zullen we gebruiken in de definitie van het begrip topologische ruimte - welk begrip dan als een generalisatie van het begrip metrische ruimte kan worden opgevat.

<u>Bewering</u>: Zij (X,ρ) een metrische ruimte.

(i) De vereniging van een willekeurige familie van open verzamelingen is een open verzameling.

(ii) De doorsnede van een eindige familie van open verzamelingen is een open verzameling.

(i) Zij $\{O_\alpha \mid \alpha \in A\}$ een collectie open verzamelingen van M en zij $O = \cup\{O_\alpha \mid \alpha \in A\}$. Laat p een willekeurig punt zijn van O. Blijkbaar bestaat er een index $\alpha \in A$ zodat $p \in O_\alpha$. Omdat O_α is open bestaat $r > 0$ zodat $p \in B_r(p) \subset O_\alpha$. Blijkbaar geldt nu ook $p \in B_r(p) \subset O$. Dus volgt dat ook O open is.
(ii) Het is voldoende aan te tonen dat de doorsnede van twee open verzamelingen weer open is (Ga dit na). Zij O_1 en O_2 open verzamelingen en $p \in O_1 \cap O_2$ willekeurig. Kies r,s zodat $p \in B_r(p) \subset O_1$ en $p \in B_s(p) \subset O_2$. Zij $t = \min(r,s)$. Dan $p \in B_t(p) \subset O_1 \cap O_2$ waaruit volgt dat inderdaad $O_1 \cap O_2$ open is.

Opmerking. De doorsnede van oneindig veel open verzamelingen is i.h.a. niet open!

Definitie. Zij X een verzameling. Een familie $\mathcal{T}$ van deelverzamelingen van X heet een topologie voor X indien
(i) $\phi \in \mathcal{T}$ $\quad X \in \mathcal{T}$
(ii) de vereniging van elke deelfamilie van $\mathcal{T}$ tot $\mathcal{T}$ behoort
(iii) de doorsnede van elke eindige deelfamilie van $\mathcal{T}$ tot $\mathcal{T}$ behoort.

Het paar $(X,\mathcal{T})$ heet dan een topologische ruimte.
De verzamelingen $O \in \mathcal{T}$ heten $(\mathcal{T})$-open.
In plaats van de elementen van X spreken we van de punten van de topologische ruimte $(X,\mathcal{T})$.

Voorbeelden.
1) Zij X een verzameling. Indien $\mathcal{T}$ de familie van alle deelverzamelingen is, dan is $(X,\mathcal{T})$ een topologische ruimte.
$\mathcal{T}$ heet dan de discrete topologie op X.
2) Zij X een verzameling. Indien $\mathcal{T} = \{\phi, X\}$ dan is $(X,\mathcal{T})$ een topologische ruimte. $\mathcal{T}$ heet de indiscrete topologie voor X.

<u>Opmerking</u>. Indien $\mathcal{T}_1$ en $\mathcal{T}_2$ twee topologieën zijn voor X zodanig dat $\mathcal{T}_1 \subset \mathcal{T}_2$, dan heet $\mathcal{T}_1$ <u>kleiner</u> (<u>grover</u>) dan $\mathcal{T}_2$; $\mathcal{T}_2$ heet <u>groter</u> (<u>fijner</u>) dan $\mathcal{T}_1$. De discrete topologie is de fijnste topologie en de indiscrete topologie is de grofste topologie voor X.

3) Zij (X,ρ) een metrische ruimte. Indien $\mathcal{T}$ de familie van alle open verzamelingen is (zoals hierboven gedefinieerd, dan is $(X,\mathcal{T})$ een topologische ruimte. $\mathcal{T}$ heet de <u>metrische topologie</u> (geïnduceerd door de metriek ρ).

Indien $(X,\mathcal{T})$ een topologische ruimte is en indien er een metriek ρ voor X bestaat, zodanig dat de door ρ geïnduceerde topologie juist $\mathcal{T}$ is, dan heet $(X,\mathcal{T})$ <u>metrizeerbaar</u>. Niet elke topologische ruimte is metrizeerbaar.

4) De reële rechte is een metrische ruimte. De metrische topologie heet wel de <u>gewone</u> of <u>euclidische</u> topologie op $\mathbb{R}$. Analoog spreken we van de gewone of euclidische topologie op $\mathbb{R}^n$.

2. Enkele fundamentele begrippen van topologische ruimten.

Zij X een topologische ruimte (wanneer geen verwarring mogelijk is zullen we i.p.v. de topologische ruimte $(X,\mathcal{T})$ steeds praten over de topologische ruimte X).

Als $U \subset X$ een punt p bevat dan heet U een omgeving van p indien er een open verzameling 0 bestaat zodat $p \in 0 \subset U$.
Ga na dat een verzameling G open is in X d.e.s.d. wanneer G een omgeving is van elk van zijn punten.
Dit begrip omgeving komt overeen met het begrip omgeving zoals gedefinieerd voor metrische ruimten op blz. 33, wanneer we (zoals gebruikelijk) de metrische toplogie beschouwen.

Een verzameling $F \subset X$ heet gesloten wanneer $X \setminus F$ open is (in de topologische ruimte X).

Voorbeelden.

I X en ϕ zijn gesloten

II In elke metrische ruimte (X,ρ) is de verzameling

$$S_r(p) = \{x \mid \rho(x,p) \leq r\}$$

gesloten ("gesloten bol") $(r > 0)$.
Bovendien is elke verzameling die uit één punt bestaat gesloten.

III In de reële rechte is de verzameling $\{0,1,\frac{1}{2},\frac{1}{4}, \ldots\}$ gesloten, de verzameling $\{1,\frac{1}{2},\frac{1}{4}, \ldots\}$ echter niet.

Stelling. Zij $(X,\mathcal{T})$ een topologische ruimte.

(i) De doorsnede van een willekeurige familie van gesloten verzamelingen is weer een gesloten verzameling.
(ii) De vereniging van een eindige familie van gesloten verzamelingen is weer een gesloten verzameling.

Bewijs. Ga dit zelf na.

Als nu A een verzameling is in een topologische ruimte X, dan is de doorsnede van alle gesloten verzamelingen die A omvatten kennelijk de kleinste gesloten verzameling die A omvat. Men noemt deze verzameling het gesloten omhulsel of de afsluiting van A (in X); notatie $\overline{A}$.

Ten aanzien van de afsluiting gelden de volgende eigenschappen.

Stelling 1. $\overline{A \cup B} = \overline{A} \cup \overline{B}$

$A \subset B \Rightarrow \overline{A} \subset \overline{B}$

A is gesloten $\Leftrightarrow \overline{A} = A$

$\overline{\overline{A}} = \overline{A}$.

Opmerking. Voor open verzamelingen heeft men iets analoogs
(i) de vereniging van alle open verzamelingen $\subset A$ is de grootste open verzameling $\subset A$. Men noemt deze verzameling het inwendige van A; notatie A^0.
(ii) A is open $\Leftrightarrow A^0 = A$.

Een verzameling A in een topologische ruimte X heet (overal)dicht als $\overline{A} = X$.

We zullen hieronder enkele andere karakteriseringen van de begrippen "gesloten verzameling" en afsluiting geven.

Zij A een verzameling in de topologische ruimte X. Een punt $p \in X$ heet verdichtingspunt van A, indien voor iedere omgeving U_p van p geldt dat $U_p \cap (A \setminus \{p\}) \neq \phi$.

Merk op dat deze definitie van verdichtingspunt overeenkomt met de vroegere definitie voor de reële getallen.

Voorbeeld. In de reële rechte is 0 verdichtingspunt zowel van de verzameling $\{0,\frac{1}{2},\frac{1}{4},\frac{1}{8}, \ldots\}$ als van de verzameling $\{\frac{1}{2},\frac{1}{4}, \ldots \}$.

Stelling 2. Een verzameling is dan en slechts dan gesloten indien A de verzameling van zijn verdichtingspunten bevat.

Bewijs. Zij eerst A gesloten en p een verdichtingspunt van A.
We moeten aantonen $p \in A$. Onderstel $p \notin A$.
Blijkbaar is dan $U = X \setminus A$ een omgeving van p. Omdat $U \cap (A \setminus \{p\}) = \phi$ kan p nooit een v.d.p. van A zijn. Tegenspraak, dus $p \in A$. Omgekeerd, als elk v.d.p. van A element is van A, moeten we bewijzen dat A gesloten is. Laat $q \in X \setminus A$ willekeurig
Blijkbaar is q geen v.d.p. van A, dus er bestaat een omgeving U_q van q zodat $U_q \cap A \setminus \{q\} = 0$ i.e. $U_q \cap A = \phi$ (want $q \notin A$) en ook $U_q' \subset X \setminus A$.
$X \setminus A$ is dus open omdat deze verzameling een omgeving is van elk van zijn punten. Inderdaad volgt nu dat A gesloten is.

Inwendige, uitwendige en rand van een verzameling.

Stel X een topologische ruimte. Met A^0 noteren we het inwendige van een deelverzameling A. Het complement van een deelverzameling A noteren we als A^c.

Bewering $A^0 = A^{c-c}$.
Bewijs. Omdat $E \subset A \Leftrightarrow A^c \subset E^c$ merken we op dat de open verzamelingen $E \subset A$ precies de complementen zijn van gesloten verzamelingen $\supset A^c$. Dus $A^0 = \cup \{F^c \mid F$ is gesloten en $F \supset A^c\} = \cap \{$ F gesloten en $F \supset A^c \}^c = A^{c-c}$.

Definitie. Is A een willekeurige deelverzameling van een topologische ruimte X dan valt X uiteen in drie disjuncte deelverzamelingen.

(a) $A^{c-c} = A^0$ het inwendige van A
(b) $A^{-c} = A^{c0}$ het uitwendige van A
(c) $(A^{-c} \cup A^{c-c})^c = \overline{A} \cap A^{c-}$ de rand van A, notatie b(A).

Punten van het inwendige, uitwendige of de rand van A heten respectievelijk <u>inwendige</u> punten, <u>uitwendige</u> punten en <u>randpunten</u> van A.

<u>Voorbeelden</u>.

I. Stel $A \subset \mathbb{R}^2$, $A = \{(x,y) \mid \frac{x^2}{a^2} + \frac{y^2}{b^2} \leq 1\}$. $a,b \neq 0$.

Het inwendige van A is $\{(x,y) \mid \frac{x^2}{a^2} + \frac{y^2}{b^2} < 1\}$; het uitwendige van A is $\{(x,y) \mid \frac{x^2}{a^2} + \frac{y^2}{b^2} > 1\}$; de rand van A is $\{(x,y) \mid \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1\}$.

Hoe vreemd de rand van een verzameling er i.h.a. uit kan zien getuige het volgende voorbeeld.

II. Stel Q de deelverzameling van R bestaande uit de rationale getallen. Het uitwendige van Q is ϕ, het uitwendige van Q is ϕ en de rand van Q is $\mathbb{R}$.

III. Beschouw $\mathbb{R}^n$ met de gewone topologie (n vast).

Zij $S_r(p) = \{x \in \mathbb{R}^n \mid \rho(x,p) \leq r\}$.

Bewijs dat het inwendige van $S_r(p)$ gelijk is aan $B_r(p) =$
$= \{x \in \mathbb{R}^n \mid \rho(x,p) < r\}$; het uitwendige van $S_r(p) = \{x \in \mathbb{R}^n \mid \rho(x,p) > r\}$ en de rand van $S_r(p) = \{x \in \mathbb{R}^n \mid \rho(x,p) = r\}$.

Geldt deze eigenschap in elke metrische ruimte ??

III. Stel A een deelverzameling van een topologische ruimte X

p is een randpunt van A $\Leftrightarrow$. Iedere omgeving van p bevat zowel van A als van A^c punten.

Bewijs ook dat $\overline{A} = A \cup b(A)$

Als A open is in X toon dan aan dat $b(A) \cap A = \phi$.

<u>Deelruimten van een topologische ruimte</u>.

Stel $(X,\mathcal{T})$ een topologische ruimte. Een deelverzameling Y van X maken we tot een topologische ruimte door als stelsel open verzamelingen te nemen $\mathcal{T}^1 = \mathcal{T} \cap Y = \{U \cap Y \mid U \in \mathcal{T}\}$. $(Y,\mathcal{T}^1)$ heet een <u>deelruimte</u> van $(X,\mathcal{T})$. De topologie op Y heet de <u>relatieve topologie</u>.

Wanneer we dus nu een willekeurige deelverzameling Y van een topologische ruimte bekijken, dan is Y blijkbaar open en gesloten in de deelruimte Y doch niet noodzakelijk open in X. Hieruit blijkt weer duidelijk dat het begrip open geen absolute betekenis heeft. We moeten expliciet de ruimte vermelden waar we de topologie van bedoelen.

Voorbeelden

I. Een segment $[a,b]$ is een deelruimte van de reële rechte R.

II. De verzameling Q der rationale getallen is een deelruimte van de reële rechte R. De deelverzamelingen $\{x \in Q \mid a < x < b$, a,b irrationale getallen$\}$ zijn zowel open als gesloten deelverzamelingen van de deelruimte Q (maar niet van R).

III. De verzameling der natuurlijke getallen is met de relatieve topologie (geïnduceerd door R) een discrete topologische ruimte.

Stel nu (X,ρ) een metrische ruimte. X is dus een topologische ruimte d.m.v. de metrische topologie

Een deelverzameling Y van X is op natuurlijke wijze een metrische ruimte (Y,ρ') als we afspreken dat de afstand $\rho'(y,z)$ tussen twee punten y,z van Y gelijk is aan $\rho(y,z)$. Op natuurlijke wijze is dus op Y de metrische topologie geïnduceerd. Nu is het gemakkelijk in te zien dat deze topologie op Y samen valt met de relatieve topologie op Y (geïnduceerd door X).

Continue afbeeldingen.

Laten X en Y topologische ruimten zijn.

Een functie $f: X \to Y$ heet continu in $x_0 \in X$ indien bij iedere omgeving V van $f(x_0)$ in Y een omgeving U van x_0 in x bestaat zodat $f(U) \subset V$. Een functie $f: X \to Y$ heet continu indien f continu is in elk punt van X.

Stelling 3. Een functie $f: X \to Y$ is dan en slechts dan continu indien voor iedere open verzameling 0 van Y de verzameling $f^{-1}(0) =$
$= \{x \in X \mid f(x) \in 0\}$ open is in X.

Deze definitie van continuiteit is zeer algemeen en deze kunnen we dus ook toepassen om het begrip continuïteit te onderzoeken bij afbeeldingen tussen metrische ruimten en i.h.b. bij afbeeldingen tussen reële getallen.

<u>Stelling 4</u>. De samenstelling van twee continue functies is weer continu. In formule: Als X,Y,Z topologische ruimten zijn en $f:X \to Y$ $g:Y \to Z$ zijn continue afbeeldingen dan is ook $g \circ f: X \to Z$ continu.

<u>Bewijs</u>. Dit volgt gemakkelijk uit de vorige stelling.

<u>Stelling 5</u>. Een afbeelding f van een metrische ruimte (X,ρ) naar een metrische ruimte (Y,σ) is dan en slechts dan continu in $x_0 \in X$ indien bij iedere $\varepsilon > 0$ een $\delta > 0$ is te vinden zodanig dat $\sigma(f(x), f(x_0)) < \varepsilon$ voor alle $x \in X$ die voldoen aan $\rho(x,x_0) < \delta$.

<u>Bewijs</u>. Veronderstel eerst dat f continu is. Stel $x_0 \in X$ en kies $\varepsilon > 0$. Blijkbaar is $B_\varepsilon(f(x_0)) = \{y \in Y \mid \sigma(f(x_0),y) < \varepsilon\}$ een omgeving van $f(x_0)$ in Y. Omdat f continu is bestaat er een omgeving U van x_0 in X zodat $f(U) \subset B_\varepsilon(f(x_0))$. Uit de definitie van omgeving en open verz. in X volgt dat er een $\delta > 0$ bestaat zodat $B_\delta(x_0) = \{x \in X \mid \rho(x,x_0)<\delta\}\subset U$.

Nu geldt als $\rho(x,x_0) < \delta$ dan $x \in U$ dus $f(x) \in B_\varepsilon(f(x_0))$ d.w.z. $\sigma(f(x), f(x_0) < \varepsilon$. En hiermede is het eerste deel van de stelling aantoond.

Zij omgekeerd f een afbeelding die aan de eis in de stelling voldoet. Zij V een omgeving van $f(x_0)$ in Y. Uit de definitie van omgeving en open verz. in de metrische ruimte Y volgt dat er een $\varepsilon > 0$ bestaat zodat $B_\varepsilon(f(x_0)) \subset V$. Kies bij deze ε een positief getal δ zodat $\rho(x,x_0) < \delta$ impliceert $\sigma(f(x), f(x_0) < \varepsilon$. Dan is $U = B_\delta(x_0)$ een omgeving van x_0 die door f binnen V wordt afgebeeld. Dus f is continu in x_0.

<u>Definitie</u>. Zij f een afbeelding van een metrische ruimte (X,ρ) naar een metrische ruimte (Y,σ). Dan heet f <u>uniform continu</u> indien

$$\forall \varepsilon > 0 \quad \exists \delta > 0 \quad \text{zodat} \quad \rho(x,a) < \delta \Rightarrow \sigma(f(x), f(a)) < \varepsilon$$

voor $\forall a \in X$.

Het begrip uniforme continuiteit is niet een topologisch begrip doch een metrisch begrip. Immers door verandering van metriek kan een uniform continue functie overgaan in een niet uniform continue functie. Maar zoals we later zien is onafhankelijk van ρ en σ een functie gedefinieërd op een compacte metrische ruimte steeds uniform continue. I.h.b. is een continue functie gedefinieërd op een kubus $\{x \in \mathbb{R}^n | a_i \leq x_i \leq b_i\}$ uniform continu.

Passen we de definitie van continuiteit en bovenstaande stelling toe op de $\mathbb{R}^n$ dan krijgen we

Stelling 6. Een afbeelding f van een deelverzameling $K \subset \mathbb{R}^n$ in $\mathbb{R}^n$ is dan en slechts dan continu in $a = (a_1,\ldots,a_n)$ indien $\forall \epsilon > 0 \; \exists \delta > 0$ zodat voor alle $x = (x_1,\ldots,x_n)$ met

$$\underbrace{\sqrt{\sum_{i=1}^{m} (x_i - a_i)^2} < \delta}_{(1)} \qquad \text{geldt} \qquad \underbrace{\sqrt{\sum_{i=1}^{m} ((f(x))_i - (f(a))_i)^2} < \varepsilon}_{(2)}$$

Hierbij stelt $(f(x))_i$ en $(f(a))_i$ de i'de coordinaat van f(x) resp f(a) voor.

Voor het geval m = 1. (reële functies) kan men i.p.v. (2) schrijven $|f(x) - f(a)| < \varepsilon$
en als ook n = 1 kan men i.p.v. (1) schrijven $|x-a| < \delta$.
Zodoende krijgt men in dit geval overeenstemming met het begrip continuiteit zoals gedefinieerd voor reële functies op blz. 29.

Ga zelf na dat ook het volgende geldt

Stelling. Een afbeelding f van een deelverzameling $K \subset \mathbb{R}^n$ in $\mathbb{R}^m$ is dan en slechts dan continu in $a = (a_1,\ldots,a_n)$ van $\mathbb{R}^n$ indien $\forall \varepsilon > 0 \; \exists \delta > 0$ zodat voor alle $x = (x_1,\ldots,x_n)$ met

$$\max_{i=1,..,n} |x_i - a_i| < \delta \qquad \text{geldt} \qquad \max_{i=1,..,m} |((f(x))_i - (f(a))_i| < \varepsilon \,.$$

Afbeeldingen van deelverzamelingen K van R^n naar $\mathbb{R}$ komen op natuurlijke wijze te voorschijn bij het bestuderen van functies van meerdere variabelen.
Immers zij $f(x_1,x_2,\ldots,x_n)$ een grootheid die afhankelijk is van de variabelen $x_1,\ldots,x_n$ waarbij $a_i \le x_i \le b_i$. Dan is deze grootheid niets anders dan een functie f van $K \subset \mathbb{R}^n$ naar $\mathbb{R}$ met $K = \{x \in \mathbb{R}^n | a_i \le x_i \le b_i\}$.
Als nu bv. gegeven is dat f continu is in alle variabelen $x_1,\ldots,x_n$ tezamen dan wil dit niets anders zeggen dat $f\colon K \to \mathbb{R}$ een continue afbeelding is.

Vervolg van de algemene topologie.

Definitie. Zij f een eenduidige continue afbeelding van een ruimte X op een ruimte Y. Als ook de inverse afbeelding $g = f^{-1}$ van Y op X continu is dan heet f een homeomorfisme. Dit is d.e.s.d. het geval wanneer f de open verzamelingen van X afbeeldt op de open verzamelingen van Y. Twee ruimten X en Y waartussen een homeomorfisme bestaat heten homeomorf (of topologisch equivalent). Topologisch kan men bij homeomorfe ruimten geen onderscheid zien!

§3. Samenhang

Definitie. Een topologische ruimte X heet samenhangend wanneer $\emptyset$ en X de enige zowel open als gesloten deelverzamelingen van X zijn.
Een deelverzameling A van een topologische ruimte X heet samenhangend wanneer A als deelruimte samenhangend is.
Voor het gemak zullen we een zowel open als gesloten deelverzameling van een topologische ruimte opgesloten noemen.

Voorbeelden.

I. Een ruimte bestaande uit één punt is samenhangend.

II. De reële rechte $\mathbb{R}$ met de gewone topologie is samenhangend.

Bewijs. Veronderstel dat er toch een niet lege opgesloten deelverzameling A van $\mathbb{R}$ bestond.
Stel $B = \mathbb{R}\backslash A$. Kies $a \in A$ en $b \in B$, we mogen onderstellen dat $a < b$.
Zij $p = \sup (A \cap [a,b])$. p behoort tot A want p is een punt van A of een verdichtingspunt van A en A is gesloten. Iedere (rechter) omgeving van p bevat punten van $\mathbb{R}\backslash A = B$, dus A is niet open. Tegenspraak. Op dezelfde wijze bewijst men dat ieder segment samenhangend is.

III. De enige samenhangende deelverzamelingen van $\mathbb{R}$ zijn intervallen $[a,b]$, $(a,b]$, $[a,b)$, (a,b) met a,b al of niet ∞ ($(\infty,b]$ betekent $\{x|x \leq b\}$ enz.).

IV. De deelverzameling Q van $\mathbb{R}$ bestaande uit alle rationale getallen is niet samenhangend. (als a,b irrationale getallen zijn, $a < b$, dan is $\{x \in Q|a < x < b\}$ een niet lege opgesloten verzameling van Q die $\neq Q$ is.

Stelling.1 Een topologische ruimte X is dan en slechts dan niet samenhangend, wanneer er twee deelverzamelingen A en B bestaan met
1) $A \neq \emptyset$, $B \neq \emptyset$
2) $A \cup B = X$; $A \cap B = \emptyset$
3) A en B gesloten in X.

Het is gebruikelijk om een niet samenhangende ruimte X splitsbaar te noemen. Als $X = A \cup B$ terwijl A en B aan 1) t/m 3) van de vorige stelling voldoen dan heet (A,B) een splitsing van X.

Stelling 2. Stel X een topologische ruimte en A een samenhangende deelverzameling van X. Dan is $\bar{A}$ samenhangend.

Bewijs. Stel C een opgesloten deelverzameling van $\bar{A}$, we moeten aantonen dat $C = \emptyset$ of $C = \bar{A}$.
$C \cap A$ is een opgesloten deelverzameling van A en omdat A samenhangend is, geldt blijkbaar $C \cap A = \emptyset$ of $C \cap A = A$. $C \cap A = \emptyset$ implieert $A \subset \bar{A}\backslash C$. Omdat $\bar{A}\backslash C$ gesloten is in $\bar{A}$ (want C is open in $\bar{A}$) en dus in X volgt $\bar{A} \subset \bar{A}\backslash C$ d.w.z. $C = \emptyset$. De tweede onderstelling $C \cap A = A$ impliceert dat $A \subset C$ en omdat C gesloten is in $\bar{A}$ (en dus in X) ook $\bar{A} \subset C$ d.w.z. $C = \bar{A}$.

Stelling 3. (generalisatie van de vorige stelling).
Zij A een samenhangende deelverzameling van een topologische ruimte X. Als $A \subset B \subset \bar{A}$, dan is B samenhangend.

Bewijs: Opgave.

Stelling 4. Zij $\{A_\alpha | \alpha \in I\}$ een familie samenhangende verzamelingen van een topologische ruimte en veronderstel dat voor elk paar (α,β) geldt $A_\alpha \cap A_\beta \neq \emptyset$. Dan is $\bigcup\{A_\alpha | \alpha \in I\}$ samenhangend.

Toepassing.

1) De deelverzameling A van $\mathbb{R}^2$ gedefinieerd door $A = \{(x,y) \in \mathbb{R}^2 | y/x$ is rationaal$\}$ is samenhangend (merk op dat $A = \bigcup \{A_\alpha | \alpha$ is rationaal$\}$ met $A_\alpha = \{(x,y) | y = \alpha x\}$ en pas de vorige stelling toe).

2) De ruimte $\mathbb{R}^m$ $m \geq 1$ is samenhangend.
Bewijs. Schrijf $\mathbb{R}^m$ als $\mathbb{R}^m = \bigcup\{S_a | a \in \mathbb{R}^m\}$ met $S_a = \{x \in \mathbb{R}^m | x_i = \lambda a_i\}$.
Elke verzameling S_a is samenhangend want $\lambda \longmapsto \lambda a_1,\ldots,\lambda a_n)$ bewerkstelligt een homeomorfisme tussen $\mathbb{R}$ (die samenhangend is) en S_a.
Omdat $S_a \cap S_b \supset \{0\}$ voor alle $a, b \in \mathbb{R}^m$ volgt uit de vorige stelling dat $\mathbb{R}^m$ is samenhangend.

Een zeer belangrijke stelling is de volgende:

Stelling 5. Zij f een continue afbeelding van de ruimte X op de ruimte Y. Als X samenhangend is, dan is ook Y samenhangend. Kortweg het continue beeld van een samenhangende ruimte is samenhangend.

Bewijs. Zij Z een zowel open als gesloten deelverzameling van Y. Dan is ook $f^{-1}(Z)$ zowel open als gesloten in X. Dus, omdat X is samenhangend, $f^{-1}(Z) = \emptyset$ of $f^{-1}(Z) = X$. Dus $Z = Y$ of $Z = \emptyset$. Hieruit volgt dat Y samenhangend is.

Gevolg. Zij f een continue afbeelding van een samenhangende ruimte X naar $\mathbb{R}$. (b.v. een continue afbeelding van $\mathbb{R}^n$ naar $\mathbb{R}$ of een gewone reële functie $f: \mathbb{R} \to \mathbb{R}$ of $f: X \to \mathbb{R}$ met X samenhangend in $\mathbb{R}$).

Dan kan $f(X)$ alleen de volgende vormen aannemen $(-\infty,a)$, $(-\infty,a]$, $(a,b]$, $[a,b]$, $[a,b)$, $[a,\infty)$, (a,∞) of $f(X)$ is een punt.
I.h.b. volgt dat de functie met elke twee waarden $f(x)$ en $f(y)$ die ze aanneemt ook elke waarde u met $f(x) < u < f(y)$ aanneemt. Dit laatste resultaat staat bekend als de <u>tussenwaardestelling voor continue functies</u>).

<u>Toepassing</u>. Zij f een monotoon dalende of stijgende functie van een samenhangende deelverzameling X van $\mathbb{R}$ naar $\mathbb{R}$.
Dan is $f^{-1}: fX \to X$ weer continu.

<u>Bewijs</u>. We merken op dat fX blijkbaar een samenhangende deelverzameling is van X. Ook merken we op dat f 1-1-duidig is immers $x_1 \neq x_2$ dan $x_1 > x_2$ of $x_1 < x_2$ dus $f(x_1) > f(x_2)$ of $f(x_1) < f(x_2)$. We zullen aantonen dat als $(a,b) \subset X$ dan $f(a,b) = (f(a), f(b))$. Hieruit volgt de continuiteit van f. Inderdaad, het is duidelijk dat als $x \in (a,b)$ dan $f(x) \in (f(a), f(b))$ omdat f monotoon is. Omgekeerd als $y \in (f(a), f(b))$ dan volgens de tussenwaarde stelling $\exists x \in X$ zodat $f(x) = y$. Uit de monotonie van f volgt weer $x \in (a,b)$.

<u>Opgave</u>. Zij f een continue functie van een samenhangende ruimte X naar $\mathbb{R}$. Als gegeven is dat f in elk punt van X een lokaal maximum aanneemt ($\exists$ omgeving U van het punt x zodat $f(y) \leq f(x)$ voor $y \in U$) bewijs dan dat f een constante is.

<u>Definitie</u>. Een open samenhangende verzameling van $\mathbb{R}^n$ heet een <u>gebied</u>

<u>Definitie</u>. Een deelverzameling A van $\mathbb{R}^n$ heet een <u>gesloten jordankromme</u> indien er een homeomorfisme bestaat van de deelruimte $\{(x,y) \in R^2 | x^2 + y^2 = 1\}$ (= cirkel) op A.

Zonder bewijs vermelden we de volgende klassieke stelling.

Stelling 6. (Jordan) Zij C een gesloten jordankromme in $\mathbb{R}^2$. Dan bestaan er precies twee gebieden G_1 en G_2 van $\mathbb{R}^2$ zodat

1) $b(G_1) = b(G_2)$ (rand van G_1 resp. G_2) $= C$
2) $G_1 \cup G_2 \cup C = \mathbb{R}^2$, $G_1 \cap G_2 = \emptyset$
3) G_1 is begrensd en G_2 is onbegrensd (in dat gevel heet G_1 het binnengebied; G_2 het buitengebied van C) óf G_2 is begrensd en G_2 is onbegrensd (in dat geval heet G_2 het binnengebied, G_1 het buitengebied van C.

De stelling van Jordan is zeer algemeen en zegt zoveel als: een gesloten kromme verdeeld het vlak in twee delen met gemeenschappelijke rand, een binnengebied en een buitengebied.

Opgave 1. Onderzoek de geldigheid van de stelling van Jordan voor het geval $C \subset \mathbb{R}^2$ is gegeven door $C = \{(x,y) \mid \frac{x^2}{a^2} + \frac{y^2}{b^2} = 1\}$.

Geldt de stelling van Jordan ook voor $\mathbb{R}^n$ $(n > 2)$??

Opgave 2. Zij $f(x_1,\dots,x_n)$ een grootheid die continu afhangt van reële variabelen $x_1,\dots,x_n$. $0 \leq x_i \leq 1$.

a Zij $f(0,0,0,\dots,0) = 0$ $f(1,1,1,\dots,1) = 1$. Bewijs dat er $0 \leq y_1,\dots,y_n \leq 1$ bestaan met $f(y_1,\dots,y_n) = \frac{1}{2}$.

b Zij $f(x_1,x_2,\dots,x_n) \neq 0$ voor zekere $x_1,\dots,x_n$ met $0 \leq x_i \leq 1$. Bewijs dat er $a_i \in R$ bestaan zodat $f(x_1',x_2',\dots,x_n') \neq 0$ voor alle x_i' met $|x_i' - x_i| < a_i$ $i = 1,\dots,n$.

§4. Compactheid.

Zij X een verzameling en zij $A \subset X$. Een familie K van deelverzamelingen van X heet een overdekking van A, indien $A \subset \bigcup K$; men zegt dan ook dat K A overdekt. Indien $H \subset K$, en indien ook H een overdekking is van A, dan heet H een deeloverdekking van K voor A.

Als $(X,\mathcal{T})$ een topologische ruimte is en $A \subset X$ dan heet een overdekking $\mathcal{K}$ van A een open overdekking indien ieder element van $\mathcal{K}$ $\mathcal{T}$-open is. Een topologische ruimte X heet compact (of bicompact) wanneer iedere open overdekking van X een eindige deeloverdekking bezit. Een deelverzameling Y van een topologische ruimte X heet compact, wanneer Y als deelruimte compact is.

Het is duidelijk dat : $Y \subset X$ is compact $\Longleftrightarrow$ Iedere overdekking van Y met open verzamelingen van X heeft een eindige deeloverdekking.

Voorbeelden.

1) De reële rechte is niet compact, want de overdekking van $\mathbb{R}$ bestaande uit alle open intervallen $(-n,n)$ $(n=1,2,\ldots)$ bezit geen eindige deeloverdekking
2) Als A een onbegrensde verzameling is van $\mathbb{R}$ dan is A niet compact.
3) Een segment $[a,b]$ is een compacte verzameling van $\mathbb{R}$.

Bewijs. Stel $\mathcal{U}$ een open overdekking van $[a,b]$ (met open verzamelingen van $\mathbb{R}$). Stel c het supremum van alle getallen x van $[a,b]$ met de eigenschap dat $[a,x]$ door een eindige deelcollectie van $\mathcal{U}$ wordt overdekt. Kies $U \in \mathcal{U}$ zodat $c \in U$ (merk op dat $c \in [a,b]$) en kies een open interval (d,c) zodat $[d,c] \subset U$. Er bestaat nu een eindige deelcollectie van $\mathcal{U}$ die $[a,d]$ overdekt en deze familie tezamen met $\{U\}$ overdekt $[a,c]$. Nu is $c = b$ want anders verkrijgen we een eindige deelcollectie van $\mathcal{U}$ die een interval $[a,c']$ overdekt met $b \geq c' > c$, wat in tegenspraak is met de definitie van c.

Stelling 1. Stel Y een gesloten deelverzameling van de compacte ruimte X. Dan is Y compact.

Bewijs. Stel $\mathcal{U}$ een willekeurige open overdekking van Y (met open verzamelingen van X). Omdat X compact is bezit de open overdekking $\mathcal{U} \cup \{X\backslash Y\}$ van X een eindige deeloverdekking $\mathcal{H}$. $\mathcal{H}\backslash\{X\backslash Y\}$ is het gevraagde eindige deelstelsel van $\mathcal{U}$ dat Y overdekt.

Definitie. Men noemt een topologische ruimte een Hausdorff-ruimte wanneer aan het volgende axioma voldaan is.

(*) Iedere twee verschillende punten bezitten disjuncte omgevingen.

De meeste van de ruimten die in deze syllabus als voorbeeld gegeven zijn, zijn Hausdorff-ruimten. Een discrete ruimte is een Hausdorff-ruimte, iedere metrische ruimte (X,ρ) met de metrische topologie is een Hausdorff-ruimte (als $p, q \in X$, $p \neq q$ en $r = \frac{1}{2}\rho(p,q)$ dan zijn $B_r(p)$ en $B_r(q)$ disjuncte omgevingen van p en q).
Verder is iedere deelruimte van een Hausdorff-ruimte weer een Hausdorff-ruimte.

Stelling 2. Stel C een compacte deelverzameling van een Hausdorff-ruimte X. Dan is C gesloten in X.

Bewijs. Stel p een vast punt $\notin C$. Kies voor ieder punt $c \in C$ een open omgeving U_c en een omgeving U_p van p met $U_c \cap U_p = \emptyset$. Omdat C compact is, bezit de overdekking $\{U_c | c \in C\}$ van C een eindige deel-overdekking $\{U_{c_i} | i=1,\ldots,n\}$. $\bigcap \{U_{p_i} | i=1,\ldots,n\}$ is nu een (open) omgeving van p die C vermijdt. Uit de willekeurigheid van p volgt dat $X \setminus C$ open is in X, dus C is gesloten in X.

Opgave. (generalisatie van de vorige stelling). Bewijs dat de doorsnede van een compacte deelverzameling en een gesloten deelverzameling van een topologische ruimte compact is.
2. Bewijs dat de doorsnede van willekeurig veel compacte verzamelingen van een Hausdorff-ruimte weer compact is. Toon aan dat deze uitspraak niet juist is in een willekeurige topologische ruimte X. (Stel $A = \{a_n | n=1,2,\ldots\}$ een aftelbare verzameling en p en q twee vaste punten $\notin A$. Kies de volgende topologie $\mathcal{T}$ op $X = A \cup \{p,q\}$:
$\mathcal{T} = \{a_1\}, \{a_2\}, \ldots, \{p\} \cup \{$bijna alle $a_n\}, \{q\} \cup \{$bijna alle $a_n\}$.
$A \cup \{p\}$ en $A \cup \{q\}$ zijn twee compacte verzamelingen van $(X,\mathcal{T})$ en hun doorsnede is niet compact.

Stelling 3. Zij f een continue afbeelding van de compacte ruimte X op de ruimte Y. Dan is Y compact.
Anders gezegd : compactheid is een continue invariant.

Bewijs. Kies een willekeurige open overdekking $\mathcal{U}$ van Y. Het stelsel $\{f^{-1}(U) | U \in \mathcal{U}\}$ is een open overdekking van X en uit de compactheid van X volgt dat voor eindig veel $U_i \in \mathcal{U}$ $(i=1,2,\ldots,n)$ reeds geldt $X = \bigcup \{f^{-1}(U_i) | i=1,\ldots,n\}$. $\{U_i | i=1,\ldots,n\}$ is nu de gevraagde deel-overdekking van $\mathcal{U}$.

Toepassing.

1. Er bestaat geen continue afbeelding van een segment [a,b] op de reële rechte R (anders zou R als continu beeld van de compacte ruimte [a,b] compact zijn en dat is niet waar. Dus volgt dat het beeld van [a,b] ook begrensd is en het is gesloten in $\mathbb{R}$ (compact in Hausdorffruimte $\Rightarrow$ gesloten); ook is het samenhangend (zie vroeger) dus volgt dat het beeld noodzakelijkerwijs weer een interval [c,d] is.
2. Als f een continue afbeelding is van een compacte ruimte X naar $\mathbb{R}$ dan bereikt f een maximum en een minimum.

   Bewijs. f(X) is een compacte verzameling van $\mathbb{R}$. Dus f(X) is gesloten in $\mathbb{R}$ (stelling). f(X) is ook begrensd want onbegrensde verzamelingen van $\mathbb{R}$ zijn nooit compact (zie blz. 48). Dus inf f(X) en sup f(X) bestaan, deze zijn resp. het maximum en het minimum van f op X.

Stelling 4. Een 1-1 continue afbeelding f van een compacte ruimte X op een Hausdorffruimte Y is een homeomorfisme. (d.w.z. f^{-1} is ook continu).

Bewijs. We zullen aantonen dat f gesloten verzamelingen van X op gesloten verzamelingen van Y afbeeldt. Dan volgt dat f^{-1} continu is. Stel A een gesloten verzameling van X. Dan is A compact volgens stelling 1 en ook f(A) is compact volgens stelling 3 en dus ook gesloten in Y omdat Y een Hausdorffruimte is.

Opgave. In een compacte ruimte bezit iedere oneindige verzameling minstens één verdichtingspunt.

Compactheid in de Euclidische ruimte.

Met behulp van vb. 1 en 2 op blz. 48 en de stelling op blz. 49 volgt dat de compacte deelverzamelingen van $\mathbb{R}$ precies de begrensde gesloten deelverzameling van $\mathbb{R}$ zijn.

Definitie. Een deelverzameling A van $\mathbb{R}^m$ heet begrensd wanneer er een positief getal M is niet de eigenschap $\rho(x,0) < M$ voor $\forall x \in A$.

Een punt $a \in \mathbb{R}^m$ heet limiet van een rij punten $\{a_n\}_{n=1}^{\infty}$ van $\mathbb{R}^m$ wanneer voor iedere omgeving U van a in $\mathbb{R}^m$ voor bijna alle indices n geldt $a_n \in U$ (of equivalent, wanneer voor iedere $\epsilon > 0$ een index $n_0(\varepsilon)$ bestaat met $\rho(a,a_n) < \varepsilon$ voor $n > n_0(\varepsilon)$).
Een rij punten van $\mathbb{R}^m$ die een limiet heeft, heet convergent. (de limiet is eenduidig bepaald).

Lemma. Iedere begrensde rij punten van $\mathbb{R}^m$ bezit een convergente deelrij.

Voor $m = 1$ volgt dit direct uit de stelling van Bolzana Weierstrass.
Stel $m > 1$ en veronderstel de eigenschap al bewezen voor $k < m$.
Stel $\{a^{(n)}\}_{n=1}^{\infty}$ een begrensde rij in $\mathbb{R}^m$. $a^{(n)} = (a_1^{(n)},\dots,a_m^{(n)})$ voor $n = 1,2,\dots$. De rij $\{b^{(n)}\}_{n=1}^{\infty} = \{(a_1^{(n)},\dots,a_{m-1}^{(n)})\}$ is een begrensde rij punten van $\mathbb{R}^{m-1}$ en bezit dus volgens de inductie-onderstelling een convergente deelrij $\{b^{(n_j)}\}_{j=1}^{\infty}$. Kies ook een convergente deelrij $\{a_m^{(k_j)}\}_{j=1}^{\infty}$ van $\{a_m^{(n_j)}\}_{j=1}^{\infty}$. Dan is $\{(a_1^{(k_j)},\dots,a_m^{(k_j)})\}_{j=1}^{\infty}$ een convergente deelrij van $\{a^{(n)}\}_{n=1}^{\infty}$.

Stelling 5. Voor een deelverzameling A van $\mathbb{R}^m$ gelden de volgende equivalenties

(i) A is compact.

(ii) A is begrensd en gesloten.

(iii) Iedere oneindige deelverzameling van A bezit minstens een verdichtingspunt dat tot A behoort.

(i) → (ii) Stel A compact. Uit een vorige stelling volgt dat A gesloten is in $\mathbb{R}^m$. A is ook begrensd want anders zou de overdekking van A bestaande uit alle open bollen van $\mathbb{R}^m$ met middelpunt 0 en straal n (n=1,2,...) geen eindige deel overdekking bezitten, hetgeen in strijd is met de compactheid van A.

(ii) → (iii) Dit volgt direct uit het vorige lemma.

(iii) → (ii) Ga zelf na.

We bewijzen nu (ii) ⇒ (i). Het is voldoende (ii) ⇒ (i) te bewijzen voor het geval A een kubus van $\mathbb{R}^m$ is $A = \{x \in \mathbb{R}^m \mid |x_i| \leq a_i\}$. Immers A is begrensd d.w.z. bevat in een (voldoend grote) kubus en is hierin ook gesloten. Toepassing van stelling 1, geeft dan het gewenste resultaat.

Onder de diameter van een kubus $L = \{x \in \mathbb{R}^m | a_i \leq x_i \leq b_i\}$ verstaan we de lengte van de langste zijde (= $\max\{(b_i - a_i) | i=1,\ldots,n)\}$

We merken op dat elke kubus L uit $\mathbb{R}^m$ vereniging is van eindig veel kubussen uit $\mathbb{R}^m$ met diameter < $\frac{1}{2}$ diameter L (ga dit na).

Zij $\mathcal{U}$ een open overdekking van K (met open verzamelingen van $\mathbb{R}^m$) en onderstel dat geen eindig deelstelsel van $\mathcal{U}$ bestaat dat K overdekt. Nu is K vereniging van eindig veel kubussen met diameter < $\frac{1}{2}$ en blijkbaar is er dan een kubus K_1 met diameter < $\frac{1}{2}$ die niet door eindig veel U's uit $\mathcal{U}$ overdekt wordt. Met volledige inductie construeren we kubussen $K \supset K_1 \supset K_2 \supset \ldots \supset K_n \supset \ldots$ zodat diameter $K_n < 1/n$ en zodat K_n niet door eindig veel elementen uit $\mathcal{U}$ wordt overdekt. Kies uit elke K_n een punt p_n. De rij $\{p_n\}_{n=1}^{\infty}$ bezit een convergente deelrij die convergeert naar p. Blijkbaar $p \in K_n$ voor alle n want elke K_n is gesloten. Kies $U \in \mathcal{U}$ zodat $p \in U$ (dit is mogelijk want $p \in K$ en $\mathcal{U}$ is een overdekking van K). Kies l zo groot dat$B_{1/l}(p) \subset U$ en zij $s > 2ml$, dan $p \in K_s \subset B_l(p) \subset U$. K_s wordt nu door {U} overdekt in tegenspraak met de onderstelling dat K_s niet door eindig veel elementen van $\mathcal{U}$ wordt overdekt. Hiermede is de stelling bewezen.

<u>Opgave</u>. Bewijs dat de compacte samenhangende deelverzamelingen van $\mathbb{R}$ juist de gesloten intervallen [a,b] zijn (al of niet ontaard).

Opmerking 1. Als f een continue afbeelding is van een compacte ruimte X naar $\mathbb{R}$ dan is blijkbaar f(X) compacte deelverzameling van $\mathbb{R}$ i.e. f(X) is begrensd en gesloten. Is bovendien X samenhangend dan is f(X) noodzakelijk een interval.

I.h.b. volgt dat wanneer A een kubus is in $\mathbb{R}^m$

$A = \{(x = (x_1,\ldots,x_m) \in \mathbb{R}^m | a_i \leq x_i \leq b_i\}$ dat dan f(A) een interval in $\mathbb{R}$ is.

Dus als $f(x_1,\ldots,x_n)$ een grootheid voorstelt die continue afhangt van de variabelen $x_1,\ldots,x_n$ waarbij $a_i \leq x_i \leq b_i$ dan neemt f een maximum en een minimum aan en elke waarde hiertussen wordt door $f(x_1,x_2,\ldots,x_n)$ bereikt.

Opmerking 2. Zij H de Hilbertruimte. Dus $H = \{(x = (x_1,x_2,\ldots,x_n,\ldots)$ $| \sum_{i=1}^{\infty} x_n^2 < \infty\}$ en zij $A \subset H$ gegeven door $H = \{(x_1,\ldots,x_n,\ldots) | \sum_{i=1}^{\infty} x_n^2 = 1\}$.
A is nu wel een begrensde gesloten verzameling van H maar toch is H niet compact. Bovenstaande stelling is dus blijkbaar een eigenschap van de euclidische ruimte $\mathbb{R}^m$ die niet door andere metrische ruimten gedeeld wordt. Ook is niet iedere continue functie op $A \subset H$ begrensd, zodat dit voor begrensde gesloten deelverzamelingen van $\mathbb{R}^m$ het geval is.

Compactheid in willekeurige metrische ruimten.

We vermelden de volgende stelling zonder bewijs.

Stelling 6. In een metrische ruimte M zijn equivalent
(i) M is compact.
(ii) Iedere oneindige deelverzameling van M heeft een verdichtingspunt.
(iii) Iedere rij punten in M heeft een convergente deelrij.

Opmerking. Het begrip "compact" in de zin van "iedere oneindige deelverzameling heeft een verdichtingspunt" is in de literatuur eerder opgetreden dan het begrip "bicompact" (iedere open overdekking een eindige deeloverdekking). Dit laatste is vaak belangrijker gebleken dan het eerste begrip compact.

<u>Stelling 7</u>. Zij f een continue afbeelding van een compacte metrische ruimte (M,ρ) naar een metrische ruimte (N,σ). Dan is f uniform continu.

<u>Bewijs</u>. Zij $\varepsilon > 0$. Voor $x \in M$ kies $k(x) > 0$ zodat
$\rho(x,y) < k(x) \Rightarrow \sigma(f(x), f(y)) < \varepsilon/2$.
Nu is de collectie $\{B_{\frac{k(x)}{2}}(x) | x \in M\}$ een open overdekking van M. Omdat M compact is bestaan er eindig veel $x_1,\ldots,x_n \in M$ zodat $\bigcup\{B_{\frac{k(x_i)}{2}}(x_i) | i=1,\ldots,n\} = M$.
Zij $\delta = \frac{1}{2}\min\{k(x_i) | i=1,\ldots,n\}$. We zullen aantonen dat
$\rho(x,y) < \delta \Rightarrow \sigma(f(x), f(y)) < \varepsilon$.
Stel $\rho(x,y) < \delta$. Er bestaat nu l met $1 \leq l \leq n$ zodat
$x \in B_{\frac{k(x_l)}{2}}(x_l)$. Dus $\rho(x,x_l) < \frac{k(x_l)}{2}$. Ook $\rho(y,x_l) \leq \rho(y,x) + \rho(x,x_l)$
$< \delta + \frac{k(x_l)}{2} < k(x_l)$.
Dus $\sigma(f(x), f(x_l)) < \varepsilon/2$ en $\sigma(f(y), f(x_l)) < \varepsilon/2$.
Hieruit volgt inderdaad $\sigma(f(x), f(y)) < \varepsilon$ (driehoeksongelijkheid).

<u>Gevolg 1.</u> Zij $K = \{x \in \mathbb{R}^m | a_i \leq x_i \leq b_i\}$. Iedere continue functie gedefinieerd op K naar $\mathbb{R}$ is uniform continu.

<u>Gevolg 2</u>. Als ρ een willekeurige metriek is voor de compacte metrische ruimte (X,ρ) dan is iedere continue functie gedefinieerd op X uniform continu.

§5. Volledige metrische ruimten.

<u>Definitie</u>. Een rij punten $\{x_i\}_{i=1}^{\infty}$ van een metrische ruimte M met metriek ρ heet een <u>fundamentaalrij</u> als bij iedere $\varepsilon > 0$ een natuurlijk getal n te vinden is met de volgende eigenschap: als $i > n$ en $j > n$ dan $\rho(x_i,x_j) < \varepsilon$.

Een rij punten $\{x_i\}_{i=1}^{\infty}$ van een metrische ruimte M met metriek ρ heet <u>convergent</u> indien er een punt $x \in M$ bestaat zodat $\rho(x_i,x) < \varepsilon$ voor $i > n_0(\varepsilon)$. Notatie $x = \lim_{n\to\infty} x_n$.

Het is duidelijk dat een convergente rij punten van een metrische ruimte een fundamentaalrij is. Omgekeerd hoeft niet iedere fundamentaalrij uit een metrische ruimte convergent te zijn.

<u>Definitie</u>. Een metrische ruimte (M,ρ) heet <u>volledig</u> wanneer iedere fundamentaalrij convergent is.

<u>Vb.1</u> $\mathbb{R}^m$ met de euclidische metriek is een volledige metrische ruimte.

<u>Vb.2</u> $Q = \{r \in \mathbb{R} | r \text{ is rationaal}\}$ is geen volledige metrische ruimte.

<u>Stelling 1</u>. Iedere compacte metrische ruimte M met metriek ρ is een volledige metrische ruimte.

<u>Bewijs</u>. Zij $\{x_i\}_{i=1}^{\infty}$ een fundamentaalrij uit M.
Als de rij slechts eindig veel verschillende elementen bevat, dan is er een natuurlijk getal n te vinden met $x_i = x_n$ voor $i \geq n$. $\{x_i\}_{i=1}^{\infty}$ convergeert dan naar x_n. Stel nu dat $\{x_i\}_{i=1}^{\infty}$ oneindig veel verschillende elementen bevat. Dan heeft de rij een verdichtingspunt x (zie stelling op blz. 53). We zullen aantonen dat $\{x_i\}_{i=1}^{\infty}$ naar x convergeert. Zij $\rho(x_i,x_j) < \varepsilon/2$ voor $i,j >$ de index n.
Blijkbaar bevat $B_{\varepsilon/2}(x)$ een punt x_l met $l > n$. Zij $n_0 = \max(n,l)$.
Dan geldt voor $i > n_0$ $\rho(x,x_i) \leq \rho(x_i,x_l) + \rho(x_l,x) \leq \varepsilon/2 + \varepsilon/2 = \varepsilon$.

<u>Stelling 2</u>. (Ga na). Een gesloten deelverzameling van een volledige metrische ruimte M is als metrische deelruimte van de metrische ruimte M dan is A gesloten in M.

<u>Voorbeeld.</u> Zij X een compacte Hausdorffruimte. Noteer met C(X) de collectie functies $f: X \to \mathbb{R}$ die continu zijn. Zoals hiervoor aangetoond is is steeds fX compact d.w.z. begrensd en gesloten in $\mathbb{R}$.

Aangezien met f en g ook de functie $|f-g|$, welke gedefinieerd wordt door

$$|f-g|(x) = |f(x)-g(x)| \qquad \text{voor alle } x \in X$$

continu is, kunnen we definieren

$$\rho(f,g) = \max_{x\in X} |f(x)-g(x)|; \text{ men gaat gemakkelijk}$$

na dat aan de drie eisen voor metriek voldaan is.
Door deze definitie wordt $C(X)$ een metrische ruimte. De geïnduceerde topologie op $C(X)$ heet de <u>topologie der uniforme convergentie.</u>
We zullen hieronder aantonen dat $C(X)$ met deze metriek een volledige metrische ruimte is.

<u>Bewijs</u>. Zij $\{f_n\}_{n=1}^{\infty}$ een fundamentaalrij in $C(X)$

$$\rho(f_m,f_n) = \max_{x\in X} |f_n(x)-f_m(x)| < \varepsilon$$

voor $m,n > N(\varepsilon)$.
Voor elk vast punt $x \in X$ geldt

$$|f_m(x)-f_n(x)| \leq \rho(f_m,f_n) < \varepsilon \quad \text{zodra } m,n > N(\varepsilon),$$

met als gevolg $\{f_n(x)\}_{n=1}^{\infty}$ een Cauchy (fundamentaal)rij in $\mathbb{R}$. Aangezien $\mathbb{R}$ volledig is (Stelling van Cauchy) bestaat voor elke vaste $x \in X$ dus $\lim_{n\to\infty} f_n(x)$, deze limiet duiden we aan met $f(x)$. We tonen nog aan dat f continu is op X en dat $\lim_{n\to\infty} \rho(f_n,f) = 0$. Voor elke $x \in X$ geldt

$$|f_m(x)-f_n(x)| \leq \rho(f_m,f_n) < \varepsilon/_3 \quad \text{zodra } m,n > N(\varepsilon);$$

daar $\lim_{n\to\infty} f_n(x)$ bestaat volgt hieruit dat

$\lim_{n\to\infty} |f_m(x)| = |f_m(x)-f(x)| \leq \varepsilon/3$ voor alle $x \in X$, als $m > N(\varepsilon)$.

Het is duidelijk dat voor elk natuurlijk getal n geldt

$$|f(x_0)-f(x)| \leq |f(x_0)-f_n(x_0)| + |f_n(x_0)-f_n(x)| + |f_n(x)-f(x)|;$$

geven we n de vaste waarde $N = N(\varepsilon) + 1$ dan is

$$|f(x)-f_N(x)| \leq \varepsilon/_3 \text{ voor alle } x \in X; \text{ wegens de}$$

continuiteit van f_N is er bij elke $\varepsilon > 0$ een omgeving O_ε van x_0 te vinden zodanig dat

$$|f_N(x_0)-f_N(x)| < \varepsilon/_3 \text{ voor } x \in O_\varepsilon .$$

Voor alle $x \in O_\varepsilon$ geldt dus $|f(x_0)-f(x)| < \varepsilon$, waaruit volgt dat f continu is in $x_0 \in X$. Uit de willekeurigheid van $x_0 \in X$ volgt dat f continu is op X.

Uit $|f_n(x)-f(x)| \leq \varepsilon$ voor alle $x \in X$, als $n > N(\varepsilon)$, volgt tenslotte dat $\rho(f_n,f) \leq \varepsilon$ zodra $n > N(\varepsilon)$, zodat $\lim\limits_{n\to\infty} \rho(f_n,f) = 0$.

Opgave: Zij X een compacte topologische ruimte; (Y,ρ) een volledige metrische ruimte en $C(X,Y)$ de verzameling van alle continue afbeeldingen van X in Y. Definiëren we voor $f, g \in C(X,Y)$

$$\sigma(f,g) = \max_{x\in X} \rho(f(x), g(x)), \text{ dan is } \sigma \text{ een}$$

metriek op $C(X,Y)$, terwijl $C(X,Y)$ met deze metriek een volledige metrische ruimte is.

Voorbeeld. Beschouw $C(X)$ met $X = [0,1]$
en beschouw nu de volgende metriek op $C(X)$

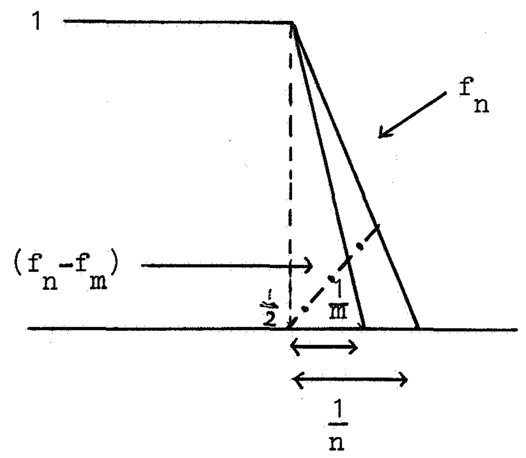
$$\rho^*(f,g) = [\int_a^b \{f(x)-g(x)\}^2 \, dx] \text{ de z.g. euclidische metriek.}$$

Het blijkt nu dat $(C(X),\rho^*)$ niet volledig is!
Immers voor $n = 1,2,\ldots$ zij f_n: $[0,1] \to \mathbb{R}$ gedefinieerd door

$$f_n(x) = 0 \qquad 0 \leq x \leq 1/2$$

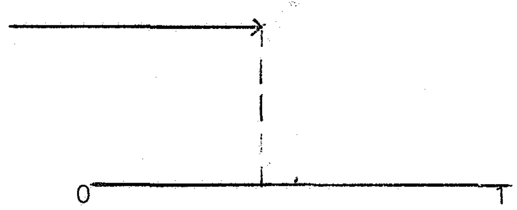
$$f_n(x) = -nx + (1+\tfrac{1}{2}n) \qquad \tfrac{1}{2} < x \leq \tfrac{1}{2}+\tfrac{1}{n}$$

$$f_n(x) = 0 \qquad \tfrac{1}{2}+\tfrac{1}{n} \leq x \leq 1$$



Het is gemakkelijk in te zien dat de f_n's een fundamentaalrij m.b.t. ρ^* vormen.

De limiet is echter de onderstaande functie



en deze is discontinu en behoort daarom niet tot onze collectie $C(X)$.

<u>Contractie afbeeldingen in volledige metrische ruimten.</u>

Zij ϕ een afbeelding van een metrische ruimte (X,ρ) in zichzelf (d.w.z. ϕ is afbeelding van X in X). Dan heet ϕ een <u>(sterke) contractie-afbeelding</u> indien er een $k \in \mathbb{R}$ bestaat , $0 \leq k < 1$ zodat

$$\rho(\phi(x), \phi(y)) < k.\rho(x,y).$$

Wanneer we slechts eisen

$$\rho(\phi(x), \phi(y)) < \rho(x,y), \text{ dan heet } \phi \text{ een}$$

(gewone) contractie afbeelding.

Stelling 3. Als ϕ een sterke contractie is in een volledige metrische ruimte X dan bezit ϕ precies één dekpunt (d.i. een punt $x \in X$ zodat $\phi(\hat{x}) = \hat{x}$)

Bewijs. Eerst tonen we aan dat zo'n punt $\hat{x}$ bestaat.
Zij x_0 een willekeurig doch vast punt van x; definieer
$x_1 = \phi(x_0)$, $x_2 = \phi(x_1)$, $x_3 = \phi(x_2)$, ...
Wegens

$$\rho(x_{n+1},x_n) = \rho(\phi(x_n), \phi(x_{n-1})) \leq k\rho(x_n,x_{n-1}) \leq \ldots \leq k^n\rho(x_1,x_0)$$

en $\rho(x_{n+p},x_n) \leq \rho(x_{n+p},x_{n+p-1}) + \ldots + \rho(x_{n+1},x_n)$

geldt voor de rij $\{x_n\}_{n=1}^{\infty}$

$$\rho(x_{n+p},x_n) \leq (k^{n+p-1}+\ldots+k^n) \cdot \rho(x_1,x_0) \leq \frac{k^n}{1-k} \cdot \rho(x_1,x_0)$$

met als gevolg $\rho(x_{n+p},x_n) < \varepsilon$ als n maar voldoende groot gekozen wordt.

De rij $\{x_n\}_{n=1}^{\infty}$ is dus een fundamentaalrij in X; daar X volledig is, is $\{x_n\}_{n=1}^{\infty}$ convergeert met als limiet $\hat{x} \in X$.
Voor elk natuurlijk getal n geldt

$$\rho(\hat{x},\phi(\hat{x})) \leq \rho(\hat{x},x_n) + \rho(x_n,\phi(\hat{x})) =$$

$$= \rho(\hat{x},x_n) + \rho(\phi(x_{n-1}), \phi(\hat{x})) \leq \rho(\hat{x},x_n) + k \cdot \rho(x_{n-1},\hat{x}).$$

Wegens $\lim_{n\to\infty} \rho(\hat{x},x_n) = 0$ kunnen we hieruit concluderen dat

$$\rho(\hat{x},\phi(\hat{x})) = 0 \quad \text{of} \quad \hat{x} = \phi(\hat{x}).$$

X bevat dus zeker één dekpunt van ϕ.
Veronderstel dat $\hat{x}_1$ en $\hat{x}_2$ dekpunten zijn van ϕ; dan geldt

$$\rho(\hat{x}_1,\hat{x}_2) = \rho(\phi(\hat{x}_1), \phi(\hat{x}_2)) \leq k.\ \rho(\hat{x}_1,\hat{x}_2),$$

maar dit is slechts mogelijk indien $\hat{x}_1 = \hat{x}_2$.
Er is dus slechts één dekpunt.

Opgave. Is (X,ρ) een metrische ruimte en ϕ een afbeelding van X in zichzelf zodanig dat
(i) $\phi(X)$ is precompact (d.w.z. de afsluiting van $\phi(X)$ is compact)
(ii) ϕ is een (gewone)contractieafbeelding,

dan bevat X precies één dekpunt van ϕ.
Aanwijzing: beschouw de functie $\rho(x,\phi(x))$.

Toepassing. Is ϕ een differentieerbare afbeelding van het interval $[0,1]$ in zichzelf, terwijl voor alle $x \in (0,1)$ geldt

$$|\phi'(x)| \leq k < 1 ,$$

dan heeft de vergelijking $x = \phi(x)$ precies een oplossing op $[0,1]$.

Bewijs. Het interval $[0,1]$ is, met de gewone metriek een volledige metrische ruimte. Met behulp van de eerste middelwaardestelling uit de uit de diff.rekening (zie later) vinden we

$$|\phi(x_1)-\phi(x_2)| = |x_1-x_2|.|\phi'(\xi)| \quad (x_1 < \xi < x_2)$$

zodat ϕ een contractieoperator is. Volgens de vorige stelling bezit ϕ precies één dekpunt.

Toepassing 2.
Gevraagd wordt aan te tonen dat de lineaire integraal vergelijking in h

$$h(x) = g(x) + \lambda \int_a^b K(x,t)\ h(t)\ dt$$

voor voldoend kleine waarden van $|\lambda|$ eenduidig oplosbaar is. Aangenomen mag worden dat h en g continu zijn op [a,b] en dat de kern K continu is op het vierkant $a \leq x, t \leq b$.

Oplossing. Zij X de verzameling van alle reële continue functies f op het interval [a,b]; nemen we als metriek

$$\rho(f,g) = \max_{x\in[a,b]} |f(x)-g(x)|$$

op X, dan is X volledig volgens stelling 1.

Definieer ϕ: X $\rightarrow$ X volgens

$$\phi(f)(x) = g(x) + \lambda \int_a^b K(x,t)\, f(t)dt \ .$$

We gaan na of er waarden van λ zijn waarvoor ϕ een contractieoperator is op X.

Uit

$$\rho(\phi(f_1), \phi(f_2)) = \max_{x\in[a,b]} |\phi(f_1)(x)-\phi(f_2)(x)| =$$

$$= \max_{x\in[a,b]} \left|\lambda \int_a^b K(x,t)(f_1(t) - f_2(t))dt\right| \leq |\lambda| \max_{x\in[a,b]} \int_a^b |K(x,t)|dt.$$

$\rho(f,f_2)$ volgt dat ϕ contrakerend is voor

$$|\lambda| < (\max_{x\in[a,b]} \int_a^b |K(x,t)|dt)^{-1} \ .$$

Volgens de stelling is de integraalvergelijking voor deze waarden van λ dus eenduidig oplosbaar.

Hoofdstuk 4. Lineaire Algebra en Lineaire Analyse

§1. Lineaire ruimten, vectorrruimten, basis en dimensie

Definitie. Een (reële) lineaire ruimte L is een verzameling L met de eigenschappen A_1 en A_2:

A_1. Er is een afbeelding van $L \times L$ in L uitverkoren. Deze heet de optelling van elementen. Het beeld van het paar $(a,b) \in L \times L$ heet de som van a en b, en wordt voorgesteld door $a+b \in L$. Onder de optelling is L een abelse groep d.w.z. aan de volgende eisen is voldaan:

1) $(a+b) + c = a + (b+c)$
2) $a + b = b + a$
3) Er bestaat één element $0 \in L$ zodat $a + 0 = 0 + a = a$ voor alle $a \in L$
4) Voor elke $a \in L$ bestaat 1 element $a^* \in L$ zodat $a + a^* = 0$. I.p.v. a^* schrijven we $-a$.

A_2. Er is een uitverkoren afbeelding van $\mathbb{R} \times L$ in L. Deze heet de scalaire vermenigvuldiging van elementen met getallen. Het beeld van het paar $(\lambda,a) \in \mathbb{R} \times L$ heet product en noteren we met $\lambda a \in L$.
Er is aan de volgende eigenschappen voldaan:

5) $\lambda(a+b) = \lambda a + \lambda b$
6) $(\lambda+\mu)a = \lambda a + \mu a$
7) $(\lambda\mu)a = \lambda(\mu a)$
8) $1.a = a$

Zeer belangrijke opmerking. De scalaire vermenigvuldiging kan ook m.b.v. een ander algebraisch systeem gedefinieerd worden. Als S een systeem is waarin een optelling en vermenigvuldiging is gedefinieerd zodat aan alle algebraïsche eisen van pagina 14 voldaan is, dan kan men een z.g. lineaire ruimte over S definiëren. Hieronder verstaat men dan een verzameling L waarin een optelling, en een vermenigvuldiging met elementen uit S is gedefinieerd zodat aan A_1 en A_2 voldaan is (met $\mathbb{R}$ vervangen door S). I.h.b. kan men voor S het systeem (lichaam) der complexe getallen nemen (zie blz. 26A). Men spreekt dan van een complexe lineaire ruimte. Hieronder beschouwen we voorlopig slechts reële ruimten.

Stelling 1. Zij L een lineaire ruimte. Voor willekeurige $\alpha,\beta\in\mathbb{R}$ en $x\in L$ geldt:

a) $\alpha.0 = 0$

b) $0.x = 0$

c) $\alpha.x = 0 \Longleftrightarrow \alpha = 0$ of $x = 0$.

d) $(-\alpha).x = -(\alpha.x) = \alpha.(-x)$

e) $\alpha.(x-y) = \alpha.x - \alpha.y$

f) $(\alpha-\beta).x = \alpha.x - \beta x$

Bewijs. a) $\alpha.0 = \alpha.(0+0) = \alpha.0 + \alpha.0 \Longrightarrow 0 = \alpha.0$

b) $0.x = (0+0).x = 0.x + 0.x \Longrightarrow 0 = 0.x$

c) Stel $\alpha x = 0$ maar $\alpha \neq 0$. Dan is $x = 1.x = (\alpha^{-1}\alpha).x = \alpha^{-1}(\alpha x) = \alpha^{-1}.0 = 0$

d) $\alpha.x + (-\alpha).x = (\alpha+(-\alpha)).x = 0.x = 0 \Longrightarrow (-\alpha).x = -(\alpha x)$

e) $\alpha(x-y) = \alpha.x + \alpha(-y) = \alpha x - \alpha y$

f) $(\alpha-\beta)x = \alpha x + (-\beta)x = \alpha x - \beta x$

Voorbeelden. 1) De vectoren in het platte vlak vormen met deze definitie een lineaire ruimte. Hetzelfde geldt voor de vectoren in de ruimte.

2) Zij $\mathbb{R}^n$ de verzameling van alle geordende n-tallen reële getallen $x = (x_1,\dots,x_n)$ met als optelling

$$(x_1,x_2,\dots,x_n) + (y_1,\dots,y_n) = ((x_1+y_1),\dots,(x_n+y_n)) .$$

Voor $\alpha\in\mathbb{R}$ en $x = (x_1,\dots,x_n)\in\mathbb{R}^n$ zij

$$\alpha x = (\alpha x_1,\dots,\alpha x_n) .$$

Dan is $\mathbb{R}^n$ met deze definities een lineaire ruimte. $\mathbb{R}^n$ heet ook wel de numerieke n-dimensionale vectorruimte.

Opmerking. Dikwijls schrijven we $x = (x_1,\dots,x_n)$ ook als

$$x = \begin{pmatrix} x_1 \\ \cdot \\ \cdot \\ x_n \end{pmatrix}$$

m.a.w. we schrijven de vector x als kolomvector.

3) Een algemener voorbeeld verkrijgen we door uit te gaan van een willekeurige verzameling V (die in plaats treedt van de $\{1,2,\ldots,n\}$ van zoëven).
Zij $\mathbb{R}^V$ de verzameling van alle functies gedefinieerd op V en met waarden in $\mathbb{R}$. Als $f,g\in\mathbb{R}^V$, dan zij $f+g\in\mathbb{R}^V$ die functie waarvoor

$$(f+g)(x) = f(x) + g(x) \qquad \text{voor } x\in V.$$

Als verder $f\in\mathbb{R}^V$ en $\lambda\in\mathbb{R}$ laat dan $\lambda f\in\mathbb{R}^V$ gedefinieerd zijn door $(\lambda f)(x)=\lambda f(x)$.
Op deze wijze wordt $\mathbb{R}^V$ tot een lineaire ruimte.

4) Stel $a < b \in \mathbb{R}$. Laat $C(a,b) = \{f:[a,b] \to \mathbb{R} \mid f \text{ continu}\}$. Als $f,g\in C(a,b)$ en $\lambda\in\mathbb{R}$ definieer dan $f+g$ en λf als volgt

$$(*) \qquad (f+g)(x)= f(x) + g(x) \qquad x\in[a,b]$$

$$(**) \qquad (\lambda f)(x) = \lambda.f(x)$$

Dan is $C(a,b)$ een lineaire ruimte.

5) Stel $g_0,g_1,g_2,\ldots,g_n$ reëelwaardige functies gedefinieerd op het interval $[a,b]$. Zij E de verzameling van alle functies die n-maal differentieerbaar zijn op $[a,b]$ en die voldoen aan de volgende differentiaalvergelijking

$$g_n(x)\,\frac{d^n f(x)}{dx^n} + g_{n-1}(x).\,\frac{d^{n-1}f(x)}{dx^{n-1}} + \ldots + g_1(x).\,\frac{df(x)}{dx} = g_0(x)\ .$$

Indien we de optelling in E en de vermenigvuldiging met een reëel getal weer definiëren m.b.v. (*) en (**) dan is E een lineaire ruimte.

Zij nu L een lineaire ruimte en $B\subset L$ heeft de eigenschap dat bij elke $a\in L$ $\exists\, b_1,\ldots,b_n\in B$ zodat

$$a = \lambda_1 b_1 + \lambda_2 b_2 + \ldots \lambda_n b_n \qquad \text{voor zekere } \lambda_i\in\mathbb{R},\ i = 1,\ldots,n,$$

dan heet $B = \{b_i \mid i\in I\}$ een stel <u>voortbrengenden</u> van L. Is geen echte deelverzameling van B weer een stel voortbrengenden van L, dan heet B een <u>basis</u> voor L.

Men kan aantonen dat iedere lineaire ruimte een basis bezit. Heeft men te maken met een lineaire ruimte die een <u>eindig aantal voortbrengenden</u> $B = \{b_1,\ldots,b_n\}$ <u>bezit</u>, dan kan men, als B geen basis is, minstens één vector uit B weg laten en toch een stel voortbrengenden overhouden. Dit herhalend vindt men in eindig veel stappen een basis.

<u>Definitie</u>. Een lineaire ruimte A heet een <u>vectorruimte</u> indien A een eindige basis bezit.
Het kleinste getal dat als aantal van een basis voorkomt heet de <u>dimensie van</u> A. Bestaat A uit precies één vector, dan heet de dimensie daarentegen nul.

<u>V.b.</u> In $\mathbb{R}^n$ beschouwen we het volgende stelsel voortbrengenden

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ldots \quad e_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 1 \end{pmatrix}$$

. Ga na dat $\{e_1,\ldots,e_n\}$ ook een basis vormen voor $\mathbb{R}^n$. Probeer te bewijzen dat de dimensie van $\mathbb{R}^n$ precies n is (zie ook later). Dus $\mathbb{R}^n$ is een vectorruimte.

<u>Opmerking</u>. We spreken dus alleen over een vectorruimte wanneer de ruimte een eindige dimensie bezit. Is er geen sprake van een eindige dimensie (zoals b.v. in de genoemde voorbeelden 3, 4, 5) dan spreken we van een lineaire ruimte. We zullen ook alleen over vectoren spreken wanneer het elementen van vectorruimten betreft.

<u>Definitie</u>. Zij L een lineaire ruimte. Dan heet een element $b \in L$ <u>lineair afhankelijk</u> van een stelsel elementen $\{a_1,a_2,\ldots,a_n\}$ van L wanneer er $\lambda_i \in \mathbb{R}$ bestaan, $i = 1,2,\ldots,n$; λ_i niet alle nul, zodat

$$(1) \qquad b = \lambda_1 a_1 + \lambda_2 a_2 + \ldots + \lambda_n a_n .$$

Een stelsel elementen $\{a_1,\ldots,a_n\}$ heet <u>lineair afhankelijk</u> indien er een i $1 \le i \le n$ bestaat zodat a_i afhankelijk is van de overige elementen $a_1, a_2, \ldots, a_{i-1}, a_{i+1}, a_{i+2}, \ldots, a_n$. Het stelsel $\{a_1, a_2, \ldots, a_n\}$ heet <u>lineair onafhankelijk</u> indien het niet lineair afhankelijk is. $\{a_1, a_2, \ldots, a_n\}$ is lineair onafhankelijk d.e.s.d. indien

$$\lambda_1 a_1 + \lambda_2 a_2 + \ldots + \lambda_n a_n = 0$$

impliceert $\lambda_1 = \lambda_2 = \ldots = \lambda_n = 0$.

Ten aanzien van lineaire afhankelijkheid gelden de volgende drie grondeigenschappen.

<u>Grondeigenschap I</u>

Ieder a_i $(i = 1, \ldots, n)$ is lineair afhankelijk van $\{a_1, \ldots, a_n\}$.

<u>Grondeigenschap II</u>

Is b lineair afhankelijk van $a_1, \ldots, a_n$, echter niet van $a_1, \ldots, a_{n-1}$, dan is a_n lineair afhankelijk van $a_1, \ldots, a_{n-1}, b$.

<u>Grondeigenschap III</u>

Is c lineair afhankelijk van $a_1, \ldots, a_s$ en is iedere a_j lineair afhankelijk van $b_1, \ldots, b_n$ dan is c lineair afhankelijk van $b_1, \ldots, b_n$.

<u>Stelling 2</u>. Een basis $a_1, \ldots, a_m$ van een vectorruimte A bestaat uit lineair onafhankelijke vectoren. Omgekeerd is elk lineair onafhankelijk stel voortbrengenden van A een basis.

<u>Bewijs</u>. Stel een der vectoren $a_1, \ldots, a_m$ van een basis is lineair afhankelijk van de anderen. Na eventueel vernummeren mag aangenomen worden dat dit a_m is. Dus a_m lineair afhankelijk van $a_1, \ldots, a_{m-1}$; omdat elke vector b uit de ruimte lineair afhankelijk is van $a_1, \ldots, a_m$ volgt uit grondeigenschap III dat b ook lineair afhankelijk is van $a_1, \ldots, a_{m-1}$ in tegenspraak met het feit dat $a_1, \ldots, a_m$ een basis is.

Was omgekeerd een onafhankelijk stelsel voortbrengenden $a_1, \ldots, a_m$ van A niet een basis, dan kan een b.v. a_m gemist worden. Maar dan zou a_m lineair afhankelijk zijn van $a_1, \ldots, a_{m-1}$.

Definitie. Twee stelsels vectoren $a_1,\dots,a_r$ en $b_1,\dots,b_s$ heten equivalent (notatie $(a_1,\dots,a_r) \sim (b_1,\dots,b_s)$) indien iedere a_k van $(b_1,\dots,b_s)$ en iedere b_i van $a_1,\dots,a_r$ lineair afhankelijk is.

Uitwisselingslemma (Steinitz)

Zij $b_1,\dots,b_s$ lineair onafhankelijk en zij elke b_j lineair afhankelijk van $a_1,\dots,a_r$ dan is er in het systeem der a_j een deelsysteem $\{a_{i_1},\dots,a_{i_s}\}$ van precies s elementen welke men tegen $b_1,\dots,b_s$ kan uitwisselen zodat het door deze uitwisseling ontstane systeem uit $\{a_1,\dots,a_r\}$ equivalent is met het oorspronkelijke systeem $\{a_1,\dots,a_r\}$. I.h.b. $s \leq r$.

Bewijs. Voor $s = 0$ is de bewering triviaal. Laat de stelling bewezen zijn voor $\{b_1,\dots,b_{s-1}\}$ en zij $\{b_1,\dots,b_{s-1}\}$ tegen $\{a_{i_1},\dots,a_{i_{s-1}}\}$ uitgewisseld. Door deze uitwisseling ontstaat een met $\{a_1,\dots,a_r\}$ equivalent systeem $\{b_1,\dots,b_{s-1},a_k,a_l,\dots\}$. Nu is b_s ook van het equivalente systeem $\{b_1,\dots,b_{s-1},a_k,a_l,\dots\}$ van $\{a_1,\dots,a_r\}$ lineair afhankelijk. Er bestaat dus ook een kleinste deelverzameling van $\{b_1,\dots,b_{s-1},a_k,a_l,\dots\}$ waarvan b_s lineair afhankelijk is. Dit kleinste systeem kan niet uit uitsluitend a_j bestaan omdat de b_j's lineair afhankelijk zijn. Dus bevat deze kleinste deelverzameling minstens één a_k, deze noemen we a_{i_s}. Uit grondeigenschap II volgt dat $a_k = a_{i_s}$ lineair afhankelijk is van het systeem dat uit $\{b_j,\dots,a_k\}$ door vervanging van a_k door b_s ontstaat en ook van het omvattende systeem dat uit $\{b_1,\dots,b_{s-1},a_k,a_l,\dots\}$ door de vervanging $a_k \to b_s$ ontstaat. Dit systeem is $\{b_1,\dots,b_s,a_k,a_l,\dots\}$. Het is met $\{b_1,\dots,b_{s-1},a_k,a_l,\dots\}$ equivalent omdat a_k van het eerste systeem en b_s van het laatste lineair afhankelijk is. Hieruit volgt dat het nieuwe systeem $\{b_1,\dots,b_{s-1},b_k,a_l,\dots\}$ equivalent is met $\{b_1,\dots,b_{s-1},a_k,a_l\}$ en ook met het oorspronkelijke systeem $\{a_1,\dots,a_r\}$.

Gevolg. Twee equivalente lineair onafhankelijke systemen bestaan uit evenveel elementen.

Een belangrijk gevolg is ook de volgende stelling.

Dimensiestelling. Elke basis van een n-dimensionale vectorruimte A telt n vectoren. Elk stelsel lineair onafhankelijke vectoren kan aangevuld worden tot een basis. Men kan dus definiëren: de dimensie van een vectorruimte is het aantal vectoren van een basis.

Stelling 3. Is $a_1,\dots,a_m$ een basis van A, dan bestaat bij elke vector $b\in A$ precies één stel getallen $\lambda_1,\dots,\lambda_m$ zodat (1) geldt.

Bewijs. Is $\mu_1 a_1 + \dots + \mu_m a_m = \nu_1 a_1 + \dots + \nu_n a_n$ dan is $(\mu_1-\nu_1)a_1 + \dots + (\mu_m-\nu_m)a_m = 0$ en wegens onafhankelijkheid van $a_1,\dots,a_m$ $\mu_1 - \nu_1 = 0,\dots,\mu_m-\nu_m = 0$.

Deelruimten

Zij B een deelverzameling van de lineaire ruimte A. Als voor elke $a,b\in B$ en $\lambda\in R$ ook $\lambda a\in B$ en $a+b\in B$, dan is ook B met de optelling en scalaire vermenigvuldiging van A een lineaire ruimte. B heet een deelruimte van de lineaire ruimte A. De dimensie van B is kleiner of gelijk aan de dimensie van A. (Ga dit na). (Voor het geval A een eindige dimensie heeft). Als dim B = dim A dan volgt ook B = A (direct gevolg van de dimensiestelling).

Uit de dimensiestelling volgt verder dat als $b_1,\dots,b_m$ een basis is voor de deelruimte B, dan kan dit stelsel vectoren aangevuld worden tot een basis voor A.

Zij B, C deelverzamelingen van een lineaire ruimte A. Dan $B + C = \{a+b \mid a\in B\ b\in C\}$. Zijn B en C deelruimten van A dan is ook B + C een deelruimte van A. B + C heet de door B en C opgespannen deelruimte van A. Verder is ook steeds $B\cap C$ een deelruimte van A.

2e dimensiestelling. Zijn B en C twee deelruimten van een vectorruimte A met doorsnede $B\cap C$, die samen B + C opspannen, dan geldt voor de dimensies:

$$\dim(B+C) = \dim B + \dim C - \dim(B\cap C) .$$

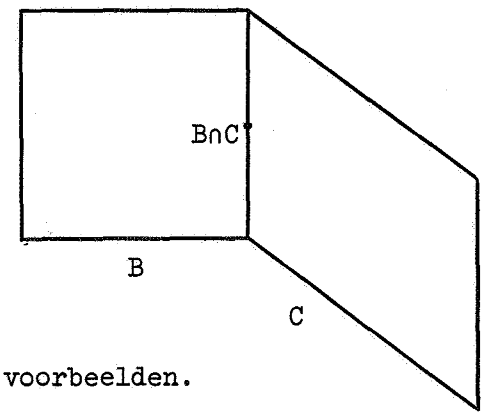
<u>Bewijs</u>. Stel $\dim(B\cap C) = r$, $\dim B = r + s$, $\dim C = r + t$. Kies een basis $a_1,\dots,a_r$ in $B\cap C$ en vul deze aan enerzijds tot een basis $a_1,\dots,a_r,b_1,\dots,b_s$ van B, anderzijds tot een basis $a_1,\dots,a_r,c_1,\dots,c_t$ van C.

Nu is $a_1,\dots,a_r,b_1,\dots,b_s,c_1,\dots,c_t$ een stelsel voortbrengenden van $B + C$. Het is een basis, immers, is

$$(*) \qquad \sum_1^r \alpha_i a_i + \sum_1^s \beta_j b_j + \sum_1^t \gamma_k c_k = 0$$

Dan is $-\sum_1^s \beta_j b_j$ niet alleen in B, maar ook in C, dus in $B\cap C$. Maar dan is deze vector nul, want anders waren $b_1,\dots,b_s$ en $a_1,\dots,a_r$ afhankelijk. Nu is $\beta_1 = \dots = \beta_s = 0$ en $\sum_1^r \alpha_i a_i + \sum_1^t \gamma_k c_k = 0$. Wegens de onafhankelijkheid van $a_1,\dots,a_r,c_1,\dots,c_t$ zijn alle coëfficiënten in (*) en nul. De dimensie van $B + C$ is dus $r + s + t$, hetgeen het gestelde inhoudt.

In de volgende figuur snijden de 2 dimensionale vectorruimten B en C elkaar volgens de 1-dimensionale ruimte $B\cap C$; zij spannen samen de 3-dimensionale ruimte $B + C$ op.



Nog enkele voorbeelden.

1) Zij A een vectorruimte, en $a_1,\dots,a_n$ vectoren uit A. Dan is de verzameling B bestaande uit alle vectoren $\lambda_1 a_1 + \dots + \lambda_n a_n$ met $\lambda_i \in \mathbb{R}$ $i = 1,\dots,n$ een lineaire deelruimte van A. Deze heet wel de <u>ruimte opgespannen</u> door $a_1,\dots,a_n$.

2) Laat n een natuurlijk getal zijn, en laat $\mathbb{R}^n$ de numerieke n-dimensionale vectorruimte zijn. Laat getallen a_{ij} $(i=1,\dots,p\ ,\ j=1,\dots,n)$

gegeven zijn, en laat B bestaan uit alle n-vectoren $(x_1,\ldots,x_n)$ die voldoen aan het stelsel vergelijkingen

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = 0 \\ \quad\vdots \qquad\qquad\qquad\qquad\quad \vdots \\ a_{p1}x_1 + a_{p2}x_2 + \ldots + a_{pn}x_n = 0 \end{cases}$$

Dan is B een lineaire deelruimte van R^n. De dimensie m van B is ten hoogste n. Als m = n dan is $R^n = B$.

§2. Homomorfismen van vectorruimten, beeld en kern.

Definitie. Een afbeelding σ van een lineaire ruimte L in een lineaire ruimte M heet lineair, indien voor elke $a_1, a_2 \in L$, $\lambda \in \mathbb{R}$ geldt:

$$\sigma(\lambda a_1) = \lambda.\sigma(a_1)$$

$$\sigma(a_1+a_2) = \sigma a_1 + \sigma a_2$$

Herhaalde toepassing van deze formules levert

$$\sigma(\lambda_1 a_1+\ldots+\lambda_n a_n) = \lambda_1 \sigma a_1 + \ldots + \lambda_n \sigma a_n .$$

Een lineaire afbeelding σ: L → M heet een homomorfisme en indien L = M ook een endomorfisme.

Enige voorbeelden

1. In het platte vlak geeft de draaiing over een vaste hoek om de oorsprong een lineaire afbeelding waarvoor bovenstaande betrekkingen gelden.

2. De meetkundige vermenigvuldiging van alle elementen in een lineaire ruimte L met een factor $\lambda \neq 0$. Dit endomorfisme wordt kort door λ aangeduid, λ: L → L.

3. Definieer de toevoeging $(x_1,\ldots,x_m) \in \mathbb{R}^m \longmapsto (y_1,\ldots,y_n) \in \mathbb{R}^n$ als volgt

$$\begin{cases} y_1 = a_{11}x_1 + \ldots + a_{1m}x_n \\ \cdot \\ \cdot \\ \cdot \\ y_n = a_{n1}x_1 + \ldots + a_{nm}x_n \end{cases} \qquad (a_{ij} \in \mathbb{R})$$

Deze toevoeging induceert een lineaire afbeelding A: $\mathbb{R}^m \longrightarrow \mathbb{R}^n$. We zullen hier nog later op terugkomen.

LEMMA. Zijn de elementen $a_1,\ldots,a_r \in L$ lineair afhankelijk en is σ: L → M een homomorfisme, dan zijn ook $\sigma a_1,\ldots,\sigma a_r$ lineair afhankelijk in M.

Bewijs. Ter inleiding merken we op $\sigma(0) = \sigma(0.a) = 0.\sigma a = 0$. Zijn $a_1,\ldots,a_r$ lineair afhankelijk dan bestaan $\lambda_1,\ldots,\lambda_r$ R met alle nul, zodat

$$\lambda_1 a_1 + \ldots + \lambda_r a_r = 0$$

dus

$$\lambda_1 \sigma a_1 + \ldots + \lambda_r \sigma a_r = \sigma(\lambda_1 a_1 + \ldots + \lambda_r a_r) = \sigma(0) = 0$$

Stelling 1. Een homomorfisme σ van een vectorruimte A naar een vectorruimte B is uniek bepaald door haar werking op de vectoren van een basis van A. Vormen $a_1,\ldots,a_m$ een basis van A en zijn $b_1,\ldots,b_m \in B$, dan bestaat precies één homomorfisme σ waarvoor $\sigma a_i = b_i$, $i = 1,\ldots,m$.

Bewijs. Is $a_1,\ldots,a_m$ een basis, dan bestaat bij elke $a \in A$ juist één stel getallen $\lambda_1,\ldots,\lambda_m$ zodat

$$a = \lambda_1 a_1 + \ldots + \lambda_m a_m$$

Het eerst gestelde volgt dan uit het feit dat

$$\sigma(a) = \lambda_1 \sigma a_1 + \ldots + \lambda_m \sigma a_m = \lambda_1 b_1 + \ldots + \lambda_m b_m .$$

Zijn omgekeerd $b_1,b_2,\ldots,b_m \in B$ dan voldoet de afbeelding σ: $a = \lambda_1 a_1 + \ldots + \lambda_m a_m \longmapsto \sigma(a) = \lambda_1 b_1 + \ldots + \lambda_m b_m$ aan de lineariteitseigenschappen en $\sigma a_i = b_i$; σ is dan het gewenste homomorfisme.

Definitie. Is $\sigma: A \to B$ een homomorfisme, dan heet σA het beeld, $\sigma^{-1}(0)$ de kern, A de bron van σ, en de dimensie van σA (voor zover die eindig is) heet de rang van σ.

Stelling 2. De kern $\sigma^{-1}(0)$ en het beeld σA van een homomorfisme $\sigma: A \to B$ zijn lineaire deelruimten van respectievelijk A en B. De som van de dimensies van kern en beeld is gelijk aan de dimensie van de bron: $\dim \sigma^{-1}(0) + \dim \sigma(A) = \dim A$

$$\dim \text{kern} + \dim \text{beeld} = \dim \text{bron}.$$

Bewijs. Dat $\sigma^{-1}(0)$ en σA lineaire deelruimten zijn van A resp. B is duidelijk. Immers $a,b\in\sigma^{-1}(0) \Longrightarrow \sigma(\lambda a+\mu b) = \lambda(\sigma a) + \mu(\sigma b) = 0 \Longrightarrow$ $\Longrightarrow \lambda a+\mu b\in\sigma^{-1}(0)$. Voor de tweede bewering geldt een analoge redenering.

Om bovenstaande formule te bewijzen stellen we de dimensies van A en $\sigma^{-1}(0)$ respectievelijk m en m-r. Kies een basis $a_{r+1},\ldots,a_m$ voor $\sigma^{-1}(0)$, en vul deze aan tot een basis $a_1,\ldots,a_m$ van A.

Dan is

$$\sigma a_i = 0 \qquad \text{voor } i > r.$$

We noemen

$$\sigma a_i = b_i \qquad \text{voor } i \leq r.$$

De vectoren $b_1,\ldots,b_r$ vormen een stelsel voortbrengenden van σA, immers voor $a\in A$ is

$$\sigma a = \sigma(\lambda_1 a_1 + \ldots + \lambda_m a_m) = \lambda_1 b_1 + \ldots + \lambda_r b_r \; .$$

Zij vormen zelfs een basis. Veronderstel n.l. dat $\lambda_1,\ldots,\lambda_r\in R$ en

$$\lambda_1 b_1 + \ldots + \lambda_r b_r = 0$$

Dan is $\quad \lambda_1\sigma a_1 + \ldots + \lambda_r\sigma a_r = \sigma(\lambda_1 a_1 + \ldots + \lambda_r a_r) = 0$

Dus $\quad \lambda_1 a_1 + \ldots + \lambda_r a_r\in\sigma^{-1}(0)$. Omdat $\{a_{r+1},\ldots,a_m\}$ een basis is voor $\sigma^{-1}(0)$ bestaan er getallen $\lambda_{r+1},\ldots,\lambda_m$ zodat

$$\lambda_1 a_1 + \ldots + \lambda_r a_r = \lambda_{r+1}a_{r+1} + \ldots + \lambda_m a_m \; .$$

Omdat $a_1,\ldots,a_m$ onafhankelijk zijn volgt dat $\lambda_1 = \lambda_2 = \ldots \lambda_r = 0$. De dimensie van $\sigma(A)$ is dus gelijk aan r, waarmede het gestelde is aangetoond.

Opmerking. Als σ een homomorfisme is van de vectorruimte A in B en dim A $\geq$ dim B, dan impliceert σ is injectief dus automatisch σ is surjectief. De corresponderende eigenschap voor lineaire ruimten is onjuist.

<u>Matrixvoorstelling van een lineaire transformatie.</u>

Zij $a_1,\dots,a_m$ een basis voor A en $b_1,\dots,b_n$ een basis voor B. Een willekeurig homomorfisme is bepaald door mn getallen σ_{ij} $i = 1,\dots,n$; $j = 1,\dots,m$, n.l. door de afbeelding

$$\sigma \begin{cases} a_1 \to \sigma_{11}b_1 + \sigma_{21}b_2 + \dots + \sigma_{n1}b_n \in B \\ a_2 \to \sigma_{12}b_1 + \sigma_{22}b_2 + \dots + \sigma_{n2}b_n \in B \\ \cdot \\ \cdot \\ \cdot \\ a_m \to \sigma_{1m}b_1 + \sigma_{2m}b_2 + \dots + \sigma_{nm}b_n \in B \end{cases}$$

Het schema

$$\begin{bmatrix} \sigma_{11} & \sigma_{12} & \dots & \sigma_{1m} \\ \sigma_{21} & \sigma_{22} & \dots & \sigma_{2m} \\ \cdot & & & \\ \cdot & & & \\ \cdot & & & \\ \sigma_{n1} & \sigma_{n2} & \dots & \sigma_{nm} \end{bmatrix}$$

van de getallen σ_{ij} $(1 \leq i \leq n,\ 1 \leq j \leq m)$ noemt men de <u>matrixvoorstelling van</u> σ <u>t.o.v. de bases</u> $a_1,\dots,a_m$ <u>en</u> $b_1,\dots,b_n$. Merk op dat de matrixvoorstelling van een homomorfisme afhankelijk is van de keuze van $a_1,\dots,a_m$ en $b_1,\dots,b_n$.

Wanneer σ: $A \to A$ een endomorfisme is en $a_1 = b_1,\dots,a_n = b_n$ dan spreken we van de <u>matrixvoorstelling van</u> σ <u>t.o.v. de basis</u> $a_1,\dots,a_n$.

Een deelschema

$$\begin{bmatrix} \sigma_{1j} \\ \sigma_{2j} \\ \cdot \\ \cdot \\ \cdot \\ \sigma_{nj} \end{bmatrix}$$

heet de j^{de} kolom van bovenstaande $n \times m$ matrix.

Een deelschema

$$[\sigma_{k1},\ldots,\sigma_{km}]$$

heet de k^{de} rij van de matrix (σ_{ij}).

Als $x = x_1a_1 + \ldots + x_ma_m$ dan bepaalt men het beeld $\sigma x = y_1b_1 + \ldots + y_nb_n$ van x m.b.v. de matrix σ_{ij} formule $y_i = \sum_{j=1}^{m} \sigma_{ij}x_j$ $(i=1,\ldots,n)$. Schematisch

$$\begin{pmatrix} \sigma_{11} & \cdots\cdots & \sigma_{1m} \\ \cdot & & \\ \boxed{} & & \\ \cdot & & \\ \cdot & & \\ \cdot & & \\ \sigma_{n1} & \cdots\cdots & \sigma_{nm} \end{pmatrix} \begin{pmatrix} x_1 \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ x_m \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \boxed{} \\ \cdot \\ \cdot \\ \cdot \\ y_n \end{pmatrix}$$

Door middel van de matrix-presentatie hebben we als het ware σ gerealiseerd tot een lineaire trafo van $\mathbb{R}^m$ in $\mathbb{R}^n$! (de bij de matrix σ_{ij} behorende afbeelding: zie blz. 80)

Beschouwen we nu i.p.v. A de numerieke m-dimensionale vectorruimte $\mathbb{R}^m$ waarin de vectoren als kolommen

$$\begin{pmatrix} x_1 \\ \cdot \\ \cdot \\ \cdot \\ x_m \end{pmatrix}$$

geschreven worden en voor B de n-dimensionale ruimte $\mathbb{R}^n$, dan kunnen we voor $a_1,\ldots,a_m$ de eenheidsvectoren

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{pmatrix}, \ldots, e_m = \begin{pmatrix} 0 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 1 \end{pmatrix}$$

nemen, en voor $b_1,\dots,b_m$ de eenheidsvectoren

$$e_1 = \begin{pmatrix} 1 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{pmatrix}, \dots, e_m = \begin{pmatrix} 0 \\ 0 \\ \cdot \\ \cdot \\ 1 \end{pmatrix} .$$

Als σ nu een lineaire transformatie is van $\mathbb{R}^m$ in $\mathbb{R}^n$ dan wordt deze weergegeven door een matrixpresentatie t.o.v. $e_1,\dots,e_m$ en $e_1,\dots,e_n$.

$$\sigma = \begin{pmatrix} \sigma_{11} & \cdots\cdots & \sigma_{1m} \\ \cdot & & \\ \cdot & & \\ \cdot & & \\ \sigma_{n1} & \cdots\cdots & \sigma_{nm} \end{pmatrix}$$

Het beeld onder σ van een vector $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ bepaald men door de formules

$$\sigma\begin{pmatrix} x_1 \\ \cdot \\ \cdot \\ x_m \end{pmatrix} = \begin{pmatrix} \sigma_{11} & \cdots\cdots & \sigma_{1m} \\ \cdot & & \\ \cdot & & \\ \sigma_{n1} & \cdots\cdots & \sigma_{nm} \end{pmatrix} \begin{pmatrix} x_1 \\ \cdot \\ \cdot \\ x_m \end{pmatrix} = \begin{pmatrix} \sigma_{11}x_1 + \sigma_{12}x_2 + \dots \sigma_{1m}x_m \\ \cdot \\ \cdot \\ \sigma_{n1}x_1 + \sigma_{n2}x_2 + \dots \sigma_{nm}x_m \end{pmatrix}$$

I.h.b. geldt

$$\sigma e_1 = \begin{pmatrix} \sigma_{11} \\ \cdot \\ \cdot \\ \cdot \\ \sigma_{n1} \end{pmatrix}, \dots, \sigma e_m = \begin{pmatrix} \sigma_{1m} \\ \cdot \\ \cdot \\ \cdot \\ \sigma_{nm} \end{pmatrix}$$

m.a.w. <u>de kolommen van de matrix zijn niets ander dan de beelden van de eenheidsvectoren van</u> $\mathbb{R}^m$.

Samenstelling en optelling van Homomorfismen, inverse homomorfisme

Zij $\sigma: L \to M$ en $\rho: M \to N$ homomorfismen van lineaire ruimten. Dan kan men de samenstelling $\rho\sigma$ van σ en ρ beschouwen het z.g. product van σ en ρ. Gemakkelijk gaat men na dat $\rho\sigma: A \to C$ weer een homomorfisme is. Net als bij afbeeldingen tussen verzamelingen hebben we de associatieve wet

$$(\tau\rho)\sigma = \tau(\rho\sigma) .$$

Als $\sigma: A \to B$ en $\rho: A \to B$ homomorfisme zijn dan is $\sigma + \rho: A \to B$ gedefinieerd door $(\sigma+\rho)(x) = \sigma x + \rho x$ weer een homomorfisme: de z.g. som der homomorfismen σ en ρ. Verder, als $\lambda \in \mathbb{R}$ en $\sigma: L \to M$ dan is $\lambda\sigma: L \to M$ gedefinieerd door $(\lambda\sigma)x = \lambda(\sigma x)$, ook weer een homomorfisme. Er gelden de distributieve wetten

$$\rho(\sigma+\tau) = \rho\sigma + \rho\tau$$

en

$$(\sigma+\tau)\rho = \sigma\rho + \tau\rho.$$

Een homomorfisme $\sigma: L \to M$ dat 1-1 en surjectief is heet een isomorfisme. In dat geval bestaat $\sigma^{-1}: M \to L$, de inverse van σ.
Ga na dat geldt σ is 1-1 d.e.s.d. $\sigma^{-1}(0) = 0$. Inderdaad, in dat geval is, wanneer $\sigma a = \sigma b$, dan is $\sigma(a-b) = 0 \Longrightarrow a-b = 0 \Longrightarrow a = b$. Het is gemakkelijk na te gaan dat ook σ^{-1} een homomorfisme is. Als σ en τ isomorfismen zijn $\sigma: L \to M$ en $\tau: M \to N$, dan geldt $(\tau\sigma)^{-1} = \sigma^{-1}\tau^{-1}$.

Stelling 3. Met de gedefinieerde optelling en scalaire vemenigvuldiging vormen de homomorfismen van L naar M zelf weer een lineaire ruimte (met deze vermenigvuldiging vormen zij zelfs een ring).

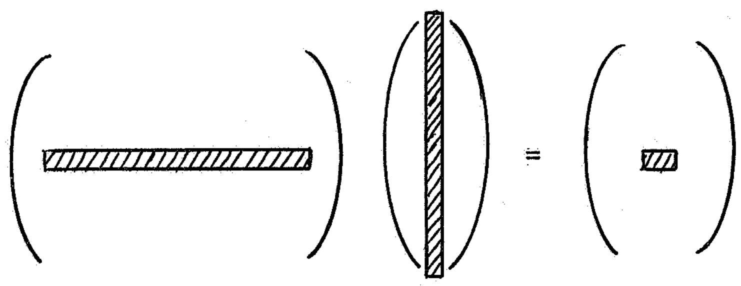
Zij nu A, B, C weer vectorruimten $\sigma: A \to B$, $\rho: B \to C$ homomorfismen; $a_1,\dots,a_m$; $b_1,\dots,b_n$; $c_1,\dots,c_p$ basis voor resp. A, B, C.
Laat $(\sigma_{ij})(i = 1,\dots,n;\ j = 1,\dots,m)$ de matrix zijn voor σ t.o.v. de a's en b's.
Laat $(\rho_{ki})(k = 1,\dots,p;\ i = 1,\dots,n)$ de matrix zijn voor ρ t.o.v. de b's en c's.

Dan geldt voor de matrix van $\rho\sigma$ t.o.v. de a's en de c's dat deze in de j'de kolom en de k'de rij het getal $(\rho\sigma)_{kj} = \sum_{i=1}^{n} \rho_{ki}\, \sigma_{ij}$ bezit.

Inderdaad, als $x = x_1a_1 + \ldots + x_ma_m$, dan voegt σ aan x toe de vector $y = y_1b_1 + \ldots + y_nb_n$ met $y_i = \sum_{j=1}^{m} \sigma_{ij}\, x_j$. σ is de transformatie die aan $y = y_1b_1 + \ldots + y_nb_n$ toevoegt $z_1c_1 + \ldots + z_pc_p$ uit C met $z_k = \sum_{i=1}^{n} \rho_{ki}\, y_i$.
De productafbeelding voegt dus aan $x_1a_1 + \ldots + x_ma_m$ toe $z_1c_1 + \ldots + z_pc_p$ en we hebben

$$z_k = \sum_{i=1}^{n} \rho_{ki} \sum_{j=1}^{m} \sigma_{ij}\, x_j = \sum_{j=1}^{m} x_j \sum \rho_{ki}\, \sigma_{ij} .$$

Schematisch



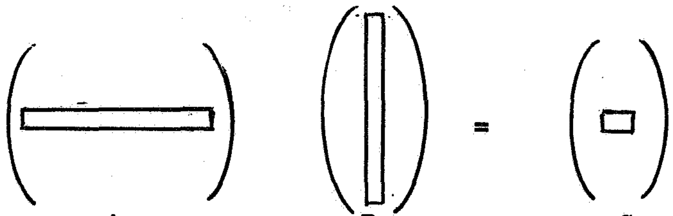
Onafhankelijk van de theorie van de lineaire transformatie kan men (n×m) matrices definiëren als zijnde een schema van mn getallen a_{ij} $1 \leq i \leq n$, $1 \leq j \leq m$ voorgesteld door

$$A = \begin{pmatrix} a_{11} & \ldots & a_{1m} \\ \cdot & & \\ \cdot & & \\ \cdot & & \\ a_{n1} & \ldots & a_{nm} \end{pmatrix}$$

Wanneer A een n×m matrix is $A = (a_{ij})$ $(1 \leq i \leq n;\ 1 \leq j \leq m)$ en B is een p×n matrix: $B = (b_{ki})$ $(1 \leq k \leq p;\ 1 \leq i \leq n)$ dan definieert men het <u>product van</u> B en A, schrijfwijze BA, als zijnde de matrix $C = (c_{kj})$ $(1 \leq k \leq p;\ 1 \leq j \leq m)$ door

$$\boxed{c_{kj} = \sum_{i=1}^{n} b_{ki}\, a_{ij}}$$

Schematisch



A B C

Nu volgt onmiddellijk: (zie blz. 79) de volgende bewering: Als σ: $A \to B$ en ρ: $B \to C$ lineaire transformaties zijn en $A = (a_{ij})$ is de matrixpresentatie van ρ t.o.v. de bases $a_1,\dots,a_m$ en $b_1,\dots,b_n$ en $B = (b_{ij})$ is de matrixpresentatie van ρ t.o.v. de bases $b_1,\dots,b_n$ en $c_1,\dots,c_p$ dan is BA de matrixpresentatie van de samengestelde afbeelding $\rho\circ\sigma$.

Stel nu $A = (a_{ij})$ $(i = 1,\dots,n;\ j = 1,\dots,n)$ een $n\times m$ matrix. Men kan A nu ook zien als een lineaire trafo $A = \mathbb{R}^m \to \mathbb{R}^n$ door $A(x) = y$ met

$$\begin{cases} y_1 = a_{11}x_1 + \dots + a_{1m}x_m \\ \quad\cdot \\ \quad\cdot \\ \quad\cdot \\ y_n = a_{n1}x_1 + \dots + a_{nm}x_m \end{cases}$$

(zie ook blz. 72)

Deze trafo A: $\mathbb{R}^m \to \mathbb{R}^n$ die bij de matrix geassocieerd is heet de bij de matrix A behorende transformatie. Merk op dat het beeld van een vector x (kolom vector) $\in \mathbb{R}^m$ onder de trafo A niets anders is dan het matrix-product van A en de matrix x.

Ook geldt: 1) het beeld van de i^{de} eenheidsvector e_i van $\mathbb{R}^m$ onder de bij A behorende lineaire transformatie A: $\mathbb{R}^m \to \mathbb{R}^n$ is precies de i^{de} kolom van de matrix A.

2) De matrix A is precies de matrixpresentatie van zijn bijbehorende transformatie A: $\mathbb{R}^m \to \mathbb{R}^n$ t.o.v. de basis $e_1,\dots,e_m$ en $e_1,\dots,e_n$ van resp. $\mathbb{R}^m$ en $\mathbb{R}^n$.

Men definieert de som A + B van A en B indien p = m door A + B = C met $c_{ij} = a_{ij} + b_{ij}$ $(1 \le i \le n;\ 1 \le j \le m)$. Analoog kunnen we een scalaire vermenigvuldiging definiëren. Als A een $n\times m$ matrix is dan definieert men λA door $C = \lambda A$ met $c_{ij} = \lambda a_{ij}$.

Men verifieert gemakkelijk dat definitie van optelling en scalaire vermenigvuldiging compatibel zijn met de optelling en scalaire vermenigvuldiging van de bijbehorende lineaire transformaties.

De optelling en vermenigvuldiging met een scalair voldoen aan de volgende eigenschappen:

1) $A + B = B + A$

$(A + B) + C = A + (B + C)$

$\lambda(A + B) = \lambda A + \lambda B \qquad (\lambda+\mu)A = \lambda A + \mu A$

$(\lambda\mu)A = \lambda(\mu A)$

2) $(AB)C = A(BC)$

$A(B + C) = AB + AC$

$(A + B)C = AC + BC$

$A(\lambda B) = \lambda(AB) = (\lambda A)B$

De $n\times n$ matrix I die op de diagonaal allemaal énen en voor de rest nullen bezit, heet de eenheidsmatrix en de bijbehorende transformatie $I:\mathbb{R}^n \to \mathbb{R}^n$ is de identieke afbeelding.

Als A een willekeurige $n\times n$ matrix is dan geldt dus $AI = IA = A$. M.a.w. I treedt op als eenheidselement t.a.v. de vermenigvuldiging van matrices.

Als A een $n\times n$ matrix is en er bestaat A^* zodat $AA^* = A^*A = I$ dan is A^* eenduidig bepaald en A heet in dit geval regulier. I.p.v. A^* schrijven we A^{-1}. Gemakkelijk gaan we na dat geldt:

A is regulier $\Longleftrightarrow$ de bij A behorende transformatie $A: \mathbb{R}^n \to \mathbb{R}^n$ is eenduidig (en dus een isomorfisme).

Rang van een matrix

Zij A een $n\times m$ matrix. De kolomvectoren spannen een r-dimensionale deelruimte op van $\mathbb{R}^n$. r heet de rang van de matrix A.

Nu volgt: de rang van een matrix A is gelijk aan de rang van de bij A behorende transformatie $\mathbb{R}^n$.

De volgende eigenschappen gelden:

1) De transformatie $A: \mathbb{R}^m \to \mathbb{R}^n$ is eeneenduidig $\Longleftrightarrow$ rang $A = m$.
2) De transformatie $A: \mathbb{R}^m \to \mathbb{R}^n$ is surjectief $\Longleftrightarrow$ rang $A = n$.

Uit het bovenstaande volgt dus:

Een $n \times n$ matrix A is regulier d.e.s.d. wanneer rang $A = n$.

Zonder bewijs vermelden we nog

Stelling 4. De dimensie van de kolommenruimte van een matrix is gelijk aan de dimensie van de rijenruimte van de matrix.

Lineaire vergelijkingen

Het stelsel

$$(1) \qquad \begin{cases} a_{11}x_1 + \dots + a_{1m}x_m = 0 \\ a_{21}x_1 + \dots + a_{2m}x_m = 0 \\ \cdot \\ \cdot \\ \cdot \\ a_{n1}x_1 + \dots + a_{nm}x_m = 0 \end{cases}$$

met bekenden a_{ij} en onbekenden x_j $(1 \le i \le n,\ 1 \le j \le m)$ heet een stelsel van n homogene lineaire vergelijkingen met m onbekenden. Men vraagt de oplossingen. Dit vraagstuk kan ook als volgt geformuleerd worden. Men vraagt de kern van de bij de matrix $A = (a_{ij})$ behorende transformatie $A: \mathbb{R}^m \to \mathbb{R}^n$.

Het stelsel kunnen we dus samenvatten in de vergelijking

$$A(x) = 0 \quad \text{en ook} \quad Ax = 0$$

($Ax = 0$ betekent dan dat het matrixproduct van de matrix A en de matrix x gelijk is aan nul).

Als rang $A = r$ dan is $\dim \ker A = m - r$; dus:
De oplossingen van (1) vormen een $m-r$ dimensionale deelruimte van $\mathbb{R}^m$.
Is $r = m$ dan is er precies één oplossing n.l. $x_1 = x_2 = \dots = x_m = 0$.

Het stelsel

$$(2)\qquad \begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1m}x_m = b_1 \\ \quad\vdots \\ a_{n1}x_1 + \dots\dots\dots + a_{nm}x_m = b_n \end{cases}$$

kan nu samengevat worden tot $Ax = b$ ($A(x) = b$) wanneer we afspreken dat $b = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$.

(2) heet een stelsel van n niet-homogene lineaire vergelijkingen in m-onbekenden. (1) heet het daarbij behorende homogene stelsel. Men onderscheidt twee gevallen

I De vector $b = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$ ligt in $A(\mathbb{R}^m)$. Er is dus een oplossing. De rangen van de volgende matrices zijn gelijk, omdat de stelsels kolommen dezelfde ruimte van dimensie r opspannen

$$A = \begin{bmatrix} a_{11} & \dots & a_{1m} \\ \vdots & & \\ a_{m1} & \dots & a_{nm} \end{bmatrix} \quad \text{en} \quad \begin{bmatrix} a_{11} & \dots & a_{1m}, b_1 \\ \vdots & & \\ a_{n1} & \dots & a_{nm}, b_m \end{bmatrix}$$

Men noemt de laatste matrix ook wel de aangevulde matrix.

Is y een oplossing van (2) dus $A(y) = b$; x een oplossing van (1) dus $A(x) = 0$ dan is $y + x$ een oplossing van (2), immers $A(y+x) = A(y) + A(x) = b + 0 = b$. Anders gezegd is y een oplossing van (2) dan vormen alle oplossingen van (2) de verzameling $y + A^{-1}(0)$. Dat dit inderdaad alle oplossingen zijn volgt gemakkelijk. Immers zij ook y' een oplossing van (2), dus $A(y') = b$, dan volgt $A(y-y') = A(y) - A(y') = b - b = 0$ dus $y - y' \in A^{-1}(0)$ dus $y' = y + (y'-y)$ waarbij $y'-y \in A^{-1}(0)$.

II De kolommen van de aangevulde matrix spannen een ruimte op, die de ruimte opgespannen door de kolommen van matrix A echt omvat, d.w.z. bevat en er niet gelijk aan is. De rangen van deze matrix zijn dan verschillend. Het stelsel vergelijkingen heet dan strijdig.

Samengevat

Stelling 5. De vergelijkingen van (2) hebben geen oplossing (strijdig stelsel) indien de rang r van de matrix A ongelijk is aan de rang van de aangevulde matrix. Zijn deze rangen wel gelijk dan zijn de oplossingen in een eeneenduidig verband te brengen met de oplossingen van het bijbehorende homogene stelsel (1) die een m-r dimensionale deelruimte van $\mathbb{R}^m$ vormen.

Stelling 6. Is het aantal vergelijkingen gelijk aan het aantal onbekenden en is ook de rang van de matrix A hieraan gelijk, n = m = r, dan heeft (2) precies één oplossing.

§3. Determinant van een endomorfisme

Multilineaire functie

Definitie: Een functie $\phi(x_1,\dots,x_r)$ van r variabelen die vectoren in A zijn,

$$\phi: A \times \dots \times A \to \mathbb{R}$$

met waarden in $\mathbb{R}$, heet multilineair (r-lineair) indien $r \geq 1$ en indien elk van de functies

$$x_i \to \phi(x_1,\dots,x_r) \qquad x_j = \text{constant voor } j \neq i$$

(homogeen) lineair is. $x_1,\dots$ en x_r heten ook wel de argumenten van de functie. De functie heet antisymmetrisch indien steeds

$$\phi(x_1,\dots,x_r) = -\phi(y_1,\dots,y_r)$$

geldt indien de rij vectoren $y_1,\dots,y_r$ uit de rij $x_1,\dots,x_r$ kan ontstaan door verwisseling van twee der vectoren.

Een voorbeeld van een multilineaire antisymmetrische functie verkrijgt men als volgt. Kies drie vectoren a, b en c met beginpunt 0 in de gewone driedimensionale ruimte $\mathbb{R}^3$. Noem de inhoud van het parallelopipedum dat a, b en c tot ribben heeft $I(a,b,c)$. $I = 0$ indien a, b en c afhankelijk zijn. Stel $\phi(a,b,c) = I(a,b,c)$ indien de draaiïng van a naar b over de kleinste hoek vanuit het uiteinde van c gezien overeenstemt met de draaiïng van de wijzers van de klok. Stel $\phi(a,b,c) = -I(a,b,c)$ wanneer dit niet het geval is. Men gaat gemakkelijk na dat ϕ multilineair en antisymmetrisch is!

Stelling 1. Zij $\phi(x_1,\dots,x_r)$ een multilineaire functie met argumenten (dat zijn de x_i's) in de vectorruimte A. Dan zijn de volgende voorwaarden equivalent:

1) ϕ is antisymmetrisch.
2) $\phi(x_1,\dots,x_r) = 0$ indien twee der argumenten gelijk zijn.
3) $\phi(x_1,\dots,x_r) = 0$ indien de vectoren $x_1,\dots,x_r$ lineair afhankelijk zijn.

Bewijs

Uit 1) zowel als uit 3) volgt 2) als speciaal geval. Geldt 2) dan is wegens de lineariteit, indien we uitsluitend het p'de en q'de argument noteren, en de andere argumenten constant houden: $\phi(a,b) = -\phi(b,a)$. Daarmee staat dus reeds vast dat 1) en 2) gelijkwaardig zijn we tonen tenslotte nog aan dat 3) uit 2) volgt. Stel ϕ heeft de eigenschap 2) en $x_1,\ldots,x_r$ zijn lineair afhankelijk. Dan bestaat $\lambda_2,\ldots,\lambda_r \in R$ zodat na eventuele vernummering

$$x_1 = \lambda_2 x_2 + \ldots + \lambda_r x_r \qquad \lambda_i \in R.$$

Dan is

$$\phi(x_1,x_2,\ldots,x_r) = \phi(\lambda_2 x_2 + \ldots + \lambda_r x_r, x_2,\ldots,x_r) =$$

$$= \lambda_2 \phi(x_2,x_2,\ldots,x_r) + \ldots \lambda_r \phi(x_r,x_2,\ldots,x_r) = 0.$$

In het volgende hebben we nog het begrip permutatie nodig. Een eeneenduidige afbeelding van een eindige verzameling op zichzelf heet een permutatie. Een permutatie "toepassen" betekent elk element op de plaats van zijn beeld zetten. Permutaties kunnen worden samengesteld (samenstellen van afbeeldingen: zie vroeger). Elke permutatie van n elementen $a_1,\ldots,a_n$ kan worden samengesteld door achtereenvolgens een aantal verwisselingen telkens van twee elementen uit te voeren, die b.v. successievelijk $a_1,a_2,\ldots,a_{n-1}$ op de uiteindelijke gewenste plaats brengen.

Stelling 2. Is σ een permutatie van een eindige verzameling en zijn τ_i en τ'_j, $i = 1,\ldots,m$; $j = 1,\ldots,m'$, verwisselingen van telkens twee elementen zodat

$$\sigma = \tau_m . \tau_{m-1} \ldots \tau_1 = \tau'_{m'} . \tau'_{m'-1} \ldots \tau'_1$$

dan zijn m en m' beide even of beide oneven, of gelijkwaardig daarmee is $(-1)^m = (-1)^{m'}$.

Door deze stelling is de volgende definitie mogelijk:
Een permutatie heet <u>even</u> resp. <u>oneven</u> indien hij door samenstelling uit een even resp. oneven aantal verwisselingen van telkens twee elementen kan ontstaan. Dan is het getal sign $\sigma = (-1)^m = (-1)^{m'}$ gelijk 1 of -1. sign betekent de <u>signatuur</u> van de permutatie.

<u>Bewijs van de stelling</u>. We nummeren de elementen van de eindige verzameling, en stellen ze tevens voor door de natuurlijke getallen $1,2,\ldots,n$. σ zij een permutatie en $\sigma^{-1}(i) = j_i$ of wel $\sigma(j_i) = i$. Toepassing van σ geeft dus op de plaats van i het getal j_i; op de plaatsen van $1,2,\ldots,n$ de getallen $j_1,j_2,\ldots,j_n$. We geven deze permutatie σ aan door de rij die ontstaat indien men de permutatie toepast op de rij $(1,2,\ldots,n)$ nl. $\sigma^{-1}(1),\sigma^{-1}(2),\ldots,\sigma^{-1}(n) = (j_1,j_2,\ldots,j_n)$.
We noemen een paar getallen j_p,j_q met $p < q$ een onordelijk paar of inversie van de permutatie σ van $1,\ldots,n$, dan en slechts dan wanneer $j_p > j_q$. Het aantal inversies zij $N(j_1,\ldots,j_n)$. Bij een verwisseling van twee naast elkaar staande indices in de rij $j_1,\ldots,j_n$ neemt $N(j_1,\ldots,j_n)$ met één toe of af. Een verwisseling van twee willekeurige indices bijv. de p^{de} of de q^{de}, kan worden teweeggebracht door een oneven aantal n.l. $2(q-p)-1$ dergelijke buurtverwisselingen. Bij een verwisseling van twee indices in de rij $(j_1,\ldots,j_n)$ keert de functiewaarde van $(-1)^{N(j_1,\ldots,j_n)}$ dus van teken om. Is m een aantal verwisselingen die samengesteld de permutatie $\sigma = (j_1,\ldots,j_n)$ leveren, dan is dus $(-1)^{N(j_1,\ldots,j_n)} = (-1)^m$. Het linkerlid hiervan is bij een gegeven permutatie σ van $1,2,\ldots,n$ onafhankelijk van het aantal m. Daaruit volgt de stelling $(-1)^m = (-1)^{m'}$. Het rechterlid is onafhankelijk van de nummering (= naamgeving der elementen). Daaruit volgt dat het getal sign $\sigma = \text{sign}(j_1,\ldots,j_n) = (-1)^{N(j_1,\ldots,j_m)} = (-1)^m$ geheel bepaald is door σ; sign σ is dus een functie van σ.

<u>Stelling 3</u>. Is A een n-dimensionale vectorruimte, dan bestaat precies één n-lineaire antisymmetrische functie $\phi(x_1,\ldots,x_n)$ met argumenten $x_i \in A$ die op de n-gerangschikte vectoren $a_1,\ldots,a_n$ van een gegeven basis van A de gegeven waarde $\lambda \in R$ aanneemt.

We bewijzen eerst dat ϕ eenduidig bepaald is bij gegevan $a_1,\ldots,a_n$ en λ.

Stel $x_1,\ldots,x_n \in A$ en $x_i = \sum_{k=1}^{n} \sigma_{ki}\, a_k$.

Dan is wegens de multilineariteit van ϕ:

$$\phi(x_1,\ldots,x_n) = \phi(\sum \sigma_{j_1}\, a_{j_1},\ldots \sum \sigma_{j_n n}\, a_{j_n}) =$$

$$= \sum \sigma_{j_1 1}.\sigma_{j_2 2},\ldots\, \sigma_{j_n n}\, \phi(a_{j_1},\ldots,a_{j_n}) \;.$$

De sommatie is over $j_1,\ldots,j_n$ die elk van 1 tot n lopen. $\phi(a_{j_1},\ldots,a_{j_n}) = 0$ wanneer twee der argumenten gelijk zijn dus wanneer twee der indices gelijk zijn. Is dit niet zo dan ontstaat $j_1,\ldots,j_n$ door een permutatie uit $1,\ldots,n$ die samengesteld is uit een m-tal verwisselingen van twee elementen. Dan is

$$\phi(a_{j_1},\ldots,a_{j_n}) = \mathrm{sign}(j_1,\ldots,j_n).\phi(a_1,\ldots,a_n) =$$

$$= \mathrm{sign}(j_1,\ldots,j_n).\lambda \;.$$

Dus volgt dat de functiewaarde van $\phi(x_1,\ldots,x_n)$ volledig bepaald is door de formule

$$(*) \qquad \phi(x_1,\ldots,x_n) = \lambda \sum \mathrm{sign}(j_1,\ldots,j_n)\, \sigma_{j_1 1}.\sigma_{j_2 2}\ldots\, \sigma_{j_n n} \;.$$

waarbij $j_1,\ldots,j_n$ alle permutaties toegepast op $1,\ldots,n$ doorloopt.

Om nu omgekeerd de existentie te bewijzen moeten we aantonen dat bovenstaande formule ook inderdaad altijd een multilineaire antisymmetrische functie voorstelt. De multilineariteit is duidelijk. Om te bewijzen dat (*) antisymmetrisch is stellen we twee der argumenten gelijk van $\phi(x_1,\ldots,x_n)$, b.v. $x_p = x_q$. Dan is ook $\sigma_{kp} = \sigma_{kq}$ voor $k = 1,\ldots,n$.

Nu is $\phi(x_1,\ldots,x_n) = \lambda \sum \sigma_{j_1 1}\, \sigma_{j_2 2} \ldots \sigma_{j_n n}\, \mathrm{sign}(j_1,\ldots,j_n)$.

Voor elke $k \neq l$ zodat $1 \leq k,l \leq n$ bevat deze som voor een keuze van j_i $(i \neq p,q)$ precies twee termen van de vorm

$$\sigma_{j_1 1}.\sigma_{j_2 2} \ldots \sigma_{kp} \ldots \sigma_{lq} \ldots \sigma_{j_n n}.\, \mathrm{sign}(j_1,\ldots,k,\ldots,l,\ldots,j_n)$$

$$\sigma_{j_1 1}.\sigma_{j_2 2} \ldots \sigma_{lp} \ldots \sigma_{kq} \ldots \sigma_{j_n n}.\, \mathrm{sign}(j_1,\ldots,l,\ldots,k,\ldots,j_n)$$

die omdat $\sigma_{kp} = \sigma_{kq}$ en $\sigma_{lp} = \sigma_{lq}$ elkaars tegengestelde zijn. De totale som bestaat dus uit termen die twee bij twee elkaars tegengestelde zijn. Dus $\phi(x_1,\ldots,x_n) = 0$. Hiermede is de stelling bewezen.

Voor het geval n = 2 uitgeschreven

$$\phi(x_1,x_2) = (\sigma_{11}\,\sigma_{22} - \sigma_{21}\,\sigma_{12}).\lambda$$

Als $\phi \not\equiv 0$ dan geldt dus blijkbaar $\phi(x_1,\ldots,x_n) = 0 \iff (x_1,\ldots,x_n)$ zijn lineair afhankelijk. Immers merk op dat ϕ identiek nul wordt zodra $\phi = 0$ voor een basis $a_1,\ldots,a_n$ van A. Verder merken we op dat de coëfficiënt van $\phi(a_1,\ldots,a_n)$ niet van ϕ afhangt. Zijn ϕ en ψ dus twee n-lineaire antisymmetrische functies niet identiek nul, dan is voor twee bases $(y_1,\ldots,y_n)$ en $(z_1,\ldots,z_n)$ steeds

$$\frac{\psi(y_1,\ldots,y_n)}{\phi(y_1,\ldots,y_n)} = \frac{\psi(z_1,\ldots,z_n)}{\phi(z_1,\ldots,z_n)} \tag{1}$$

m.a.w.: de verhouding van twee n-lineaire antisymmetrische functies op een n-dimensionale vectorruimte is een constante.

Zij nu ϕ een niet triviale antisymmetrische functie en σ een endomorfisme σ: $A \to A$. De functie $\phi(\sigma y_1,\ldots,\sigma y_n)$ is ook lineair in de argumenten $y_1,\ldots,y_n$ en neemt de waarde nul aan als twee der argumenten gelijk zijn. Het is dus ook een antisymmetrische n-lineaire functie in $y_1,\ldots,y_n$ en slechts identiek nul indien de rang van σ < n is. Het getal

$$\frac{\phi(\sigma y_1,\ldots,\, y_n)}{\phi(y_1,\ldots,y_n)} \tag{2}$$

is nu afhankelijk van de basis $y_1,\ldots,y_n$ zowel voor geval rang $\sigma < n$ als rang $\sigma = n$. Volgens (1) verandert de waarde van (2) ook niet als ϕ door een andere lineaire functie ψ vervangen wordt.

(2) heet de determinant van het endomorfisme σ, det σ of ook wel $|\sigma|$.

Blijkbaar kunnen we det σ in formule schrijven als we de matrix (σ_{ij}) van σ t.o.v. een basis $y_1,\ldots,y_n$ van A bepalen door $\sigma y_j = x_j = \sum \sigma_{ij}\, y_i$.

Dan volgt uit het bewijs van de vorige stelling

$$\boxed{\det \sigma = \sum \sigma_{j_1 1}.\sigma_{i_2 2}.\ \ldots\ \sigma_{i_n n}.\ \mathrm{sign}(i_1,\ldots,i_n)}$$

Stelling 4. Zijn σ en τ endomorfismen $A \to A$ dan is $\det(\sigma\tau) = \det \sigma.\ \det \tau$. det identiek = 1 ; $\det(\sigma^{-1}) = (\det \sigma)^{-1}$.

Bewijs. Is de rang van σ of τ kleiner dan n dan is ook de rang van $\sigma\tau$ kleiner dan n en $\det \sigma\tau = \det \sigma.\ \det \tau = 0$. Anders is bij een basis $y_1,\ldots,y_n \in A$ een niet triviale n-lineaire antisymmetrische functie ϕ, volgens het voorgaande:

$$0 \neq \phi(\sigma\tau y_1,\ldots,\sigma\tau y_n) = \det \sigma\tau.\phi(y_1,\ldots,y_n) = (\text{anderzijds}) =$$

$$\det \sigma.\phi(\tau y_1,\ldots,\tau y_n) = \det \sigma.\ \det \tau.\ \phi(y_1,\ldots,y_n)$$

waaruit het eerst gestelde volgt. De andere beweringen zijn nu triviaal.

Determinant van een matrix.

Zij $A = (a_{ij})$, $i,j = 1,\ldots,n$, een n×n matrix

$$A = \begin{bmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \\ a_{n1} & \ldots & a_{nn} \end{bmatrix}$$

Dan stelt deze matrix een lineaire transformatie A voor A: $R^n \to R^n$. (zie blz. 80). Onder de determinant van de matrix A verstaat men nu per definitie det A. Notatie

$$\begin{vmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \\ a_{n1} & \ldots & a_{nn} \end{vmatrix}$$

Blijkbaar $\boxed{\det A = \sum a_{i_1 1}.a_{i_2 2} \dots a_{i_n n}.\ \mathrm{sign}(i_1,i_2,\dots,i_n)}$.

Gevolg. De determinant van een n×n matrix A is de functiewaarde op de rij kolomvectoren van A, van die n-lineaire antisymmetrische functie op R^n welke aan de rij kolommen van de eenheidsmatrix de waarde 1 toevoegt.

Blijkbaar geldt nu voor de getransponeerde matrix, d.i. de matrix A^T die uit A ontstaat door verwisseling van de rijen en kolommen (dus $(A^T)_{ij} = a_{ji}$, $i,j = 1,\dots,n$), dat

$$\det A^T = \sum a_{1j_1}.a_{2j_2}\dots a_{ni_n}\ \mathrm{sign}(j_1,\dots,j_n)$$

$$= \sum a_{i_1 1}.a_{i_2 2}\dots a_{i_n n}.\ \mathrm{sign}(i_1,\dots,i_n) = \det A$$

(Ga dit na).
det A = det A^T is ook een multilineaire antisymmetrische functie van zijn kolomvectoren = rijen van A.

Nu hebben we dus de volgende zeer algemene stelling.

Stelling 5. De determinant van een matrix is nul dan en slechts dan, wanneer de kolommen (of rijen) lineair afhankelijk zijn. Verwisselt men in een matrix twee kolommen (rijen) dan keert de determinant van teken om. Telt men bij een kolom (rij) van een matrix een veelvoud van een andere kolom (rij) op dan verandert de determinant niet van waarde. De determinant van een n×n matrix is een antisymmetrische lineaire functie van de rijen (idem van de kolommen).

Stelling 6. Als A en B n×n matrices zijn, dan geldt

$$\det AB = \det A.\ \det B\ .$$

Bewijs. Dit volgt direct uit Stelling 4.

Stelling 7. Zij j een der getallen $1,\ldots,n$. Is elk element van de j^{de} kolom van een vierkante matrix

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$$

als som van twee termen geschreven: $a_{ij} = b_{ij} + c_{ij}$ $(i = 1,\ldots,n)$ dan is det A = det B + det C. Hierin is B de matrix die uit A ontstaat door de j^{de} kolom

$$\begin{pmatrix} a_{1j} \\ \vdots \\ a_{nj} \end{pmatrix} \quad \text{te vervangen door} \quad \begin{pmatrix} b_{1j} \\ \vdots \\ b_{nj} \end{pmatrix} ,$$

terwijl C analoog is gedefinieerd.

Een analoge redenering geldt voor de rijen van de matrix A.

De ontwikkeling van een determinant naar een kolom.

Laat men uit de matrix

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$$

de i^{de} rij en de j^{de} kolom weg, dan ontstaat er weer een vierkante matrix die we met A_{ij} aanduiden. De determinant van A_{ij} heet de onderdeterminant van a_{ij}. Het getal $(-1)^{i+j}$ det $A_{ij} = m_{ij}$ wordt de minor van a_{ij} genoemd.

Stelling 8. Is j een der getallen $1,\ldots,n$ dan is

$$\det A = a_{1j}\, m_{1j} + a_{2j}\, m_{2j} + \ldots + a_{nj}\, m_{nj}$$

(ontwikkeling van de determinant van een matrix naar de elementen van een kolom). De overeenkomstige stelling geldt voor ontwikkeling naar een rij.

Bewijs. 1. Zij B een n×n matrix met $b_{in} = 0$ voor $j = 1,\dots,n-1$. Dan zijn in de som, waarin det B volgens de formule op blz. 91 geschreven kan worden, alleen die termen mogelijk ongelijk nul, die de gedaante

$$\mathrm{sign}(i_1,\dots,i_n)\ b_{i_1 1} \cdots b_{i_n n} \quad \text{met } i_n = n$$

hebben, waarbij $i_1,\dots,i_{n-1}$ een permutatie is van $1,\dots,n-1$. Daar $\mathrm{sign}(i_1,\dots,i_{n-1},n) = \mathrm{sign}(i_1,\dots,i_{n-1})$, stellen we nu vast, dat

$$\det B = b_{nn} \det B_{nn}$$

waarbij $\det B_{nn}$ de onderdeterminant is van b_{nn}.

2. We keren nu terug tot onze matrix A. Volgens de vorige stelling is

$$\det A = \det P_1 + \dots + \det P_n$$

waarbij de matrix P_i ontstaat door in de j^{de} kolom van A (voor de ene bepaalde j van de stelling) alle elementen a_{kj} met $k \neq i$ door 0 te vervangen.

Door in ieder der P_i's de i-de rij n-i plaatsen naar beneden te schuiven en de j^{de} kolom n-j plaatsen naar rechts komen we tot

$$\det P_i = (-1)^{(n-i)+(n-j)} \det \begin{bmatrix} & 0 \\ & 0 \\ A_{ij} & 0 \\ & \vdots \\ \cdots\cdots & a_{ij} \end{bmatrix}$$

Passen we nu 1 toe dan zien we dat

$$\det P_i = (-1)^{i+j}\, a_{ij} \det A_{ij} = a_{ij}\, m_{ij}$$

waaruit het gestelde volgt.

Toepassing. Berekening van een determinant:

$$\begin{vmatrix} 1 & 2 & 8 & 4 \\ 0 & 2 & 1 & 3 \\ 3 & 6 & 0 & 6 \\ 4 & 1 & 0 & 2 \end{vmatrix} = \begin{vmatrix} -7 & -14 & 0 & -20 \\ 0 & 2 & 1 & 3 \\ 3 & 6 & 0 & 6 \\ 4 & 1 & 0 & 2 \end{vmatrix}$$

(de eerste rij vervangen door er 8× tweede van af te trekken)

$$= \text{(ontwikkeling naar derde kolom)} = \begin{vmatrix} -7 & -14 & -20 \\ 3 & 6 & 6 \\ 4 & 1 & 2 \end{vmatrix} =$$

$$= 2 \begin{vmatrix} -7 & -14 & -10 \\ 3 & 6 & 3 \\ 4 & 1 & 1 \end{vmatrix} = 2.3 \begin{vmatrix} -7 & -14 & -10 \\ 1 & 2 & 1 \\ 4 & 1 & 1 \end{vmatrix} = 6 \begin{vmatrix} -7 & 0 & -10 \\ 1 & 0 & 1 \\ 4 & -7 & 1 \end{vmatrix}$$

(2× de eerste kolom van de tweede kolom afgetrokken)

$$= \text{(ontw. naar tweede kolom)} = 6.7 \begin{vmatrix} -7 & -10 \\ 1 & 1 \end{vmatrix} = -42 \begin{vmatrix} 7 & 10 \\ 1 & 1 \end{vmatrix} = -42.-3 = 126.$$

Opgave: Bewijs dat

$$\begin{vmatrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & \dots & x_n \\ x_1^2 & x_2^2 & \dots & x_n^2 \\ \vdots & \vdots & & \vdots \\ x_1^{n-1} & x_2^{n-1} & \dots & x_n^{n-1} \end{vmatrix} = (x_2-x_1)\,(x_3-x_2)\,(x_3-x_1)\,(x_4-x_3)\,(x_4-x_2)\,\dots\,(x_n-x_1)$$

Deze determinant staat bekend als de determinant van Van der Monde.

Toepassing op de theorie van de lineaire vergelijkingen, regel van Cramer.

I. Beschouw het homogene stelsel vergelijkingen

$$\begin{array}{l} a_{11}\, x_1 + \ldots + a_{1n}\, x_n = 0 \\ \vdots \\ a_{n1}\, x_1 + \ldots + a_{nn}\, x_n = 0 \end{array} \qquad \text{en stel} \quad A = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \\ a_{n1} & \ldots & a_{nn} \end{pmatrix}$$

We hebben vroeger gezien dat dit stelsel een oplossing $\neq 0$ bezit d.e.s.d. rang $A < n$.
Omdat rang $A < n \Longleftrightarrow$ de kolomvectoren zijn lineair afhankelijk volgt det $A = 0 \Longleftrightarrow$ rang $A < n$ (stelling 5).

Dus: het stelsel heeft dan en slechts dan een oplossing $\neq 0$ indien det $A = 0$.

II. Gegeven het niet homogene stelsel lineaire vergelijkingen

$$\begin{array}{l} a_{11}\, x_1 + \ldots + a_{1n}\, x_n = b_1 \\ \vdots \\ a_{n1}\, x_1 + \ldots + a_{nn}\, x_n = b_n \end{array} \qquad \text{en stel} \quad A = \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \\ a_{n1} & \ldots & a_{nn} \end{pmatrix}$$

en det $A \neq 0$. In dat geval bestaat er een éénduidige oplossing (zie vorige paragraaf).
Nu volgt als deze oplossing $x = (x_1, \ldots, x_n)$ is:

$$\begin{array}{l} a_{11}\, x_1 + \ldots (a_{1i}\, x_i - b_1) + \ldots + a_{1n}\, x_1 = 0 \\ \vdots \\ a_{n1}\, x_1 + \ldots (a_{n1}\, x_i - b_n) + \ldots + a_{nn}\, x_n = 0 \end{array}$$

De kolommen van de volgende matrix zijn dus lineair afhankelijk

$$\begin{bmatrix} a_{11} & \dots & a_{1i}\, x_i - b_1 & \dots & a_{1n} \\ \cdot & & & & \\ \cdot & & & & \\ \cdot & & & & \\ a_{n1} & \dots & a_{ni}\, x_i - b_n & \dots & a_{nn} \end{bmatrix}$$

Dan is de determinant van deze matrix nul.

Indien we de determinant van de matrix, die uit A ontstaat door de i^{de} kolom te vervangen door de kolom der getallen b_j, $j = 1,\dots,n$, voorstellen door D_i en det A = D stellen, dan volgt

$$D.x_i - D_i = 0 \implies x_i = \frac{D_i}{D} \ .$$

Deze vroeger veel gebruikte formule voor de oplossing is de <u>regel van Cramer</u>.

Determinant als inhoud

Inhoud van parallelopipedum en simplex

Onder een maat, soms genaamd oppervlakte of inhoud op een puntverzameling V verstaat men een functie I(G) die aan elke (G) van een collectie deelverzamelingen van V een niet negatief getal I(G) toevoegt, en met de eigenschap

$$G_1 \cap G_2 = 0 \quad \text{dan is} \quad I(G_1 \cup G_2) = I(G_1) + I(G_2) \ .$$

Het zou ons veel te ver voeren om maattheorie, of een onderdeel van deze theorie, de integraalrekening, te bespreken. We geven alleen formules voor de inhoud van parallelopipedum en simplex. De lezer beschouwe eerst zelf het geval n = 2, n = 3.

Onder een <u>parallelopipedum</u> (par.) in $\mathbb{R}^n$ verstaat men de verzameling vectoren x van $\mathbb{R}^n$ met

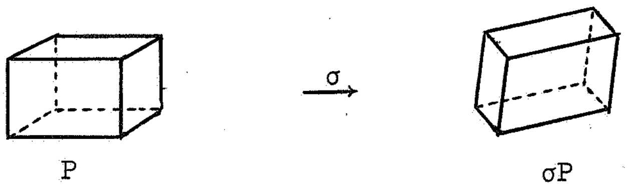
$$P\colon x = u + \sum_{i=1}^{n} \lambda_i b_i \qquad 0 < \lambda_i < 1 \ ; \ u, b_i \in \mathbb{R}^n$$

(i.h.b. n = 2 of 3, n = 2, P is een parallelogram).

Onder de inhoud van het par. verstaat men per definitie de absolute waarde van de determinant van de matrix die als kolommen $b_1,\ldots,b_n$ heeft. (vectoren in $\mathbb{R}^n$ als kolommen geschreven).

Ga na dat voor n = 2,3 de gewone definitie van inhoud hiermee samenvalt.

Als σ nu een lineaire transformatie is van $\mathbb{R}^n$ naar $\mathbb{R}^n$ en P is een parallelopipedum met inhoud I(P), dan is de inhoud van het par. σP precies $I(\sigma P) = |\det\sigma|.I(P)$. Hieruit zien we dus dat de determinant van σ in feite meet hoe de inhoud van een blok onder invloed van σ verandert.



Een simplex (driehoek, viervlak voor n = 2 resp. 3) bestaat uit een verzameling punten

$$S: x = u + \sum_{i=1}^{n} \lambda_i b_i \qquad 0 < \lambda_i < 1 \;,\; \sum \lambda_i < 1 \;;\; u,b_i \in R^n \;.$$

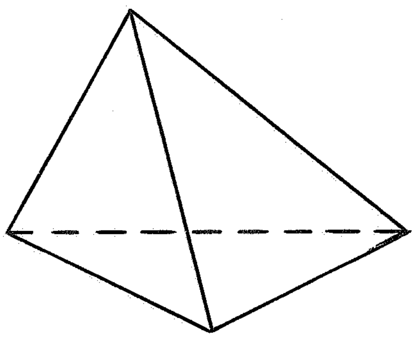
De inhoud van S is

$$\frac{1}{n!}\,|\det(b_1,\ldots,b_n)| \qquad \left(\text{dus } \frac{1}{n!} \times \text{de inhoud van bovenstaande par.}\right)$$

Hierbij weer $b_1,\ldots,b_n$ als kolommen geschreven.

Men kan gemakkelijk inzien dat elke par. bestaat uit n! disjuncte simplices, elk met gelijke inhoud.

Kontroleer zelf dat voor n = 2,3 aan bovenstaande formule voldaan is.



n = 3

Berekening van de inverse van een matrix

Zij A een $n \times n$ matrix. Op blz. 95 hebben we reeds laten zien dat A dan en slechts dan regulier is (d.w.z. A^{-1} bestaat) indien rang $A < n$. M.a.w. A is regulier d.e.s.d. $\det A \neq 0$. We zullen hieronder een methode geven om A^{-1} werkelijk uit te rekenen.

We merken eerst op dat als $AB = I$ dan A,B regulier zijn en elkaars inverse. Inderdaad, $\det AB = 1 = \det A . \det B$ d.w.z. $\det A \neq 0$; $\det B \neq 0$. Verder geldt $A^{-1}(AB) = A^{-1} = (A^{-1}A)B = IB = B$. Evenzo $(AB)B^{-1} = B^{-1} =$ $= A(BB^{-1}) = A$.

Stelling 9. Stel

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \qquad B = \begin{pmatrix} b_{11} & \cdots & b_{1n} \\ \vdots & & \\ b_{n1} & \cdots & b_{nn} \end{pmatrix}$$

en dat $b_{ij} = m_{ji}/(\det A)$, waarin m_{ji} de minor van a_{ji} in A is. Dan is $AB = I$.

Bewijs. Vroeger bewezen we

$$\det A = a_{i1}\, m_{i1} + \ldots + a_{in}\, m_{in} .$$

Ook geldt $$0 = a_{i1}\, m_{k1} + \ldots + a_{in}\, m_{kn} \qquad i \neq k \quad \text{(Ga dit na).}$$

Hetgeen gelezen kan worden als

$$\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} m_{11} & \cdots & m_{n1} \\ & & \\ m_{1n} & \cdots & m_{nn} \end{pmatrix} = \det A . I$$

Gevolg. De matrix $B = (b_{ij})$ met $b_{ij} = m_{ji}/\det A$ is de inverse van A. Dit geeft ons een methode om $A^{-1} = B$ werkelijk te berekenen.

<u>Stelling 10</u>. Als A en B regulier zijn dan is ook AB regulier en $(AB)^{-1} = B^{-1} A^{-1}$.

<u>Bewijs</u>. $(AB)(B^{-1}A^{-1}) = A(BB^{-1})A^{-1} = (AI)A^{-1} = AA^{-1} = I$.

<u>Coordinatentransformaties</u>.

Laat A een vectorruimte zijn waarvan $b_1,\ldots,b_n$ een basis is. Aan elke x kunnen we de coördinaten toekennen t.o.v. $b_1,\ldots,b_n$ n.l. de getallen $x_1,\ldots,x_n$ eenduidig bepaald door $x = x_1b_1 + \ldots + x_nb_n$.

Gaan we uit van een tweede basis $c_1,\ldots,c_n$ dan behoren bij dezelfde vector x andere getallen, die we door $x_1^*,\ldots,x_n^*$ zullen voorstellen.

We kunnen de c's in de b's uitdrukken

$$\begin{aligned} c_1 &= s_{11}\, b_1 + \ldots + s_{n1}\, b_n \\ &\vdots \\ c_n &= s_{1n}\, b_1 + \ldots + s_{nn}\, b_n \end{aligned}$$

Dan vinden we de $x_1,\ldots,x_n$ uit de $x_1^*,\ldots,x_n^*$ door schematisch

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = S \begin{pmatrix} x_1^* \\ \vdots \\ x_n^* \end{pmatrix} \qquad \text{met} \qquad S = \begin{pmatrix} s_{11} & \ldots & s_{1n} \\ \vdots & & \\ s_{n1} & \ldots & s_{nn} \end{pmatrix}$$

S heet de <u>matrix van deze coördinatentransformatie</u> (d.i. deze overgang van $x_1,\ldots,x_n$ op $x_1^*,\ldots,x_n^*$).

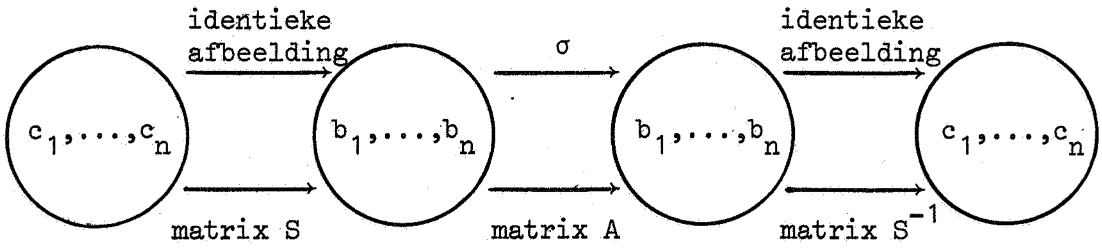
Het is gemakkelijk in te zien dat S regulier is; immers het is de matrix van de eeneenduidige lineaire trafo die de b_i's in de c_i's overvoert.

<u>Stelling 11</u>. Laat σ een lineaire transformatie zijn van A in zichzelf. Als A de matrix is van σ t.o.v. de basis $b_1,\ldots,b_n$ en A^* de matrix van σ t.o.v. de basis $c_1,\ldots,c_n$, dan geldt als S de matrix is van de coördinatentransformatie:

$$A^* = S^{-1}AS \; .$$

<u>Bewijs</u>. Ga dit zelf na.

Schematisch is de situatie als volgt:



Samenstelling geeft $A^* = S^{-1}AS$.

§4. Eigenwaarden en eigenvectoren van een lineaire transformatie.

Zij σ een lineaire afbeelding van een lineaire ruimte A in zichzelf. Een getal λ heet eigenwaarde van σ, indien er een element $a \in A$, $a \neq 0$ bestaat, zodat $\sigma(a) = \lambda a$.

In het geval dat A uit functies bestaat, dan heet a een eigenfunctie van σ. Is A een vectorruimte, dan heet a een eigenvector van de transformatie σ; een eigenvector behorende bij de eigenwaarde λ.

Voorbeeld. Beschouw de ruimte A van alle continu differentieerbare functies op R^4 naar de complexe getallen. Beschouw de volgende lineaire afbeelding H (lineaire operator H) gedefinieerd door

$$H\Psi = \left(\frac{\partial^2\Psi}{\partial x^2} + \frac{\partial^2\Psi}{\partial y^2} + \frac{\partial^2\Psi}{\partial z^2}\right) + \underset{\uparrow\ \text{potentiële energie.}}{V(r)} \quad (\Psi = \psi(x,y,z,t))$$

Stel we zoeken naar oplossingen van de volgende differentiaalvergelijking (Schrödingergolfvergelijking)

$$i\,\bar{h}\,\frac{\partial}{\partial t}\,\Psi = H\Psi\ .$$

Bij de z.g. stationaire toestanden zoeken we oplossingen van deze vergelijking in de vorm $\Psi = \psi\,\exp(-iEt/\bar{h})$, waarin ψ onafhankelijk is van t. E is energie.

Dus $$i\,\bar{h}\,\frac{\partial}{\partial t}\,\Psi = i\,\bar{h}\ .\ -i/\bar{h}\ E\ .\ \psi\ \ .\ \exp(-iEt/\bar{h})$$

en $$H\Psi = H\psi\,\exp(-iEt/\bar{h})\ .$$

De vergelijking gaat nu over in

$$H\psi = E\psi\ . \tag{1}$$

M.a.w. het probleem van de stationaire toestanden is gereduceerd tot het zoeken van de eigenwaarden van de lineaire trafo H!. De eigenwaarden vormen het z.g. spectrum van de operator H. De functies ψ heten eigenfuncties.

Hieronder zullen we nu uitsluitend eigenwaarden en eigenvectoren beschouwen bij lineaire transformaties van vectorruimten (= lineaire ruimten met eindige dimensie).

Zij σ een lineaire transformatie van een vectorruimte A in zichzelf. $\sigma: A \to A$. We willen hieronder een methode schetsen om de eigenwaarden van σ te bepalen.

Zij $\lambda \in R$. De eis dat er een $x \neq 0$ bestaat zodat $\sigma x = \lambda x$ is equivalent met de eis dat voor zekere $x \neq 0$ $(\sigma-\lambda e)x = 0$ (e is de identieke afbeelding van A op A). Nu geldt $(\sigma-\lambda e)x = 0$ voor zekere $x \neq 0 \Longleftrightarrow \ker(\sigma-\lambda e) \neq 0$ $\Longleftrightarrow \operatorname{rang}(\sigma-\lambda e) < n \Longleftrightarrow \det(\sigma-\lambda e) = 0$.
Al met al

$$\lambda \text{ is eigenwaarde van } \sigma \Longleftrightarrow \det(\sigma-\lambda e) = 0$$

<u>Het rechterlid heet de eigenwaarde v.g.l. van σ.</u>
Om inzicht te krijgen in de rechtse vergelijking beschouwen we een basis $a_1,\dots,a_n$ van A en bepalen we de matrix presentatie van σ t.o.v. deze basis

$$\begin{aligned}
\sigma a_1 &= \sigma_{11}a_1 + \sigma_{21}a_2 + \dots + \sigma_{n1}a_n \\
&\vdots \\
\sigma a_n &= \sigma_{1n}a_1 + \dots \qquad \dots + \sigma_{nn}a_n
\end{aligned}$$

De matrix presentatie van $\sigma-\lambda e$ is nu

$$\begin{pmatrix}
\sigma_{11}-\lambda & \sigma_{12} & \sigma_{13} & \dots & \sigma_{1n} \\
\vdots & \sigma_{22}-\lambda & \ddots & & \vdots \\
 & & & \ddots & \\
\sigma_{n1} & & & & \sigma_{nn}-\lambda
\end{pmatrix}$$

De vergelijking $\det(\sigma-\lambda e)$ is nu niets anders dan de vergelijking

$$\det((\sigma_{ij})-\lambda I) = 0 \, .$$

Het linkerlid is een polynoom in de graad n van λ dat onafhankelijk is van de matrixpresentatie van σ.

Is A een n×n matrix dan stelt A een lineaire trafo voor van R^n in R^n.

$\det(A-\lambda I)=0$ heet de eigenwaarde vergelijking van A en ook wel karakteristieke vergelijking van A. Het linkerlid is een n^{de} graads polynoom in λ, de karakteristieke veelterm van A.

Als S een reguliere n×n matrix is dan geldt

$$\det(S^{-1}AS-\lambda I) = \det(S^{-1}AS-S^{-1}\lambda IS) = \det(S^{-1}(A-\lambda I)S) =$$

$$= \det S^{-1}.\ \det(A-\lambda I)\det S = \det S^{-1}.\ \det S.\ \det(A-\lambda I) = \det(A-\lambda I)\ .$$

De karakteristieke vergelijking van A verandert dus niet als we A door $S^{-1}AS$ (S regulier) vervangen. Dit laatste is eigenlijk duidelijk omdat $S^{-1}AS$ de matrix is van dezelfde transformatie A alleen t.o.v. een andere basis beschreven.

Zij $\lambda_1,\ldots,\lambda_n$ eigenwaarden van A dan vindt men de bijbehorende eigenvectoren door oplossing van de vergelijking $(A-\lambda I)(x) = 0$ voor $\lambda = \lambda_1,\ldots,\lambda_n$.

Stelling 1. Laat $b_1,\ldots,b_n$ een basis zijn en laat $b_1,\ldots,b_n$ tevens eigenvectoren zijn van een lineaire transformatie σ. Dan is de matrix die σ t.o.v. deze basis beschrijft de diagonale matrix

$$\begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}$$

nullen buiten de hoofddiagonaal

Ga dit na.

Gevolg. In dit geval bestaat er een reguliere matrix S zodat $S^{-1}AS$ de diagonaalvorm heeft!

We merken op, dat niet elke matrix n-lineaire onafhankelijke eigenvectoren heeft. Voorbeeld $\begin{pmatrix}1&0\\0&1\end{pmatrix}$. Karakteristieke vergelijking $(\lambda-1)^2 = 0$. En bij $\lambda = 1$ behoren slechts de eigenvectoren $\binom{a}{0}$. Ook moeten we rekening houden met het feit dat niet elke algebraïsche vergelijking reële wortels heeft.

Dit bezwaar is te ondervangen door op complexe getallen over te gaan. Tot nu toe hebben we alles met reële getallen gedaan maar de gehele lineaire algebra, voor zover we die tot nu toe hebben ontwikkeld, blijft onverminderd van kracht als we ook complexe getallen toelaten. Dit zullen we nu in het vervolg van de paragraaf doen. We kunnen dan de z.g. hoofdstelling van de algebra toepassen die zegt dat elke veelterm van de graad n in lineaire factoren is te ontbinden.

Stelling 2. Bij elke (n×n) matrix A kan men een complexe reguliere (n×n) matrix S vinden, zodat $S^{-1}AS$ een driehoeksmatrix is, dit is een matrix in de vorm

$$\begin{pmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ 0 & b_{22} & \dots & b_{22} \\ 0 & 0 & & \\ \cdot & 0 & & \\ \cdot & & & \\ 0 & 0 & \dots & b_{nn} \end{pmatrix} \qquad \text{(d.w.z. } b_{ij} = 0 \text{ als } i > j\text{)}$$

In de diagonaal staan dan de eigenwaarden van A, want de karakteristieke vergelijking is $(\lambda - b_{11}) \dots (\lambda - b_{nn}) = 0$.

Bewijs. --

Ondanks dat voor niet elke lineaire transformatie er een basis van eigenvectoren bestaat kunnen we de volgende stelling als een voldoende voorwaarde hiertoe rekenen.

Stelling 3. Een collectie eigenvectoren die behoren bij verschillende eigenwaarden van een lineaire trafo σ zijn lineair onafhankelijk. Zijn i.h.b. de eigenwaarden van σ allen verschillend, dan is er een basis van eigenvectoren van σ.

Bewijs. Stel $x_1, \dots, x_n \neq 0$ en zij

$$\begin{aligned} \sigma x_1 &= \lambda_1 x_1 \qquad \text{met } \lambda_i \neq \lambda_j \text{ voor } i,j = 1,\dots,n. \\ &\vdots \\ \sigma x_n &= \lambda_n x_n \end{aligned}$$

Om te bewijzen dat $x_1,\dots,x_n$ lineair onafhankelijk zijn, laat

$$\alpha_1 x_1 + \dots + \alpha_n x_n = 0 \qquad \alpha_i \in R \tag{1}$$

We moeten aantonen $\alpha_1 = \alpha_2 = \dots = \alpha_n = 0$.
Pas σ toe op (1) dan verkrijgen we

$$\alpha_1 \lambda_1 x_1 + \dots + \alpha_n \lambda_n x_n = 0$$

en successievelijk door toepassing van k' keer de transformatie σ:

$$\alpha_1 \lambda_1^{k-1} x_1 + \dots + \alpha_n \lambda_n^{k-1} x_n = 0$$

Dus voor $k = 1,\dots,n$ hebben we

$$\begin{cases} \lambda_1 \alpha_1 x_1 + \lambda_2 \alpha_2 x_2 + \dots + \lambda_n \alpha_n x_n = 0 \\ \vdots \\ \lambda_1^{n-1} \alpha_1 x_1 + \lambda_2^{n-1} \alpha_2 x_2 + \dots + \lambda_n^{n-1} \alpha_n x_n = 0 \end{cases}$$

In matrix notatie:

$$\begin{pmatrix} 1 & 1 & \dots & 1 \\ \lambda_1 & \lambda_2 & \dots & \lambda_n \\ \vdots & & & \\ \lambda_1^{n-1} & \lambda_2^{n-1} & \dots & \lambda_n^{n-1} \end{pmatrix} \begin{pmatrix} \alpha_1 x_1 \\ \alpha_2 x_2 \\ \vdots \\ \alpha_n x_n \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

Stel B de matrix uit het linkse deel van het linkerlid. Dan
det B = $(\lambda_2-\lambda_1)(\lambda_3-\lambda_2)(\lambda_3-\lambda_1) \dots = \prod_{i>j} (\lambda_i-\lambda_j) \neq 0$ omdat de eigenwaarden van σ allen verschillend zijn. Dus B^{-1} bestaat. Hieruit volgt formeel

$$\begin{pmatrix} \alpha_1 x_1 \\ \vdots \\ \alpha_n x_n \end{pmatrix} = B^{-1} \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

ofwel $\alpha_1 x_1 = \dots = \alpha_n x_n = 0$.
Omdat $x_1,\dots,x_n \neq 0$ volgt inderdaad
$\alpha_1 = \dots = \alpha_n = 0$.

In het algemeen kan men formules opstellen voor het aantal lineair onafhankelijke eigenvectoren van een transformatie σ. We zullen hier niet nader op ingaan.

Een zeer interessante stelling is de stelling van Cayley en Hamilton die het volgende zegt:

Stelling 4. Zij A een n×n matrix en $b_n\lambda^n + b_{n-1}\lambda^{n-1} + \ldots + b_0 = 0$ de karakteristieke vergelijking van A. Dan geldt

$$b_nA^n + b_{n-1}A^{n-1} + \ldots + b_1A + b_0I = 0 .$$

Toepassingen

Laat f(A) een matrixpolynoom zijn van (willekeurige) graad m.
Dus $f(A) = \alpha_mA^m + \alpha_{m-1}A^{m-1} + \ldots + \alpha_1A + \alpha_0I$; A is een n×n matrix.
We zullen hieronder een methode aangeven om f(A) snel te berekenen.

Laat $f(\lambda)$ het corresponderende polynoom zijn in λ en $d(\lambda)$ de karakteristieke polynoom van A.

Een stelling uit de algebra zegt dat

$$f(\lambda) = d(\lambda)\ q(\lambda) + r(\lambda)$$

waarin $r(\lambda)$ de restterm heet. $r(\lambda)$ heeft graad $\leq n-1$. Als q(A) en r(A) de corresponderende veeltermen zijn in A dan geldt

$$f(A) = d(A)\ q(A) + r(A).$$

Omdat d(A) = 0 volgens de vorige stelling, volgt

$$f(A) = r(A)$$

m.a.w. f(A) is te schrijven als n-1-graadspolynoom in A. Dus

$$f(A) = a_{n-1}A^{n-1} + a_{n-2}A^{n-2} + \ldots + a_1A + a_0I .$$

Voor het geval A n verschillende eigenwaarden bezit kunnen we de coëfficiënten $a_0,\ldots,a_{n-1}$ bepalen. Dit gaat als volgt:

Laat $\lambda_1,\ldots,\lambda_n$ de eigenwaarden zijn van A en $x_1,\ldots,x_n$ de bijbehorende eigenvectoren.
Als $1 \leq i \leq n$ dan $f(A).x_i = f(\lambda_i).x_i$ en $r(A).x_i = r(\lambda_i)x_i)$. Dus $r(\lambda_i) = f(\lambda_i)$ of wel

$$\begin{cases} f(\lambda_1) = a_{n-1}\lambda_1^{n-1} + a_{n-2}\lambda_1^{n-2} + \ldots a_0 \\ \vdots \\ f(\lambda_n) = a_{n-1}\lambda_n^{n-1} + a_{n-2}\lambda_n^{n-2} + \ldots a_0 \end{cases}$$

Dit is een stelsel van n lineaire vergelijkingen in $a_0,\ldots,a_{n-1}$ en de bijbehorende matrix heeft de determinant

$$\begin{vmatrix} 1 & 1 & \ldots & 1 \\ \lambda_1 & \lambda_2 & \ldots & \lambda_n \\ \vdots & & & \\ \lambda_1^{n-1} & \lambda_2^{n-1} & \ldots & \lambda_n^{n-1} \end{vmatrix} = \prod_{i>j} (\lambda_i - \lambda_j) \neq 0 \ .$$

Dus is het stelsel eenduidig oplosbaar en $a_0,\ldots,a_{n-1}$ zijn bekend.

Voorbeeld. Zij $A = \begin{bmatrix} -3 & -4 \\ 2 & 3 \end{bmatrix}$. Te berekenen A^{593}.

Oplossing: De eigenwaarden van A zijn $\lambda_1 = 1$, $\lambda_2 = -1$. Laat $f(A) = A^{593}$.
Dus $A^{593} = a_1 A + a_0 I$.
Nu volgt

$$f(1) = 1^{593} = 1 = a_1 1 + a_0$$

$$f(-1) = \quad = -1 = a_1(-1) + a_0$$

Ofwel $a_0 = 0$, $a_1 = 1$.
Substitutie geeft $\begin{pmatrix} -3 & -4 \\ 2 & 3 \end{pmatrix}^{593} = \begin{pmatrix} -3 & -4 \\ 2 & 3 \end{pmatrix}$.

Exponentiële machten van matrices.

Zij A een n×n matrix. Dan definiëren we

$$e^A = I + \frac{A}{1!} + \frac{A^2}{2!} + \frac{A^3}{3!} + \dots + \frac{A^m}{m!} + \dots$$

e^A is dus de matrix die op de ij'de plaats $\lim_{m\to\infty} (S_m)_{ij}$ heeft staan.
Hierbij is $S_m = \sum_{k=0}^{m} \frac{A^k}{k!}$.

We gaan niet op convergentie problemen in. De functie e^X voor X een n×n matrix is een generalisatie van de functie e^x die natuurlijk en nuttig blijkt te zijn.

O.a. geldt:

$$\begin{cases} e^{A+B} = e^A.e^B & \text{als } AB = BA \text{ !} \\ e^{t_1A}.e^{t_2A} = e^{(t_1+t_2)A} & \text{als } t_1,t_2 \in R \text{ ; } A \text{ willekeurig.} \end{cases}$$

Uit de tweede gelijkheid volgt o.a. dat e^A altijd regulier is (neem $t_1 = -t_2 = 1$ in de tweede bewering).
We kunnen bovenstaande methode nu ook toepassen om e^A snel te berekenen.

Blijkbaar is elke partiële som $S_m = \sum_{k=0}^{m} \frac{A^k}{k!}$ te schrijven als n-1 -graadspolynoom in A volgens het bovenstaande. Nu geldt deze eigenschap ook voor $\lim_{m\to\infty} S_m = e^A$.

Dus

$$e^A = a_{n-1}A^{n-1} + a_{n-2}A^{n-2} + \dots a_1A + a_0I \text{ .}$$

Zijn $\lambda_1,\dots,\lambda_n$ nu weer verschillende eigenwaarden van A dan zijn $a_0,\dots,a_{n-1}$ weer te berekenen uit het stelsel vergelijkingen

$$\begin{cases} e^{\lambda_1} = a_{n-1}\lambda_1^{n-1} + a_{n-2}\lambda_1^{n-1} + \dots + a_0 \\ \vdots \\ e^{\lambda_n} = a_{n-1}\lambda_n^{n-1} + \dots \qquad \dots + a_0 \end{cases}$$

Voorbeeld. Zij $A = \begin{bmatrix} -3 & -4 \\ 2 & 3 \end{bmatrix}$. Te berekenen e^A.

We moeten nu oplossen

$$e^{1} = \alpha_1 . 1 + \alpha_0$$

$$e^{-1} = \alpha_1(-1) + \alpha_0$$

Dus $\alpha_1 = \frac{1}{2}(e-e^{-1})$, $\alpha_0 = \frac{1}{2}(e+e^{-1})$

i.e. $e^A = \frac{1}{2}(e-e^{-1})\begin{pmatrix} -3 & -4 \\ 2 & 3 \end{pmatrix} + \frac{1}{2}(e+e^{-1})\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} -e+2e^{-1} & -2e+2e^{-2} \\ e-e^{-1} & 2e-e^{-1} \end{pmatrix}$

Men kan zich afvragen of het zin heeft om de functie e^X in te voeren. Deze functie komt te voorschijn bij het oplossen van stelsels lineaire differentiaalvergelijkingen.

$$(**)\begin{cases} \dfrac{dx_1(t)}{dt} = a_{11}(t)x_1(t) + a_{12}(t)x_2(t) + \ldots + a_{1n}(t)x_n(t) + f_1(t) \\ \vdots \\ \dfrac{dx_n(t)}{dt} = a_{n1}(t)x_1(t) + a_{n2}(t)x_2(t) + \ldots + a_{nn}(t)x_n(t) + f_n(t) \end{cases}$$

Definieer

$$x(t) = \begin{pmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{pmatrix} \quad \text{en} \quad A(t) = \begin{pmatrix} a_{11}(t) a_{12}(t) & \ldots & a_{1n}(t) \\ \vdots & & \\ a_{n1}(t) a_{n2}(t) & \ldots & a_{nn}(t) \end{pmatrix}$$

Dit stelsel kan dan geschreven worden als

$$\frac{dx(t)}{dt} = A(t)x(t) + f(t) \qquad x(t_0) = C.$$

Bij het oplossen van deze matrixvergelijking kunnen we nu in principe op dezelfde wijze te werk gaan als ware $A(t)$, $x(t)$ en $f(t)$ gewone reële functies.

Als oplossing vinden we nu ook

$$\boxed{x(t) = e^{A(t-t_0)}\, C + e^{At} \int_{t_0}^{t} e^{-As}\, f(s)\, ds}$$

Zie [ ] voor verdere informatie.

§5. Ruimten met inwendig product, Hilbertruimten

Definitie Een inwendig product op een complexe lineaire ruimte L is een complexwaardige functie ϕ op L×L met de volgende eigenschappen

$$\begin{cases} 1.\ \phi(x,y) = \overline{\phi(y,x)} & \text{voor alle } x,\ y \in L \\ 2.\ \phi(x,x) \geq 0 & \text{voor alle } x \in L \\ 3.\ \phi(x,x) = 0 & \text{dan en slechts dan indien } x = 0 \\ 4.\ \phi(\lambda_1 x_1 + \lambda_2 x_2, y) = \lambda_1\, \phi(x_1,y) + \lambda_2\, \phi(x_2,y) & \text{voor alle } x_1, x_2,\ y \in L \\ \quad \text{en } \lambda_1, \lambda_2 \in C. \end{cases}$$

Is een inproduct op L gegeven dan heet L een complexe lineaire ruimte met inwendig product.

Zij nu L een reële lineaire ruimte. Dan heet een reëel-waardige functie op L×L een inwendig product op L indien 1, 2, 3 en 4 gelden. Blijkbaar geldt dan:

$$\phi(x,y) = \phi(y,x) \qquad \text{voor alle } x,\ y \in L, \text{ m.a.w. } \phi \text{ is}$$

symmetrisch.

In beide gevallen schrijven we $<x,y>$ of ook wel (x,y) i.p.v. $\phi(x,y)$. $\sqrt{(x,x)}$ noteren we wel met $||x||$, de norm van x. Merk op dat uit de eerste eis al volgt dat $||x||$ steeds reëel en groter of gelijk nul is.

Het inproduct is, behalve lineair in de eerste variabele ook lineair in de tweede variabele voor een reële lineaire ruimte en antilineair in de tweede variabele voor het geval van een complexe lineaire ruimte met inwendig product:

$$\boxed{(x,\lambda_1 y_1 + \lambda_2 y_2) = \overline{(\lambda_1 y_1 + \lambda_2 y_2,x)} = \overline{\lambda_1}(x,y_1) + \overline{\lambda_2}(x,y_2)}\,.$$

Stelling 1. Zij L een reële of complexe lineaire ruimte met inproduct (,). Dan geldt

$$||(x,y)|| \leq ||x||.||y||.$$

Opmerking. Bovenstaande ongelijkheid heet de ongelijkheid van Cauchy-Schwarz of Cauchy-Boenjakowski.

Bewijs van Stelling 1. Neem twee punten $x,y \in L$. We mogen aannemen dat $(x,y) \neq 0$. We onderzoeken nu de uitdrukking

$$Q(\tau) = ||x + \tau y||^2 = (x + \tau y,\ x + \tau y) = (x,x) + |\tau|^2(y,y)$$
$$+ \overline{\tau}\,(x,y) + \tau(y,x)\ ;$$

τ een complexe variabele. Voor willekeurige τ is $Q(\tau)$ reëel en ≥ 0; de laatste twee termen uit het rechterlid zijn elkaars complex geconjugeerde. Stellen we

$$\tau = \frac{\overline{(x,y)}}{|(x,y)|}\ t \quad \text{(t reëel) dan komt er}$$

$$(x,x) + t^2(y,y) + 2t|(x,y)|.$$

Dit is een kwadratische vorm in t met uitsluitend reële coefficienten, die steeds ≥ 0 zijn. Dan is de discriminant

$$(x,x)\ .\ (y,y) - |(x,y)|^2 \geq 0.$$

Daarmee is het gestelde bewezen.

Ga zelf na dat dit bewijs voor reële inwendige producten (op reële lineaire ruimten) sterk vereenvoudigd kan worden.

Stelling 2. Zij L een complexe of reële lineaire ruimte met inwendig product. Dan geldt

1) $||x|| \geq 0 \qquad \forall x \in L$ en $= 0 \Leftrightarrow x = 0$

2) $||\alpha x|| = |\alpha|.||x|| \qquad \forall \alpha \in \mathbb{R}$ (of $\mathbb{C}$)

3) $||x+y|| \leq ||x|| + ||y||$. (Ongelijkheid van Minkowski)

Zo'n functie $||\ \ ||$ met deze 3 eigenschappen heet wel een norm op L. Dus het inw. product (,) op L induceert een norm op L.

Bewijs van stelling 2. (we bewijzen slechts 3))

$$||x+y||^2 = (x+y,x+y) = (x,x) + (x,y) + (y,x) + (y,y)$$

$$\leq (x,x) + (y,y) + 2|(x,y)| \leq ||x||^2 + ||y||^2 + 2||x|| \; ||y||$$

$$= (||x|| + ||y||)^2, \text{ dus}$$

$$||x+y|| \leq ||x|| + ||y||.$$

Opgave: Bewijs de ongelijkheid $||x-y|| \geq \Big| ||x|| - ||y|| \Big|$

De norm $|| \ ||$ induceert nu een metriek op de ruimte L met inwendig product (,) door de volgende definitie

$$\rho(x,y) = ||x-y||.$$

(Ga na dat de driehoeks-ongelijkheid een regelrecht gevolg is van de ongelijkheid van Minkowski).

Definitie. Stel L een (reële of complexe) lineaire ruimte met inwendig product. Dan heet L een Hilbertruimte indien de door het inproduct geinduceerde metriek L tot een volledige metrische ruimte maakt.

Al met al geeft een inproduct dus aanleiding tot een topologie. In deze topologie is de functie (,): $L \times L \to \mathbb{R}$ of $\mathbb{C}$ continu in beide variabelen.

(Ga dit zelf na).:

$$(x,y) - (x_0,y_0) = (x - x_0,y) + (x_0,y - y_0),$$

dus

$$|(x,y) - (x_0,y_0)| \leq ||x - x_0||. \; ||y|| + ||x_0|| \; ||y - y_0||$$

$$\leq ||x - x_0|| \; (||y_0|| + ||y - y_0||) + ||x_0||.||y - y_0||.$$

Ook de afbeeldingen $L\times L \to \mathbb{C}$(of $\mathbb{R}$) met $(x,y) \to x+y$ en $\mathbb{R}\times L \to L$ (of $\mathbb{C}\times L \to L$) met $(\lambda,x) \to \lambda x$ zijn continu.
Dit heeft o.a. de volgende consequenties

I. Als G een lineaire deelruimte is van L dan is de afsluiting $\overline{G}$ van G in L weer een lineaire deelruimte van L.

II. Als $x_n \to x$ in L en $y_n \to y$ in L (gewone convergentie in de metrische ruimte (L,ρ)); en $(x_n,y_n) = 0$ voor alle n, dan geldt ook $(x,y) = 0$.

III. Als $x_n \to x$ in L dan $\lim_{n\to\infty} ||x_n|| = ||x||$ (continuiteit van de norm $||\ ||$).

<u>Voorbeelden</u>

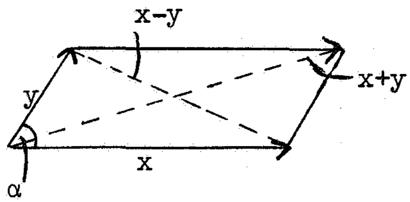
1) De ruimte $\mathbb{C}^n$. De vorm $(x,y) = \sum_{i=1}^{n} x_i\overline{y_i}$ is een inproduct op $\mathbb{C}^n$; we hebben $(x,y) = \overline{(y,x)}$. De bijbehorende norm is $||x|| = (\sum_{i=1}^{n} |x_i|^2)^{\frac{1}{2}}$; in deze norm is $\mathbb{C}^n$ volledig. Dus is $\mathbb{C}^n$, met het genoemde inproduct een Hilbertruimte.

2) De ruimte $\mathbb{R}^n$. De vorm $(x,y) = \sum_{i=1}^{n} x_iy_i$ is een inproduct op $\mathbb{R}^n$; we hebben $(x,y) = (y,x)$. De bijbehorende norm is $||x|| = (\sum_{i=1}^{n} x_i{}^2)^{\frac{1}{2}}$, in deze norm is $\mathbb{R}^n$ volledig. Dus is $\mathbb{R}^n$, met het genoemde inproduct een Hilbertruimte. De ongelijkheid in stelling 1 is niets anders dan de gewone ongelijkheid van Cauchy-Schwarz zoals we die reeds kenden. De metriek geinduceert door dit inproduct is precies dezelfde als de gewone metriek die we op $\mathbb{R}^n$ reeds kenden.
Om de <u>meetkundige betekenis van het inproduct</u> op $\mathbb{R}^3$ te onderzoeken beschouwen we vectoren in de "stereometrische ruimte", met basis e_1, e_2, e_3 bestaande uit drie vectoren die (in stereometrische zin) twee aan twee loodrecht op elkaar staan, en de lengte 1 hebben. Het inwendige product van de vector $x = (x_1,x_2,x_3)$ en $y = (y_1,y_2,y_3)$ is $x_1y_1 + x_2y_2 + x_3y_3$. Bekend is dat $x_1^2 + x_2^2 + x_3^2 = ||x||$ de lengte van de vector x is. Als de hoek tussen x en y recht is dan geldt

volgens de stelling van Pythagoras $||x-y||^2 = ||x||^2 + ||y^2||$.
Daar $(x-y,x-y) = (x,x-y) - (y,x-y) = (x,x) - 2(x,y) + (y,y)$ blijkt nu dat $(x,y) = 0$. Algemener is α de hoek tussen x en y dan geldt volgens de projectiestelling $||x-y||^2 = ||x||^2 + ||y||^2 - 2||x||.||y|| \cos\alpha$
dus $\boxed{\cos\alpha = (x,y)/||x||.||y||.}$

Stereometrisch en Planimetrisch kennen we ook de parallologram wet d.w.z. er geldt:

$$||x+y||^2 + ||x-y||^2 = 2||x||^2 + 2||y||^2.$$



Deze eigenschap geldt nu in willekeurige ruimten met inwendig product:

Stelling 3. Zij L een lineaire (complexe of reële) ruimte met inwendig product. Dan geldt

$$||x-y||^2 + ||x+y||^2 = 2||x||^2 + 2||y||^2.$$

Bewijs $\quad ||x-y||^2 = (x-y,x-y) = (x,x) - (x,y) - (y,x) + (y,y)$

$$||x+y||^2 = (x+y,x+y) = (x,x) + (x,y) + (y,x) + (y,y)$$

optelling geeft het gewenste resultaat.

Voorbeeld 3) Laat $L=\{f : [0,1] \to C | f \text{ continu }\}$. L is een lineaire ruimte m.b.v. de gewone optelling van functies en gewone scalaire vermenigvuldiging.
Definieer:

$$(f,g) = \int_0^1 f(x) \,.\, \overline{g(x)}\, dx$$

$(f,g) = \overline{(g,f)}$ komt neer op $(f,g) = \int_0^1 f(x)\, \overline{g(x)}\, dx = \int_0^1 \overline{g(x) \,.\, \overline{f(x)}}\, dx$
$= \overline{(g,f)}$. Ook geldt $(f,g) = \int_0^1 f(x) \,.\, \overline{f(x)}\, dx = \int_0^1 |f(x)|^2\, dx$.
Dus (,) vormt een inproduct op L.

We kunnen ook definieren

$$(f,g) = \int_0^1 p(x)\, f(x)\, \overline{g(x)}dx$$ als $p(x)$ een vaste positieve continue functie is. Met deze definitie wordt L ook een ruimte met inwendig product. Het is geen Hilbertruimte als we ons slechts tot continue-functie beperken, zie blz. 58. Wanneer we alle kwadratisch integreerbare functie's toelaten op L dan wordt met dit inproduct wel een Hilbertruimte (zie later).

Voorbeeld 4. Laat A een reële vectorruimte zijn en $a_1,\dots,a_n$ een basis voor A. Als $x = x_1a_1+\dots+x_na_n$ en $y = y_1a_1+\dots+y_na_n$ definieer dan

$(x,y) = \sum_{i=1}^{n} x_iy_i$. (,) is een inproduct op A (vergelijk met voorbeeld).

Voorbeeld 5. Zij $H = \{x = (x_1,\dots,x_n,\dots) \mid \sum_{i=1}^{\infty} x_i^2 < \infty\}$.

Definieer voor $x,y \in H$, $x = (x_1,\dots,x_n,\dots)$ $y = (y_1,\dots,y_n,\dots;)$

$$(x,y) = \sum_{i=1}^{\infty} x_iy_i \quad .$$

Dan is (,) een inwendig product op H. H is met de door (,) geïnduceerde norm een volledige metrische ruimte. Dus H is een Hilbertruimte. Merk op dat de door $\|\ \|$ geïnduceerde metriek dezelfde is als het voorbeeld op blz. 33 .

I.p.v. H schrijven we meestal l_2.

Definitie. Zij H een Hilbertruimte. Twee elementen $x,y \in H$ heten onderling loodrecht of orthogonaal en we schrijven $x \perp y$, indien $(x,y) = 0$. Als G een deelruimte is van H en $x \perp y$ voor alle $y \in G$ dan zeggen we dat x loodrecht staat op G en we schrijven $x \perp G$.

Stelling H. De verzameling elementen $z \in H$ met $z \perp G$ is een gesloten lineaire deelruimte van H. Deze verzameling geven we dikwijls aan met $G^{\perp}$; het z.g. orthocomplement van G.

Bewijs. Als $z_1 \perp G$ en $z_2 \perp G$ dan is $(z_1,x) = (z_2,x) = 0 \quad \forall x \in G$
dus $(\alpha_1 z_1+\alpha_2 z_2,x) = 0 \quad \alpha_1, \alpha_2 \in \mathbb{C}$ (of $\mathbb{R}$) en $x \in G$.
dus $\alpha_1 z_1 + \alpha_2 z_2 \perp G$ voor $\alpha_1, \alpha_2 \in \mathbb{C}$.
Is $z_n \perp G$ voor $n = 1,2,\ldots$ en convergeert (z_n) naar z, dan hebben we $(z_n,x) = 0$ voor alle n en voor $x \in G$ dus $(z,x) = 0$ voor $x \in G$ ofwel $z \perp G$ vanwege de continuiteit van het inproduct.

Opmerking.
De term loodrecht sluit direkt aan met de meetkundige betekenis die we hieraan kunnen toekennen in $\mathbb{R}^3$. Immers, als we de $\mathbb{R}^3$ beschouwen dan staan meetkundig de vectoren x en y loodrecht op elkaar dan en slechts dan indien $(x,y) = x_1y_1 + x_2y_2 + x_3y_3 = 0$.
I.h.a. moeten we deze meetkundige situatie uit ons hoofd laten verdwijnen. Immers bij voorbeeld 3 geldt $f \perp g$ dan en slechts dan indien
$$\int_0^1 f(x)\,\overline{g(x)}\,dx = 0.$$
We kunnen hieraan geen meetkundige consequenties verbinden.

T.a.v. loodrechtheid vermelden we nog de stelling van Pythagoras:

Stelling 5. Als $x \perp y$ dan geldt $||x+y||^2 = ||x||^2 + ||y||^2$
Bewijs.

$$(x+y,x+y) = (x,x) + (x,y) + (y,x) + (y,y) = (x,x) + (y,y).$$

Opmerking. Ook geldt dan $||x-y||^2 = ||x||^2 + ||y||^2$ (gebruik de parallelogramwet.

Orthonormale stelsels.
In het volgende doorloopt i een vast indexverzameling I, mogelijk zelfs overaftelbaar.

Definitie. Een stelsel elementen $\{e_i | i \in I\}$ in een Hilbertruimte H heet een orthonormaal stelsel als geldt

1) $||e_i|| = 1$ voor alle $i \in I$

2) $(e_i, e_{i'}) = 0$ als $i \neq i'$.

M.a.w. de vectoren e_i hebben lengte 1 en zijn onderling loodrecht. Deze voorwaarden hebben tengevolge dat ieder eindig deelstelsel van $\{e_i | i \in I\}$ lineair onafhankelijk is.
Immers, als

$$\alpha_1 e_{i_1} + \alpha_2 e_{i_2} + \ldots \alpha_n e_{i_n} = 0 \quad \text{voor zekere } \alpha_i \in \mathbb{R} \text{ (of } \mathbb{C})$$

dan geldt voor elke index i_k $(1 \leq k \leq n)$

$$(\alpha_1 e_{i_1} + \ldots + \alpha_n e_{i_n}, e_{i_k}) = 0 = \alpha_{i_k}(e_{i_k}, e_{i_k}) = \alpha_{i_k} .$$

<u>Voorbeeld</u>. We beschouwen de numerieke n-dimensionale vectorruimte $\mathbb{R}^n$. Als $x = (x_1, \ldots, x_n)$; $y = (y_1, \ldots, y_n)$ dan $(x,y) = x_1 y_1 + \ldots + x_n y_n$. Met dit inwendig product is nu het stelsel $\{e_1, \ldots, e_n\}$ bestaamde uit de n eenheidsvectoren een orthonormaal stelsel. Het is zelfs een <u>orthonormale basis</u>.

<u>Definitie</u>. Zij $\{a_i | i \in I\}$ een stelsel elementen van een lineaire ruimte. L. Onder het <u>lineaire omhulsel</u> van het stelsel $\{a_i | i \in I\}$ verstaat men per definitie de verzameling van alle $x \in L$ met $x = \lambda_1 a_{i_1} + \ldots + \lambda_n a_{i_n}$; $\lambda_1, \ldots, \lambda_n \in \mathbb{R}$ (of $\mathbb{C}$ als L een complexe lineaire ruimte is) en $i_1, \ldots, i_n \in I$.

Het lineaire omhulsel bestaat dus uit alle <u>eindige</u> lineaire combinaties van elementen van het stelsel.
Nu geldt de volgende stelling.

<u>Stelling 6</u>. Zij H een Hilbertruimte en $\{x_k\}_{k=1}^{\infty}$ een willekeurige <u>rij</u> elementen van H. (de indexverzameling I is dus nu aftelbaar). Dan bestaat een afbrekende of oneindige orthonormale rij $\{e_k\}_{k=1}^{\infty}$ zo dat de verzameling elementen $\{x_k | k=1,2,\ldots\}$ hetzelfde lineair omhulsel heeft als de verzameling der elementen $\{e_k | k=1,\ldots\}$.

Dus met een voor de hand liggende notatie

$$L(x_1,x_2,\dots,x_k,\dots) = L(e_1,e_2,\dots,e_k,\dots).$$

<u>Bewijs</u>. We lopen de rij $x_1,x_2,\dots,x_k,\dots$ langs en schrappen daarbij elk element dat een lineaire combinatie is van de vorige elementen. We schrappen dus x_1 als $x_1 = 0$; x_2 als $x_2 = \alpha x_1$ voor zekere α, enz. Op deze wijze ontstaat een afbrekende of oneindige rij van lineair onafhankelijke elementen. We geven deze rij weer met $\{x_k\}_{k=1}^{\infty}$ aan. (Merk op dat de rij d.e.s.d afbreekt indien H een vectorruimte is = eindige dimensie bezit).

In het bijzonder is $x_1 \neq 0$. We kunnen dus nemen $e_1 = x_1/||x_1||$. Zij nu y_2 een element van de vorm

$$x_2 - \alpha e_1 \qquad \alpha \in \mathbb{R} \ (\text{of } \mathbb{C}) .$$

Dan is $y_2 \neq 0$. Verder hebben we $(y_2,e_1) = (x_2,e_1) - \alpha$. Dus als we nemen $\alpha = (x_2,e_1)$ en daarna stellen $e_2 = y_2/||y_2||$ dan geldt

$$(e_2,e_1) = 0 \qquad e_2 = \beta_2 x_2 - \beta_1 e_1$$

met zekere coefficienten β_1, β_2 zodat $\beta_2 \neq 0$. Uit het laatste volgt dat $e_2 \in L(x_1,x_2)$.
Laat nu k-1 "eenheidsvectoren" $e_1,\dots,e_{k-1}$ gekozen zijn zodat de e_i $i = 1,\dots,k-1$ onderling loodrecht zijn en alle tot $L(x_1,\dots,x_{k-1})$ behoren. Voor alle $\alpha_1,\dots,\alpha_{k-1}$ is dan

$$y_k = x_k - (\alpha_1 e_1 + \alpha_2 e_2 + \dots \alpha_{k-1}e_{k-1}) \text{ een vector } \neq 0 \quad .$$

Verder is
$(y_k,e_i) = (x_k,e_i) - \alpha_i$ voor $i = 1,\dots,k-1$. Nemen we $\alpha_i = (x_k,e_i)$ $(i = 1,\dots,k-1)$ en stellen we $e_k = y_k/||y_k||$ dan is dus

(1) $(e_k,e_i) = 0$ voor $i = 1,\dots,k-1$

(2) $$\frac{y_k}{||y_k||} = e_k = \beta_{k,k}x_k - (\beta_{k,1}e_1 + \dots \beta_{k,k-1}e_{k-1}) \text{ met zekere}$$

coefficienten $\beta_{k,i}$ waarbij $\beta_{k,k} \neq 0$.

Er volgt dat er een rij (e_k) bestaat zodat voor alle k voldaan is aan (1) en (2). Uit (2) volgt dat we $e_1, e_2, \ldots$ opvolgend kunnen uitdrukken in $x_1, x_2, \ldots$; daar $\beta_{k,k} \neq 0$ voor alle k is ook het omgekeerde waar. Daarmee is de stelling bewezen.

Opmerkingen. Het beschreven proces heet orthonormaliseren en is afkomstig van Gramm-Schmidt. Op grond van deze stelling mogen we bij kwesties betreffende lineaire deelruimten voortgebracht door een rij elementen, er van uitgaan dat deze rij orthonormaal is.

Ongelijkheid van Bessel.

We merken op dat als $\{p_\alpha | \alpha \in A\}$ een collectie reële getallen is dan $\sum\limits_{p \in A} p_\alpha$ gedefinieerd is als het supremum van alle eindige deelsommen

$$\sum_{i=1}^{n} p_{\alpha_i}, \quad \alpha_1, \ldots, \alpha_n \in A.$$

Stelling 7. (ongelijkheid van Bessel).

Als $\{e_i | i \in I\}$ een orthonormaal stelsel elementen van een Hilbertruimte H is en $x \in H$, dan geldt

$$||x||^2 \geq \sum_{i \in I} |(x,e_i)|^2$$

Het getal (x,e_i) heet de i^e Fouriercoefficient van x (t.o.v. dit orthonormale stelsel).

Bewijs. Zij $i_1, \ldots, i_n \in I$ en zij voor $1 \leq k \leq n$ $\lambda_k = (x, e_{i_k})$.
Het is gemakkelijk in te zien dat

$$(x - \sum_{k=1}^{n} \lambda_k e_{i_k}) \perp \sum_{k=1}^{n} \lambda_k e_{i_k} .$$

Ook geldt nu volgens de stelling van Pythagoras (zie blz. 117)

$$||x||^2 = ||x - \sum_{k=1}^{n} \lambda_k e_{i_k}||^2 + ||\sum_{k=1}^{n} \lambda_k e_{i_k}||^2$$

$$= ||x - \sum_{k=1}^{n} \lambda_k e_{i_k}||^2 + \sum_{k=1}^{n} |\lambda_k|^2 \geq \sum_{k=1}^{n} |\lambda_k|^2 = \sum_{k=1}^{n} |(x,e_{i_k})|^2 .$$

Opmerkingen 1. Wanneer de ongelijkheid van Bessel overgaat in een gelijkheid, dan spreken we van de gelijkheid van Parceval. Hieronder zullen we zien dat dit geval slechts optreedt indien het stelsel $\{e_i | i \in I\}$ volledig is.

2. Als $\{e_i | i \in I\}$ een aftelbaar orthonormaal stelsel elementen is en $\{i_1, i_2, \ldots\}$ is een aftelling van I dan convergeert $\sum_{k=1}^{\infty} |x, e_{i_k}|^2$ voor alle $x \in H$ en de limiet is o.a. van de aftelling van I (absolute convergentie).

Zij nu $\{e_i | i \in I\}$ een orthonormaal stelsel elementen van H en $x \in H$ omdat $\sum_{i \in I} |(x,e_i)|^2 < \infty$ volgt dat $(x,e_i) \neq 0$ voor hoogstens aftelbaar veel $i \in I$. Immers voor elke $n = 1,2,\ldots$ is de collectie $\{i \in I | \, |(x,e_i)| > \frac{1}{n}\}$ eindig.

Er bestaat dus een aftelbare deelverzameling $I^* \subset I$ zodat voor $i \in I \backslash I^*$ $(x,e_i) = 0$; dus $\sum_{i \in I} |(x,e_i)|^2 = \sum_{i \in I^*} |(x,e_i)|^2$.

Laat nu $i_1, i_2, \ldots, i_n, \ldots$ een aftelling zijn voor I^*.

Voor elke n zij $s_n \in H$ gedefinieerd door $s_n = \sum_{k=1}^{n} (x,e_{i_k}) e_{i_k}$.

We zullen aantonen dat $\lim_{n \to \infty} s_n$ in H bestaat en we noteren dan deze limiet $\bar{x}$ met

$$\bar{x} = \sum_{i \in I} (x,e_i) e_i \quad .$$

Omdat de ruimte H met norm $|| \quad ||$ volledig is is het voldoende te laten zien dat

$$||s_n - s_m|| < \varepsilon \qquad \text{voor } n,m > n_0 \quad .$$

Nu geldt $||s_n - s_m||^2 = || \sum_{k=m+1}^{1} (x,e_{i_k}) e_{i_k} ||^2 = \sum_{k=m+1}^{n} |(x,e_{i_k})|^2$ (Pythagoras).

Uit opmerking 2 volgt dat het $\sum_{k=1}^{\infty} |(x,e_{i_k})|^2$ convergeert, dus het rechterlid is $< \varepsilon$ voor $n,m > n_0$.

N.B. Merk op dat de waarde van $\bar{x}$ niet afhangt van de aftelling van I^* en van I^* zelf.

Zij nu A een n-dimensionale vectorruimte en (,) een (inwendig product op A. Uit de vorige stelling volgt nu dat er een orthonomaal stelsel vectoren $(e_1,\ldots,e_n)$ (m.b.t. dit inproduct) dat tegelijk een basis is voor A. Als x een vector is uit A, dan geldt

$$(*) \qquad x = \sum_{i=1}^{n} (x,e_i)e_i \quad .$$

Immers x is te schrijven als lineaire combinatie der e_i's :
$x = x_1e_1+\ldots+x_ne_n$ en uit $(x,e_i) = (x_1e_1) = (x_1e_1+\ldots+x_ne_n;e_i) = x_i$ volgt het gestelde.

Blijkbaar is nu aan de gelijkheid van Parseval voldaan want n× toepassing van de stelling van Pythagoras geeft direct

$$(**) \qquad ||x||^2 = \sum_{i=1}^{n} ||(x,e_i)e_i||^2 = \sum_{i=1}^{n} |x,e_i|^2 \quad .$$

Tenslotte merken we nog op dat voor $x = x_1e_1+\ldots+x_ne_n$ het inproduct (x,y) gegeven is door de formule

$$(***) \qquad (x,y) = \sum_{i=1}^{n} (x_ie_i)(y,e_i)$$

Hieronder willen we nu onderzoeken of in een willekeurige Hilbertruimte een geschikt orthonormaal stelsel elementen bestaat zodat formules analoog aan (*), (**) en (***) voor het oneindig dimensionale geval bestaan.

<u>Definitie</u>. Een orthonormaal stelsel elementen $\{e_i | i \in I\} = D$ van een Hilbertruimte H heet <u>maximaal</u> indien D niet uit te breiden is tot een groter (i.e. echt omvattend) orthonormaal stelsel elementen van H.

Anders gezegd : $\{e_i | i \in I\}$ is maximaal $\Longleftrightarrow$ als $(x,e_i) = 0$ voor alle i dan $x = 0$.

<u>Nu geldt de volgende stelling.</u>

<u>Stelling 8</u>. Zij $\{e_i | i \in I\}$ een orthonormaal stelsel elementen van de Hilbertruimte H. Dan zijn de volgende voorwaarden equivalent:

i) $\{e_i | i \in I\}$ is maximaal

ii) Voor elke $x \in H$ geldt $x = \sum_{i \in I} (x,e_i)e_i$

iii) Voor elke $x \in H$ geldt $||x||^2 = \sum_{i \in I} |(x,e_i)|^2$ (gelijkheid van Parseval).

<u>Bewijs</u>. i$\Rightarrow$ii. Stel $\bar{x} = \sum_{i \in I} (x,e_i)e_i$. Als $I^* = \{i \in I | (x,e_i) \neq 0\}$ dan $I^* = i_1, i_2, \ldots$ en geldt $\bar{x} = \sum_{k=1}^{\infty} (x,e_{i_k})e_{i_k}$. Zij nu $i \in I$. Als $i \in I^*$ dan $i = i_s$ voor zeker natuurlijk getal s. Dus voor $n \geq s$ geldt

$$(1) \qquad x - \sum_{k=1}^{n} (x,e_{i_k})e_{i_k} \perp e_i$$

omdat $\bar{x}$ limiet is van bovenstaande sommen voor $n \geq s$ volgt uit de continuiniteit van het inproduct $x-\bar{x} \perp e_i$. Wanneer $i \notin I^*$ dan geldt (1) zelfs voor elke n, dus ook $x-\bar{x} \perp e_i$. Uit de maximaliteit van $\{e_i | i \in I\}$ volgt nu $x-\bar{x} = 0$.
ii $\Rightarrow$ iii. Zij $I^* = \{i \in I | (x,e_i) \neq 0\}$; $I^* = \{i_1, i_2, i_3, \ldots\}$. Dus $x = \lim_{n\to\infty} \sum_{k=1}^{n} (x,e_{i_k})e_{i_k}$. Uit de continuïteit van de norm volgt nu

$$||x||^2 = \lim_{n\to\infty} || \sum_{k=1}^{n} (x,e_{i_k})e_{i_k} ||^2 = (\text{Pythagoras}) = \lim_{n\to\infty} \sum_{k=1}^{n} |(x,e_{i_k})|^2$$

$$= \sum_{i \in I} |(x,e_i)|^2 .$$

iii)$\Rightarrow$i) Triviaal.

<u>Definitie</u>. Zij H een Hilbertruimte en $\{e_i | i \in I\}$ een orthonormaal stelsel elementen. Dan heet het stelsel $\{e_i | i \in I\}$ <u>volledig</u>, indien het lineaire omhulsel der e_i's dicht ligt in H. Anders gezegd: $\{e_i | i \in I\}$ is volledig d.e.s.d. indien $\forall \varepsilon > 0 \ \exists i_1, \ldots, i_n \in I$ en $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ (of $\mathbb{R}$) zodat $||x - \sum_{k=1}^{n} \lambda_k e_{i_k}|| < \varepsilon$.

Indien H een separabele Hilbert ruimte is d.w.z. er bestaat een aftelbare dicht liggende deelverzameling (zie pagina 37) dan betekent dat dat er aftelbaar veel elementen $x_1, x_2, \ldots, \ldots$ bestaan zodat $\{x_1, x_2, \ldots, \ldots\}$ dicht licht in H en à fortiori het lineair omhulsel der x_k's dicht ligt in H. Door orthogonalisering verkrijgen we zodoende een aftelbaar volledig orthonormaal stelsel in H.
Dus:

Een separabele Hilbertruimte bezit een aftelbaar volledig orthonormaal stelsel elementen.
Nu geldt de volgende stelling:

Stelling 9. Zij $\{e_i | i \in I\}$ een orthonormaal stelsel elementen van een Hilbertruimte H, dan zijn de volgende voorwaarden equivalent

1) $\{e_i | i \in I\}$ is een volledig stelsel
2) $\{e_i | i \in I\}$ is maximaal.

Bewijs. 1) $\Longrightarrow$ 2). Onderstel $\{e_i | i \in I\}$ is een volledig stelsel. Zij $x \in H$ en $(x,e_i) = 0 \quad \forall i \in I$. Als $\varepsilon > 0$ dan bestaan $\lambda_1, \lambda_2, \ldots, \lambda_n \in \mathbb{C}$ (of $\mathbb{R}$) en $i_1, i_2, \ldots, i_n \in I$ zodat voor $x_n = \sum_{k=1}^{n} \lambda_k e_{i_k}$ geldt

$||x-x_n||^2 < \varepsilon$. Blijkbaar $x \perp x_n$ dus $||x-x_n||^2 = ||x|| + ||x_n||^2$ volgens Pythagoras. Dus $||x||^2 < \varepsilon$ i.e. $x = 0$.

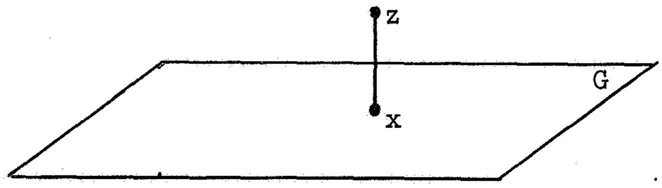
2) $\Longrightarrow$ 1). Zij $x \in H$ vast. Uit stelling 8 volgt dat $x = \sum_{i \in I} (x,e_i)e_i$.

Als $I^* = \{i \in I | (x,e_i) \neq 0\}$ en $I^* = i_1, i_2, \ldots, \ldots$ dan volgt

$x = \lim_{n \to \infty} \sum_{k=1}^{n} (x,e_{i_k})e_{i_k}$. d.w.z. voor $\varepsilon > 0 \ \exists n_0$ zodat $||x - \sum_{k=1}^{n_0} (x,e_{i_k})e_{i_k}||$ $< \varepsilon$. Omdat deze som in het lineaire omhulsel der ε_i's ligt volgt het gestelde.

Een zeer interessante aangelegenheid bij Hilbertruimten is het z.g. projecteren van een punt op een gesloten lineaire deelruimte. (Merk op dat elke lineaire deelruimte van een vectorruimte met inproduct (,) gesloten is).

<u>Stelling 10</u>. Zij G een gesloten lineaire deelruimte van een Hilbertruimte H en $z \in H\backslash G$, dan bestaat er precies één element $x \in G$ zodat $||z-x|| = \inf\{||z-y|| \mid y \in G\}$. Er geldt $(z-x) \perp G$; x heet de <u>projectie</u> van z op G



x heet de <u>projectie</u> van z op G $\qquad$ $||z-x||$ heet de <u>afstand</u> van z tot G.

<u>Bewijs</u>. Stel $d = \inf\{||z-y|| \mid y \in G\}$.
Er bestaat een rij punten $x_n \in G$ zodat $\lim_{n\to\infty} ||z-x_n|| = d$. We zullen eerst aantonen dat dan $\{x_n\}_{n=1}^{\infty}$ een fundamentaalrij is.
De parallelogramwet geeft voor $n,m \in \mathbb{N}$

$$||\frac{z-x_n}{2} + \frac{z-x_m}{2}||^2 + ||\frac{z-x_m}{2} - \frac{z-x_n}{2}||^2 =$$

$$= 2(||\frac{z-x_n}{2}||^2 + ||\frac{z-x_m}{2}||^2) \qquad \text{ofwel}$$

$$||z - \frac{x_n+x_m}{2}||^2 + \tfrac{1}{4}||x_n-x_m||^2 = \tfrac{1}{2}(||z-x_n||^2 + ||z-x_m||^2).$$

Zij $\eta > 0$ willekeurig, dan $||z-x_n||$ en $||z-x_m|| < d+\eta$ voor $n,m > N_0$.
Omdat $\tfrac{1}{2}(x_n+x_m) \in G$ volgt $||z-\tfrac{1}{2}(x_n+x_m)|| \geq d$.

Dus $||x_n-x_m||^2 = 2||z-x_n||^2 + 2||z-x_m||^2 - 4||z-\tfrac{1}{2}(x_n+x_m)||^2$

$< 2(d+\eta)^2 + 2(d+\eta)^2 - 4d^2 = 8d\eta + 4d\eta^2$ voor $n,m > N_0$.
Hieruit volgt dat $\{x_n\}_{n=1}^{\infty}$ een fundamentaalrij is in H. Omdat $(H, ||\ ||)$ volledig is bestaat nu $x = \lim_{n\to\infty} x_n$. $x \in G$ omdat G gesloten.
Nu geldt inderdaad $||z-x|| = d$.

Nu nog aan te tonen $z-x \perp G$. Zij $y \in G$ willekeurig. Beschouw $Q(\tau) =$
$= ||z-x-\tau y||^2 - ||z-x||^2$, $\tau \in \mathbb{C}$.
Blijkbaar $Q(\tau) \geq 0$ voor alle τ, want $x-\tau y \in G$ voor alle τ.
$Q(\tau) = |\tau|^2 \, ||y||^2 - \tau(z-x,y) - \tau(y,z-x)$.
Stel $(z-x,y) \neq 0$. Zij $\tau = \frac{(z-x,y)}{|(z-x,y)|} t$ $\quad$ t reëel.

Dan volgt $Q(\tau) = t^2||y||^2 - 2t|(z-x,y)| \geq 0$ voor alle t. Contradictie.

Éénduidigheid. Stel $x_1, x_2 \in G$ $\quad ||z-x_1|| = d, ||z-x_2|| = d$.
Blijkbaar volgens het bovenstaande $z-x_1 \perp G$ $\quad z-x_2 \perp G$; volgens Pythagoras: $||z-x_1||^2 = ||z-x_2||^2 + ||x_1-x_2||^2$, omdat $||z-x_1||^2 =$
$= ||z-x_2||^2 = d^2$ volgt $||x_1-x_2||^2 = 0$ i.e. $x_1 = x_2$.

Stelling 11. Zij $\{e_i | i \in I\}$ een orthonormaal stelsel elementen van een Hilbertruimte H en L het lineaire omhulsel der e_i's.
Als $x \in H$ en x' is de projectie van x op de afsluiting I van L dan geldt

$$x' = \sum_{i \in I} (x,e_i)e_i \quad .$$

Bewijsschets. Er geldt weer (continuiteit van inproduct)
$x-\bar{x} \perp G$, als $\bar{x} = \sum_{i \in I} (x,e_i)e_i$. Omdat blijkbaar $\bar{x} \in G$ en dus ook
$x'-\bar{x} \in G$ volgt uit de Stelling van Pythagoras.

$$||x-x'||^2 = ||x-\bar{x}||^2 + ||x'-\bar{x}||^2.$$

Omdat $||x-x'|| \leq ||x-\bar{x}||$ ($||x-x'||$ is kortste afstand) geldt $x' = \bar{x}$.

Voorbeeld. Beschouw de Hilbertruimte uit Voorbeeld 3. Het stelsel $\{e^{2\pi inx} | n=1,2,\ldots\}$ ($x \in [0,1]$ is een volledig orthonormaal stelsel.
Als f een complexwaardige integreerbare functie is op $[0,1]$ dan geldt
$f(x) = \sum_{n=1}^{\infty} a_n e^{2\pi inx}$ met a_n de n'e Fouriercoefficient van f(x):

$$a_n = \int_0^1 f(x)\, e^{-2\pi inx} \quad .$$

Dit is de toepassing van deze theorie op de Fourierreeksen (zie Computer wiskunde deel II).

§6 Het inproduct bij vectorruimten; toepassingen.

Eerst beschouwen we $\mathbb{C}^n$, de numerieke n-dimensionale complexe getallenruimte. Als gewoonlijk geven we de elementen aan door kolommen getallen uit $\mathbb{C}$.

Het gewone inproduct van twee vectoren $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ en $y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ wordt gedefinieerd door

$$(x,y) = x_1\bar{y}_1 + x_2\bar{y}_2 + \ldots + x_n\bar{y}_n \quad .$$

M.a.w. (x,y) is het matrixproduct $x^T\bar{y}$.
Hierbij is x^T de getransponeerde matrix $(x_1,\ldots,x_n)$ en $\bar{y}$ is de toegevoegd complexe matrix $\begin{pmatrix} \bar{y}_1 \\ \vdots \\ \bar{y}_n \end{pmatrix}$.

Zij nu V een willekeurige n-dimensionale complexe vectorruimte en $a_1,\ldots,a_n$ een basis. Als $x = x_1a_1+\ldots+x_na_n$ en $y = y_1a_1+\ldots+y_na_n$ dan kunnen we (x,y) definieren door

$$(x,y) = x_1\bar{y}_1 + \ldots + x_n\bar{y}_n \ .$$

$(\, , \,)$ wordt dan een inproduct en $a_1,\ldots,a_n$ een orthonormale basis.
Omgekeerd, als $(\, , \,)$ een inproduct is op V dan bestaat er m.b.t. $(\, , \,)$ een orthonormale basis $a_1,\ldots,a_n$ en (x,y) wordt met behulp van bovenstaande formules berekent.

Dus een inproduct in een complexe vectorruimte is volledig analoog aan het gewone inproduct in $\mathbb{C}^n$.

In de numerieke n-dimensionale reële getallenruimte definieren we het gewone inproduct analoog met bovenstaande formule. Omdat steeds geldt $\bar{y}_1 = y_1$; ..., $\bar{y}_n = y_n$ volgt $(x,y) = x^Ty$.
Een inproduct in een reële vectorruimte is op dezelfde manier bepaald als in een complexe vectorruimte.

<u>Definitie</u>. Zij V een complexe (reële) vectorruimte en (,) een inproduct. Een lineaire transformatie σ: V $\to$ V heet <u>unitair</u> (<u>orthogonaal</u>) indien voor alle $x,y \in V$

(2) $$(\sigma x, \sigma y) = (x,y)$$

(3) I.h.b. volgt $$||\sigma x - \sigma y|| = ||x-y||$$

m.a.w. een unitaire (orthogonale) transformatie laat afstanden invariant. Merk op dat overigens de formules (2) en (3) equivalent zijn.

Ook geldt de volgende stelling:

<u>Stelling 1</u>. Als V een complexe vectorruimte is met inproduct dan geldt als $a_1,\dots,a_n$ een orthonormale basis is ; σ: V $\to$ V een lineaire afbeelding

σ is unitair $\Longleftrightarrow$ $\sigma a_1,\dots,\sigma a_n$ is een orthonormale basis voor V.

Het analogon geldt voor een reële vectorruimte.

<u>Voorbeelden</u>. 1) Definieer S: $\mathbb{R}^2 \to \mathbb{R}^2$ door

$$Sx = y \quad \text{met} \quad \begin{cases} y_1 = x_1 \cos\phi - x_2 \sin\phi \\ y_2 = x_1 \sin\phi + x_2 \cos\phi \end{cases} .$$

S is een orthogonale trafo van $\mathbb{R}^2$ in $\mathbb{R}^2$ (rotatie om hoek ϕ).

2) Laat V een vectorruimte zijn met inproduct; $a_1,\dots,a_n$ een orthonormale basis en laat ϕ: V $\to$ V gedefinieerd zijn door $\phi a_i = \pm a_{\tau(i)}$ $i = 1,\dots,n$ waarbij τ een permutatie is van $\{1,\dots,n\}$ (dus verwisseling van basis elementen). Dan is ϕ orthogonaal.

I.h.b. is een spiegeling in de $\mathbb{R}^3$ t.o.v. een vlak door de oorsprong een orthogonale afbeelding.

<u>Definitie</u>. Een complexe n×n - matrix S heet <u>unitair</u> indien de bij S behorende transformatie S: $\mathbb{C}^n \to \mathbb{C}^n$ unitair is, relatief het gewone inproduct in $\mathbb{C}^n$.

Een unitaire matrix die slechts reële coëfficienten bezit heet <u>orthogonaal</u>. Duidelijk is dat S orthogonaal $\Longleftrightarrow$ S opgevat als S: $\mathbb{R}^n \to \mathbb{R}^n$ is orthogonaal.
Voorbeeld 1. De eenheidsmatrix is orthonaal.
Voorbeeld 2. Beschouw in de $\mathbb{R}^3$ een transformatie die de samenstelling is van een rotatie en een spiegeling. Dan wordt deze lineaire afbeelding beschreven met een orthogonale matrix.

<u>Stelling 2</u>. Zij V een complexe vectorruimte; (,) een inproduct en $a_1,\dots,a_n$ een $b_1,\dots,b_n$ orthonormale bases . Als σ: V → V lineair is, dan geldt

$$\sigma \text{ is unitair} \Longleftrightarrow \text{de matrix van } \sigma \text{ t.o.v. } a_1,\dots,a_n \text{ en } b_1,\dots,b_n \text{ is unitair.}$$

<u>Bewijs</u>. Zij S de matrix van σ t.o.v. $a_1,\dots,a_n$ en $b_1,\dots,b_n$. Als $x = x_1a_1+\dots+x_na_n$ en $y = y_1a_1+\dots+y_na_n$ dan is $(x,y) = x_1\bar{y}_1+x_2\bar{y}_2+\dots+x_n\bar{y}_n$ en $(\sigma x,\sigma y)$ is ook gemakkelijk te bepalen want de i^e coordinaat t.o.v. $b_1,\dots,b_n$ van σx en σy vinden we door vermenigvuldiging van de i^e rij van S met de kolomvectoren

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \quad \text{respectievelijk} \quad \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} .$$

Dus $(\sigma x,\sigma y) = (Sx,Sy)$ waarbij de x en y rechts bovenstaande kolomvectoren voorstellen. Het gestelde volgt nu.

Een analoog resultaat verkrijgen we bij een reële vectorruimte.

<u>Gevolg</u>: Zij S de transformatie matrix behorende bij een coördinaten transformatie van basis $a_1,\dots,a_n$ op $b_1,\dots,b_n$. Dan geldt $b_1,\dots,b_n$ orthonormaal $\Longleftrightarrow$ S is unitair. Als $x_1,\dots,x_n$ de oude coördinaten zijn van x t.o.v. $a_1,\dots,a_n$ en $x_1',x_2',\dots,x_n'$ de nieuwe t.o.v. $b_1,\dots,b_n$ dan geldt als $b_1,\dots,b_n$ orthonormaal $(x,y) = x_1\bar{y}_1+\dots+x_n\bar{y}_n = x_1'\bar{y}_1'+\dots+x_n'\bar{y}_n'$.

Als A een complexe n × n matrix is dan heet de matrix B gedefinieerd door $b_{ij} = \overline{a_{ji}}$ de <u>hermitisch getransponeerde matrix van</u> A. Notatie A^H.

We voeren dit begrip in omdat voor een willekeurige complexe n × n matrix A geldt

$$(Ax,y) = (x,A^H y) \ .$$

<u>Bewijs</u>.

$$(Ax,y) = (Ax)^T\overline{y} = x^T A^T \overline{y}$$
$$(x,A^H y) = x^T \overline{A^H y} = x^T A^T \overline{y}.$$

N.B. Als A reële coëfficienten bezit, dan geldt voor het reële of complexe) inproduct

$$(Ax,y) = (x,A^T y).$$

<u>Stelling 3</u>. Zij S een complexe n × n matrix; dan geldt

S is unitair $\Longleftrightarrow SS^H = S^H S = I$ (d.w.z. $S^{-1} = S^H$).

En i.h.b. voor een reële n × n- matrix.

S is orthogonaal $\Longleftrightarrow SS^T = S^T S = I$ (d.w.z. $S^{-1} = S^T$).

<u>Bewijs</u>. S unitair impliceert $(Sx,Sy) = (x,y)$ voor alle $x,y \in \mathbb{C}^n$. Dus $(x,S^H Sy) = (x,y)$ en ook $(x,S^H Sy-y) = 0$ voor alle x en y. Nemen we y vast en $x = S^H Sy-y$ dan volgt $S^H Sy-y = 0$, i.e., $S^H Sy = y$ voor alle y. Dus $S^H S = I$.
Evenzo $SS^H = I$.
Is omgekeerd $S^H S = I$ dan volgt $(Sx,Sy) = (x,S^H Sy) = (x,y)$ d.w.z. S is orthogonaal.

<u>Gevolg</u>. Als S een reële of complexe unitaire matrix is dan geldt $SS^H = I$, dus

1) Het inwendig product van twee verschillende kolommen van S = nul
2) Het inwendig product van een kolom van S met zichzelf is 1.

Verder geldt $\quad \det S = \pm 1$.

<u>Voorbeelden</u>. 1. Zij S een 2 × 2 - matrix die orthogonaal is.
Dus

$$S = \begin{pmatrix} s_{11} & s_{12} \\ s_{21} & s_{22} \end{pmatrix}. \qquad S^{-1} \text{ is blijkbaar } \pm \begin{pmatrix} s_{22} & -s_{12} \\ -s_{21} & s_{11} \end{pmatrix}.$$

Omdat $S^{-1} = S^T$ volgt dat S een der volgende vormen heeft

$$\begin{pmatrix} -a & b \\ b & a \end{pmatrix} \text{ of } \begin{pmatrix} a & -b \\ +b & a \end{pmatrix} \text{ met } a^2+b^2 = 1. \text{ Dus als } a = \sin\phi,\ b = \cos\phi$$

dan

$$S = \begin{pmatrix} \cos\phi & -\sin\phi \\ +\sin\phi & \cos\phi \end{pmatrix} \text{ of } \quad S = \begin{pmatrix} -\cos\phi & \sin\phi \\ \sin\phi & \cos\phi \end{pmatrix}.$$

In het eerste geval zijn de eigenwaarden toegevoegd complex i.e. S stelt een rotatie voor over een hoek ϕ om de oorsprong.
In het tweede geval ($\det S = -1$) zijn de eigenwaarden ± 1 i.e. S is een spiegeling t.o.v. de lijn $x_1 \sin\phi + x_2 \cos\phi = x_2$.

2. Als σ een orthogonale transformatie is in de $\mathbb{R}^n$ $(n = 1,2,\ldots)$ dan hebben alle eigenwaarden van σ absolute waarde 1 (modulus!). Immers, als λ een eigenwaarde is van σ met eigenvector x, dan geldt
$(x,x) = (\sigma x,\sigma x) = (\lambda x,\lambda x) = \lambda\bar{\lambda}(x,x) = |\lambda|^2 (x,x)$, dus $|\lambda|^2 = 1$,
i.e. $|\lambda| = 1$.

3. Als σ een orthogonale transformatie is van $\mathbb{R}^n$ en n is oneven, dan heeft σ een eigenwaarde + of -1. Immers, de karakteristieke v.g.l. van σ heeft een oneven graad en uit de algebra is bekend dat zo'n vergelijking altijd een reële oplossing bezit.
Ook geldt $\det \sigma = +1 \Longrightarrow \exists$ eigenwaarde +1; $\det \sigma = -1 \Longrightarrow \exists$ eigenwaarde -1.

4. Zij σ een orthogonale transformatie van $\mathbb{R}^3$. Kies volgens 3 een eigenwaarde ± 1 al naar gelang $\det \sigma = \pm 1$ zeg met eigenvector b. Vul b aan met vectoren b_1 en b_2 uit R^3 tot een orthonormale basis b, b_1, b_2 voor $\mathbb{R}^3$.
De matrix van σ t.o.v. b, b_1, b_2 heeft nu de volgende vorm

$$T = \begin{pmatrix} \pm 1 & s_{12} & s_{13} \\ 0 & s_{22} & s_{23} \\ 0 & s_{32} & s_{33} \end{pmatrix} .$$

Omdat T kennelijk orthogonaal is volgt (zie het gevolg hierboven) $s_{12} = s_{13} = 0$. Ook is de (2×2) matrix

$\begin{pmatrix} s_{22} & s_{23} \\ s_{32} & s_{33} \end{pmatrix}$ orthogonaal met determinant 1 en heeft dus de vorm $\begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix}$

Dus

$$T = \begin{pmatrix} \pm 1 & 0 & \\ 0 & \cos\phi & -\sin\phi \\ 0 & \sin\phi & \cos\phi \end{pmatrix} = \underbrace{\begin{pmatrix} \pm 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}}_{A} \underbrace{\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos\phi & -\sin\phi \\ 0 & \sin\phi & \cos\phi \end{pmatrix}}_{B} .$$

B is de matrix van de rotatieafbeelding in $\mathbb{R}^3$ om de as langs de eigenvector b over de hoek ϕ en A is de matrix van óf de identiteit (als det $\sigma=1$) óf van de spiegeling der vectoren van $\mathbb{R}^3$ t.o.v. het vlak opgespannen door b_1 en b_2. <u>We hebben dus bewezen</u>: Als $\sigma: \mathbb{R}^3 \to \mathbb{R}^3$ is orthogonaal dan det $\sigma=1$ impliceert dat σ simpelweg een rotatie is om een één dimensionale deelruimte van $\mathbb{R}^3$ en det $\sigma=-1$ impliceert dat σ is samengesteld uit zo'n rotatie om een as en een spiegeling t.o.v. een vlak loodrecht op die as.

<u>Physisch betekent deze stelling nu het volgende</u>: De beweging in een tijdsinterval van een star lichaam in de ruimte geschiedt altijd volgens een translatiebeweging en een rotatie om een as.

<u>Definitie</u>. Een complexe n × n-matrix A heet <u>hermitisch</u> indien $A = A^H$. Een complexe n × n-matrix heet <u>symmetrisch</u> indien $A = A^T$. I.h.b. is een reële symmetrische matrix hermitisch.

Dus A is hermitisch $\quad (Ax,y) = (x,Ay)$ voor alle $x,y \in \mathbb{C}^n$.

Hieronder onderzoeken we een paar eigenschappen van reële symmetrische matrices. Het zal blijken dat deze belangrijk zijn bij het bestuderen van kwadratische oppervlakken in de Analytische meetkunde.

Stelling 4. 1) De eigenwaarden van een reële symmetrische matrix A zijn reëel. 2) Twee eigenvectoren behorende bij verschillende eigenwaarden staan loodrecht op elkaar.

Bewijs. 1) Beschouw A als lineaire transformatie in $\mathbb{C}^n$. Omdat A reëel is geldt $A = A^H$; dus als x eigenvector is van A met bijbehorende eigenvector λ, dan geldt: $(Ax,x) = (\lambda x,x) = (x,Ax) = (x,\lambda x) = \bar{\lambda}(x,x)$. Dus $\lambda = \bar{\lambda}$ d.w.z. λ is reëel.

2) Als x en y eigenvectoren zijn van A bij verschillende eigenwaarden λ en μ dan volgt: $\lambda(x,y) = (Ax,y) = (x,Ay) = (x,\mu y) = \bar{\mu}(x,y) = \mu(x,y)$. Omdat $\lambda \neq \mu$ volgt $(x,y) = 0$.

Gevolg. Als A reëel en symmetrisch is en we beschouwen A als transformatie $A: \mathbb{R}^n \to \mathbb{R}^n$ dan zijn er n reële eigenwaarden en n bijbehorende eigenvectoren, die hiermee corresponderen.

Voor symmetrische matrices geldt verder nog de volgende belangrijke stelling:

Stelling 5. Als A een reële symmetrische $n \times n$-matrix is, dan bestaat er een reële orthogonale $n \times n$-matrix S zodat $S^{-1}AS$ de diagonaalvorm heeft.

Bewijs. Volledige inductie naar n.

We merken allereerst op dat als $S^{-1}AS = D$ dan is

$$A \begin{pmatrix} s_{11} & \cdots & s_{1n} \\ \vdots & & \\ s_{n1} & \cdots & s_{nn} \end{pmatrix} = \begin{pmatrix} s_{11} & \cdots & s_{1n} \\ \vdots & & \\ s_{n1} & \cdots & s_{nn} \end{pmatrix} \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} \qquad \text{zodat}$$

$$A \begin{pmatrix} s_{1i} \\ \vdots \\ s_{ni} \end{pmatrix} = \lambda_i \begin{pmatrix} s_{1i} \\ \vdots \\ s_{n1} \end{pmatrix}$$. Dus in dat geval bestaat er dus

een orthonomaal stelsel van n eigenvectoren van A n.l.

$$v_1 = \begin{pmatrix} s_{11} \\ \vdots \\ s_{n1} \end{pmatrix} ,\ldots\ldots, \quad v_n = \begin{pmatrix} s_{1n} \\ \vdots \\ s_{nn} \end{pmatrix}$$ terwijl $\lambda_1,\ldots,\lambda_n$ de bijbehorende

eigenwaarden zijn.

We kunnen dus ons probleem oplossen door eigenvectoren als basisvectoren te kiezen.

Voor 1×1 matrices is dit nu triviaal. Zij nu A een $n \times n$ matrix. Kies daarvan een eigenvector b_1 met eigenwaarde λ_1 (merk op dat λ_1 reëel en b_1 reëel). We mogen aannemen dat $|b_1| = 1$. Vul nu aan tot een orthonormale basis $b_1,\ldots,b_n$ voor $\mathbb{R}^n$. We passen coördinaten-transformatie toe naar deze nieuwe basis. De matrix wordt dan $S^{-1}AS$ met een orthogonale S. De nieuwe coördinaten van $|b_1|$ zijn $(1,0,\ldots,0)$ zodat

$$S \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = b_1 \qquad AS \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = Ab_1 = \lambda_1 b_1 = \lambda_1 S \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} .$$

We concluderen nu

$$(S^{-1}AS) \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \lambda_1 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$ zodat $S^{-1}AS$ de volgende vorm

heeft

$$S^{-1}AS = \begin{pmatrix} \lambda_1 & p_{12} & \cdots & p_{1n} \\ 0 & \vdots & & \vdots \\ \vdots & \vdots & & \vdots \\ 0 & p_{n2} & & p_{nn} \end{pmatrix} .$$

Daar S orthogonaal is, is $S^{-1}AS = S^TAS$, dus $S^{-1}AS$ is symmetrisch, dus $p_{12} = \ldots = p_{1n} = 0$, en de $(n-1) \times (n-1)$ matrix rechts onderaan is symmetrisch.

Blijkens de inductie aanname is er nu een orthogonale matrix t_{ij} die deze in de diagonaalvorm transformeert

$$\begin{pmatrix} t_{22} & \cdots & t_{2n} \\ \vdots & & \\ t_{n2} & \cdots & t_{nn} \end{pmatrix}^T \begin{pmatrix} p_{22} & \cdots & p_{2n} \\ \vdots & & \\ p_{n2} & \cdots & p_{nn} \end{pmatrix} \begin{pmatrix} t_{22} & \cdots & t_{2n} \\ \vdots & & \\ t_{n2} & \cdots & t_{nn} \end{pmatrix} = \begin{pmatrix} \lambda_2 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}.$$

Nu is ook

$$\begin{pmatrix} 1 & 0 & 0 & 0 \ldots 0 \\ 0 & t_{22} & \cdots & t_{2n} \\ \vdots & & & \\ 0 & t_{n2} & \cdots & t_{nn} \end{pmatrix}^T \begin{pmatrix} \lambda_1 & 0 & 0 \ldots 0 \\ \vdots & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & & \\ 0 & p_{n2} & & p_{nn} \end{pmatrix} \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & t_{22} & \cdots & t_{2n} \\ \vdots & & & \\ 0 & t_{n2} & \cdots & t_{nn} \end{pmatrix} = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix}.$$

De derde matrix in het linkerlid is weer een orthogonale; stel die door S_1 voor. We zien nu dat $(SS_1)^T A(SS_1)$ diagonaal is; aangezien ook SS_1 diagonaal is, is hiermee de inductiestap voltooid.

Opmerking. In de praktijk zal men niet de bovengeschetste methode volgen. Uit het eindresultaat blijkt n.l. als λ een k-voudige wortel van de karakteristieke vergelijking is, dan is er een orthonormaal stel eigenvectoren met eigenwaarden λ. Deze voldoen aan de v.g.l. $(A-\lambda I)x = 0$. In de oplossingsruimte daarvan heeft men een orthonormale basis te kiezen. Doen we dit voor iedere eigenwaarde van A_1 dan vinden we zodoende gemakkelijk een orthonormale basis van eigenvectoren van A. De matrix S verkrijgt men door de eigenvectoren als kolommen achter elkaar te zetten.

Toepassing : Kwadratische vormen.

Definitie. Een kwadratische vorm is een homogene veelterm van de graad 2 in een aantal veranderlijken $x_1,\ldots,x_n$.

De <u>matrix</u> van de kwadratische vorm in $x_1,\ldots,x_n$ is de $n \times n$ matrix (a_{ij}) met

$$a_{ii} = \text{coëfficient van de } x_i^2$$
$$2a_{ij} = \text{coëfficient van } x_i x_j \quad (\text{als } x_i x_j)\ .$$

Voorbeeld $\quad x_1^2 - 2x_1x_3 + 4x_2x_3 - 4x_3^2$ heeft de matrix $\begin{pmatrix} 1 & 0 & -1 \\ 0 & 0 & 2 \\ -1 & 2 & -4 \end{pmatrix}$.

<u>Stelling 6</u>. Is $(a_{ij}) = A$ de matrix van een kwadratische vorm, dan is de vorm

$$\sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}x_ix_j = x^T Ax = (x,Ax) = (Ax,x).$$

Verder is de matrix A symmetrisch !!

In een vectorruimte V zij nu een basis $b_1,\ldots,b_n$ gegeven. Zij verder A een symmetrische $n \times n$ matrix en p een reëel getal. We bekijken nu de verzameling C van alle vectoren $\bar{x} \in V$ met $\bar{x} = x_1b_1+\ldots+x_nb_n$ zodat voor de kolomvector $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ geldt:

$$(x,Ax) = p \qquad \text{ofwel} \qquad \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}x_ix_j = p \quad \ldots\ldots \quad (1)$$

Wanneer we dit in de stereometrische ruimte doen vormen de uiteinden van de genoemde vectoren een z.g. kwadratisch oppervlak met middelpunt O. I.h.a nemen we voor $b_1,\ldots,b_n$ een <u>orthonormale</u> basis m.a.w. we beschrijven zo'n kwadratisch oppervlak d.m.v. een rechthoekig assenkruis.

Als $s_1,\ldots,s_n$ een lineair onafhankelijk stel eigenvectoren is van de matrix A en $\lambda_1,\lambda_2,\ldots,\lambda_n$ de bijbehorende eigenwaarden dan kunnen we de vergelijking (1) ook schrijven als

$$\sum_{i=1}^{n} \lambda_i (x,s_i)^2 = p \quad \ldots\ldots \quad (2)$$

Hierbij is (x,s_i) het inproduct van de kolomvector $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$

en de eigenvector s_i van de matrix A.

Immers $x = \sum_{i=1}^{n} (x,s_i)s_i$ dus

$$(Ax,x) = (\sum_{i=1}^{n} (x,s_i)As_i,x) = \sum_{i=1}^{n} \lambda_i(x,s_i)(s_i,x) =$$

$$= \sum_{i=1}^{n} \lambda_i(x,s_i)^2.$$

Stel voor $i = 1,2,\ldots,n$ $\quad s_i = \begin{pmatrix} s_{1i} \\ \vdots \\ s_{ni} \end{pmatrix}$ en beschouw de overgang van

van coördinaten $x_1,\ldots,x_n$ t.o.v. $b_1,\ldots,b_n$ op coördinaten $x'_1,\ldots,x'_n$ t.o.v. de nieuwe orthonormale basis $c_1,\ldots,c_n$ door

$$\begin{cases} x_1 = s_{11}x'_1 +\ldots s_{1n}x'_n \\ \vdots \\ x_n = s_{n1}x'_n +\ldots s_{nn}x'_n \end{cases}.$$

De vergelijking (2) gaat nu over in $\sum_{i=1}^{n} \lambda_i(x'_i)^2 = p$. Deze heet d.z.g. kanonieke vorm van (1). en hieraan kan men dikwijls de aard van het kwadratisch oppervlak bepalen.

APPENDIX Vectorrekening en enkele toepassingen.

We beschouwen hieronder vectoren uit $\mathbb{R}^3$; een vector $a \in \mathbb{R}^3$ schrijven we als $a = (a_1,a_2,a_3)$; $e_1 = (1,0,0)$ $e_2 = (0,1,0)$ $e_3 = (0,0,1)$.

Definitie. Onder het uitwendig product van twee vectoren $a = (a_1,a_2,a_3)$ en $b = (b_1,b_2,b_3)$ schrijfwijze $a \times b$ verstaan we de vector

$$\left(\begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix} , \begin{vmatrix} a_3 & a_1 \\ b_3 & b_1 \end{vmatrix} , \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} \right).$$

Formeel kunnen we dit onthouden door de volgende symbolische formule

$$a \times b = \begin{vmatrix} e_1 & e_2 & e_3 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{vmatrix}.$$ (Dan is het eventueel handig pijlen op de eenheidsvectoren te schrijven).

Eigenschappen 1): $e_2 \times e_3 = e_1$; $e_1 \times e_2 = e_3$; $e_3 \times e_1 = e_2$.

2) Kennelijk geldt $(a\times b,a) = (a\times b,b) = 0$ (Ga na), dus $a\times b \perp a$ en $a\times b \perp b$.

3) Oriëntatiekwestie van $a \times b$:

We definiëren voor een orthonormaalstelsel a,b,c in $\mathbb{R}^3$ de oriëntatie positief indien

$$D = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} > 0$$ en negatief indien $D < 0$.

Blijkbaar heeft e_1,e_2,e_3 positieve oriëntatie. Meetkundig betekent dat a,b,c positief georiënteerd is dat de richting van c door toepassing van de kurketrekkerregel op a en b bepaald is.

Wanneer $c = a\times b$ dan geldt $\begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} =$

$= c_1^2 + c_2^2 + c_3^2 > 0$.

Dus $(a,b,a\times b)$ is positief georiënteerd.

4) Zij O de oppervlakte van de driehoek ingesloten door de vectoren a en b. Dan geldt

$$2O = |a|.|b|.\ \sin(a,b) = |a|.|b|\sqrt{1-\cos^2(a,b)} =$$

$$= \sqrt{|a|^2|b|^2 - |a|^2|b|^2\cos^2\phi} =$$

$$= \sqrt{(a_1^2+a_2^2+a_3^2)(b_1^2+b_2^2+b_3^2) - (a_1b_1+a_2b_2+a_3b_3)^2} =$$

$$= \sqrt{(a_1b_2-a_2b_1)^2 + (a_1b_3-a_3b_1)^2 + (a_2b_3-a_3b_2)^2} = |a\times b|$$

5) Verdere Rekenregels:

a) $a\times b = -(b\times a)$

b) $a \times (b+c) = a\times b + a\times c$

c) $(b+c) \times a = b\times a + c\times a$

d) $(\lambda a) \times b = \lambda(a\times b) = a \times (\lambda b)$.

6) Verband van uitwendig product en inwendig product: zij $a,b,c \in \mathbb{R}^3$. Definieer

$$(a,b,c) = \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} .$$

Dan geldt $(a\times b,c) = (a,b\times c) = (a,b,c) = 6 \times$ Opp viervlak (a,b,c).

7) $a \times (b\times c) = (a,c)b - (a,b)c$.

Al deze rekenregels zijn gemakkelijk door uitschrijven te verifiëren.
Bewijs van 7): Stel $w = b\times c$ en $s = a\times w$

$$w = \left(\begin{vmatrix} b_2 & b_3 \\ c_2 & c_3 \end{vmatrix}, \begin{vmatrix} b_3 & b_1 \\ c_3 & c_1 \end{vmatrix}, \begin{vmatrix} b_1 & b_2 \\ c_1 & c_2 \end{vmatrix}\right)$$

$$s_1 = \begin{vmatrix} a_2 & a_3 \\ w_2 & w_3 \end{vmatrix} = a_2\begin{vmatrix} b_1 & b_2 \\ c_1 & c_2 \end{vmatrix} - a_3\begin{vmatrix} b_3 & b_1 \\ c_3 & c_1 \end{vmatrix} = b_1(a_2c_2+a_3c_3)-c_1(a_2b_2+a_3b_3)$$

$$= b_1(a_1c_1+a_2c_2+a_3c_3) - c_1(a_1b_1+a_2b_2+a_3b_3) = b_1(a,c) - c_1(a,b).$$

Evenzo $\quad s_2 = b_2(a,c) - c_2(a,b)$

$$s_3 = b_3(a,c) - c_3(a,b).$$

Hieruit volgt het gestelde.

Toepassingen in de Mechanica.

1) In de Mechanica komt het uitwendig product te voorschijn bij het bestuderen van rotatiebewegingen van systemen van massapunten (b.v. lichamen). Als we een systeem bestuderen bestaande uit de massapunten $m_1,\dots,m_n$ met radiele posities $r_1,\dots,r_n$; snelheden $v_1,\dots,v_n$, dan verstaat men onder het impulsmoment van het systeem de vector

$$J = \sum_i m_i(r_i \times v_i) .$$

De verandering per seconde van deze vector J i.e. de vectorlimiet $\frac{dJ}{dt} = \lim_{\Delta t \to 0} \frac{J(t+\Delta t)-J(t)}{\Delta t}$ wordt dan in verband gebracht met de uitwendige koppels die op de massapunten werken:

$$\frac{d}{dt} J = \sum m_i(\frac{d}{dt} r_i \times v_i) + \sum m_i(r_i \times \frac{d}{dt} v_i) =$$

$$= \sum (r_i \times F_{ui}) \quad \text{(volgens Newton).}$$

Hierin is F_{ui} de (uitwendige) kracht die op het massapunt m_i werkt en $r_i \times F_{ui}$ is het moment van die kracht t.o.v. de positievector r_i.

We zullen nu J berekenen voor het geval het systeem massapunten een stijf lichaam is. In dat geval is er over een tijdsinterval t, t+Δt sprake van een rotatie van het lichaam om een vaste as (zie ook blz. 132).
We nemen nu een positief geöriënteerd rechthoekig coördinatensysteem in $\mathbb{R}^3$ waarvan de oorsprong ligt op deze draaias. We noteren met w de vector die als richting de asrichting heeft van het systeem en als lengte de hoeksnelheid. Als r_i en v_i de posities en snelheden zijn van de massapunten m_i (t.o.v. dit coördinatensysteem) dan volgt blijkbaar wiskundig

$$v_i = w \times r_i .$$

Dus, als s_i de component is van r_i langs w en d_i de component van r_i loodrecht op w:

$$J = \sum_i m_i(r_i \times (w \times r_i)) = \sum_i m_i(d_i+s_i) \times (w \times (d_i+s_i))$$

$$= \sum_i m_i(d_i+s_i) \times (w \times d_i) = \sum_i m_i(d_i \times (w \times d_i) + s_i \times (w \times d_i))$$

$$= \sum_i m_i[(d_i,d_i)w - (d_i,w)d_i + (s_i,d_i)w - (s_i,w)d_i]$$

$$= \sum_i m_i|d_i|^2 w - \sum_i m_i(s_i,w)d_i .$$

In de meeste gevallen valt de richting van J samen met die van w, dus

$$J = \underbrace{(\sum m_i|d_i|^2)w}_{(1)} .$$

(1) heet wel het <u>traagheidsmoment</u> van het systeem om deze as; deze is in veel gevallen onafhankelijk van de tijd.

$|d_i|^2$ = het kwadraat van de afstand van het massapunt m_i tot de draaias.

2) De schijnbare kracht op een stoffelijk punt bij een roterend coördinatenstelsel met hoeksnelheidsvector w berekenen we door

$$mr = F - \underbrace{m[w \times [w \times r]]}_{\text{centrifugaal-kracht}} - \underbrace{2m(w \times r)}_{\text{corioliskracht}}$$

r is hier de positievector van het stoffelijk punt t.o.v. het roterende coördinatensysteem.

Literatuurlijst

Topologie

1. W. Franz, Topologie I, Sammlung Göschen
   (Zeer elementair beknopt leerboekje voor beginners).

2. W.J. Pervin, Foundations of general topology, Academic Press
   (Meer uitgebreid elementair leerboek).

3. J. Dugundi, Topology, Allyn and Bacon, 1966
   (Standaardwerk, voor gevorderden).

Lineaire algebra, matrix rekening

1. R. Bronson, Matrix methods, an introduction, Academic Press, New York 1969 (toepassingen matricesrekening).

2. H.J. Kowalsky, Lineaire Algebra, Berlin 1963
   (uitgebreid theoretisch werk over lineaire algebra).

3. P.R. Halmos, Finite dimensional vector spaces, Princeton, 1958.

Lineaire Algebra en Lineaire Analyse

1. G.F. Simmons, Introduction to topology and modern analysis, McGraw Hill 1963.
   (Mooi exact overzicht, beknopt, met geen toepassingen).

2. S.K. Berberian, Introduction to Hilbertspace, New York, 1961.

Differentiaal en Integraalrekening

1. R. Courant, Differential and Integral Calculus
   Blackie & Sons, London.
   (Standaardwerk, weinig exact, vele toepassingen).

2. T.M. Apostol, Mathematical Analysis, Addison Wesley 1957.