

TECHNISCHE HOGESCHOOL EINDHOVEN

Afdeling Algemene Wetenschappen

Onderafdeling der Wiskunde

WISKUNDE 20

bestemd voor

BDK-II, WSK-II, N-II, E-II en T-II

Voorjaarssemester 1975



Technische Hogeschool Eindhoven

Jo de Gooij

Onderafdeling der Wiskunde

Wiskunde 20

bestemd voor BDK-II, WSK-II, N-II, W-II, E-II en T-II

TECHNISCHE HOGESCHOOL EINDHOVEN

Onderafdeling der Wiskunde

WISKUNDE 20

bestemd voor

BDK-II, WSK-II, N-II, W-II, E-II en T-II

Voorjaarssemester 1975

INHOUDSOPGAVE WISKUNDE 20

Hoofdstuk 6.	Functies $\mathbb{R} \rightarrow \mathbb{R}^2$, $\mathbb{R} \rightarrow \mathbb{R}^3$ en functies $\mathbb{R}^2 \rightarrow \mathbb{R}$, $\mathbb{R}^3 \rightarrow \mathbb{R}$	1
6.1.	Functies $\mathbb{R} \rightarrow \mathbb{R}^2$, $\mathbb{R} \rightarrow \mathbb{R}^3$	1
6.2.	Functies van twee variabelen	5
6.3.	Differentieerbaarheid van functies van twee variabelen	9
6.4.	Functies van drie variabelen	21
6.5.	Impliciete functies	23
6.6.	Richtingsafgeleide en gradiënt	27
6.7.	Extrema	30
6.8.	Extrema onder nevenvoorwaarden	34
Hoofdstuk 7.	Lineaire algebra	36
7.1.	Voorwoord	36
7.2.	Bewerkingen met matrices	36
7.3.	Voorbeelden en toepassingen	48
7.4.	Analytische meetkunde in \mathbb{R}^3 en meetkundige terminologie in \mathbb{R}^n met $n > 3$	58
7.5.	Vectorruimten	65
7.6.	Afhankelijkheid en onafhankelijkheid, bases	70
7.7.	Lineaire afbeeldingen	85
7.8.	Lineaire afbeeldingen $\mathbb{R}^n \rightarrow \mathbb{R}^n$	93
7.9.	Lineaire vergelijkingen en ongelijkheden	106

Hoofdstuk 8. Meervoudige integralen	126
8.1. Integratiegebieden en infinitesimale elementen	126
8.2. Lijnintegralen	132
8.3. Oppervlakte-integralen	136
8.4. Volume-integralen	145
8.5. Oneigenlijke meervoudige integralen	151
8.6. Gammafunctie, betafunctie	157
8.7. Integralen met een parameter	163

Hoofdstuk 6. Functies $\mathbb{R} \rightarrow \mathbb{R}^2$, $\mathbb{R} \rightarrow \mathbb{R}^3$ en functies $\mathbb{R}^2 \rightarrow \mathbb{R}$, $\mathbb{R}^3 \rightarrow \mathbb{R}$

6.1. Functies $\mathbb{R} \rightarrow \mathbb{R}^2$, $\mathbb{R} \rightarrow \mathbb{R}^3$

Een vectorfunctie \underline{f} van één reële variabele is een afbeelding $\underline{f}: A \rightarrow \mathbb{R}^n$, waarbij $A \subset \mathbb{R}$. A heet de definitieverzameling van \underline{f} ; notatie $\text{DOM } \underline{f}$.

In deze paragraaf beperken we ons tot de gevallen $n = 2$ en $n = 3$. Een vectorfunctie \underline{f} is dan een afbeelding, waarbij aan elke $t \in \text{DOM } \underline{f} \subset \mathbb{R}$ een vector $\underline{f}(t) = (x(t), y(t)) \in \mathbb{R}^2$ resp. $\underline{f}(t) = (x(t), y(t), z(t)) \in \mathbb{R}^3$ wordt toegevoegd. Naar believen zullen we $(x(t), y(t))$ resp. $(x(t), y(t), z(t))$ ook opvatten als de cartesische coördinaten van een punt in \mathbb{R}^2 resp. \mathbb{R}^3 . De functies $x(t), y(t)$ resp. $x(t), y(t), z(t)$ heten de componentfuncties of componenten van $\underline{f}(t)$.

Voorbeelden. 1) $\underline{f}: [0, 2\pi) \rightarrow \mathbb{R}^2$, $\underline{f}(t) = (\cos t, \sin t)$.

2) $\underline{f}: \mathbb{R} \rightarrow \mathbb{R}^2$, $\underline{f}(t) = (a_1 t + b_1, a_2 t + b_2)$, $(a_1, a_2) \neq (0, 0)$.

3) $\underline{f}: \mathbb{R} \rightarrow \mathbb{R}^3$, $\underline{f}(t) = (\cos t, \sin t, t)$.

6.1.1. Parametervoorstelling van krommen. Kies in \mathbb{R}^2 resp. \mathbb{R}^3 het gebruikelijke assenstelsel. De puntverzameling

$$S := \{ \underline{f}(t) \mid t \in \text{DOM } \underline{f} \}$$

heet een kromme. Indien deze kromme in een vlak ligt spreken we van een vlakke kromme, anders van een ruimte-kromme.

De functie $\underline{f}(t)$ heet een parametervoorstelling van de kromme S , met parameter t .

De functie \underline{f} uit voorbeeld 1 is een parametervoorstelling van de cirkel met middelpunt $\underline{0}$ en straal 1. Immers de kromme bestaat uit alle punten $(x, y) \in \mathbb{R}^2$ met $x = \cos t$, $y = \sin t$ voor zekere $t \in [0, 2\pi)$. Na eliminatie van t volgt $x^2 + y^2 = 1$, d.i. de vergelijking van de cirkel met middelpunt $\underline{0}$ en straal 1. Dezelfde functie \underline{f} met $\text{DOM } \underline{f}$ beperkt tot $[0, \pi]$, is een parametervoorstelling van een halve cirkel met middelpunt $\underline{0}$ en straal 1. Deze halve cirkel is ook de grafiek van de reële functie $y = \sqrt{1-x^2}$. De parametervoorstelling voor de halve cirkel is dus te herleiden tot een voorstelling door middel van een reële functie. Voor de volledige cirkel als boven is een dergelijke herleiding niet mogelijk.

① Opgave. Toon aan, dat $\underline{f}(t) = (a + r \cos t, b + r \sin t)$ met $r > 0$ een parametervoorstelling is van de cirkel met middelpunt (a, b) en straal r .

④ De functie \underline{f} uit voorbeeld 2 is een parametervoorstelling van een rechte. Deze rechte wordt ook beschreven door de vergelijking $a_2x - a_1y = a_2b_1 - a_1b_2$, zoals volgt door eliminatie van t uit de vergelijkingen $x = a_1t + b_1$, $y = a_2t + b_2$.

⑤ De functie \underline{f} uit voorbeeld 3 is een parametervoorstelling van een zg. schroeflijn. Deze schroeflijn bestaat uit alle punten $(x, y, z) \in \mathbb{R}^3$ met $x = \cos t$, $y = \sin t$, $z = t$ voor zekere $t \in \mathbb{R}$. Na eliminatie van t volgt dat de schroeflijn ook beschreven wordt door de twee vergelijkingen $x = \cos z$, $y = \sin z$. Elimineer z , dan is duidelijk dat de schroeflijn ligt op de cilinder met vergelijking $x^2 + y^2 = 1$.

Voorbeeld 4. De functie $\underline{f} : [0, 2\pi) \rightarrow \mathbb{R}^2$ gegeven door $\underline{f}(t) = (a \cos t, b \sin t)$ met $a > 0$, $b > 0$, is een parametervoorstelling van een ellips. Immers de kromme bestaat uit alle punten $(x, y) \in \mathbb{R}^2$ met $x = a \cos t$, $y = b \sin t$ voor zekere $t \in [0, 2\pi)$. Na eliminatie van t volgt $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$, d.i. de vergelijking van een ellips.

Uit de parametervoorstelling volgt dat de ellips ontstaat door de y -coördinaten van de punten van de cirkel $x^2 + y^2 = a^2$ te vermenigvuldigen met b/a of ook door de x -coördinaten van de punten van de cirkel $x^2 + y^2 = b^2$ te vermenigvuldigen met a/b .

Voorbeeld 5. De baan van een massapunt is te beschrijven door een parametervoorstelling waarbij de tijd t als parameter wordt gebruikt.

Beschouw bijv. de beweging van een massapunt langs de cirkel met middelpunt O en straal r . Op het tijdstip $t = 0$ zal het massapunt zich bevinden in het punt $(r, 0)$. De beweging geschiedt met constante hoeksnelheid ω . De positie van het massapunt op het tijdstip t wordt dan gegeven door: $x = r \cos \omega t$, $y = r \sin \omega t$. De functie $\underline{f}(t) = (r \cos \omega t, r \sin \omega t)$ is een parametervoorstelling van de baan van het massapunt.

⑥ Indien we de omlooptijd $2\pi/\omega$ als tijdseenheid kiezen dan wordt de parametervoorstelling van de baan: $\underline{f}(\tau) = (r \cos 2\pi\tau, r \sin 2\pi\tau)$, met parameter $\tau = \omega t / 2\pi$.

We definiëren nu de begrippen continuïteit en differentieerbaarheid voor een vectorfunctie.

6.1.2. Definitie. Zij $\underline{f} : \text{DOM } \underline{f} \rightarrow \mathbb{R}^2$ gegeven door $\underline{f}(t) = (x(t), y(t))$.
 \underline{f} heet continu in $a \in \text{DOM } \underline{f}$ als de componenten x, y continu zijn in a .
 \underline{f} heet continu op een verzameling $V \subset \mathbb{R}$ als \underline{f} continu is in elk punt van V .
 \underline{f} heet continu als \underline{f} continu is op $\text{DOM } \underline{f}$.

6.1.3. Definitie. Zij $\underline{f} : \text{DOM } \underline{f} \rightarrow \mathbb{R}^2$ gegeven door $\underline{f}(t) = (x(t), y(t))$.
Zij a inwendig punt van $\text{DOM } \underline{f}$ dan heet \underline{f} differentieerbaar in a als de componenten x, y differentieerbaar zijn in a ; $\underline{f}'(a) := (x'(a), y'(a))$ heet de afgeleide van \underline{f} in a .
 \underline{f} heet differentieerbaar op een open interval $I \subset \mathbb{R}$ als \underline{f} differentieerbaar is in elk punt van I .
Is bovendien de afgeleide \underline{f}' continu op I dan heet \underline{f} continu differentieerbaar op I .

De definities van continuïteit en differentieerbaarheid voor een vectorfunctie $\underline{f} : \text{DOM } \underline{f} \rightarrow \mathbb{R}^3$ zijn analoog.

In plaats van $\underline{f}'(t)$, $x'(t)$, $y'(t)$ gebruikt men ook vaak de notaties $\dot{\underline{f}}(t)$, $\dot{x}(t)$, $\dot{y}(t)$.

6.1.4. Raaklijn. Beschouw een vectorfunctie $\underline{f} : \text{DOM } \underline{f} \rightarrow \mathbb{R}^3$ gegeven door $\underline{f}(t) = (x(t), y(t), z(t))$, waarbij $\text{DOM } \underline{f}$ een open interval is.
Laat \underline{f} differentieerbaar zijn in $a \in \text{DOM } \underline{f}$, dan zijn de componenten x, y, z differentieerbaar in a . Volgens 2.6.1 geldt dan:

$$x(a+h) = x(a) + \dot{x}(a)h + h\rho(h),$$

$$y(a+h) = y(a) + \dot{y}(a)h + h\sigma(h),$$

$$z(a+h) = z(a) + \dot{z}(a)h + h\tau(h),$$

voor alle h in een omgeving V van 0 , en

$$\lim_{h \rightarrow 0} \rho(h) = \lim_{h \rightarrow 0} \sigma(h) = \lim_{h \rightarrow 0} \tau(h) = 0.$$

In een omgeving van a is $\underline{f}(t)$ lineair benaderbaar door de functie

$$\underline{q}(t) := \underline{f}(a) + \dot{\underline{f}}(a)(t-a). \quad a+h=t \rightarrow h=t-a$$

Indien $\dot{\underline{f}}(a) \neq \underline{0}$ dan is \underline{q} een parameterrepresentatie van een rechte: de raaklijn aan de kromme in het punt $\underline{f}(a)$. De richtingsvector van deze raaklijn is $\dot{\underline{f}}(a) = (\dot{x}(a), \dot{y}(a), \dot{z}(a)) \neq \underline{0}$. Als $\dot{x}(a) = 0$ is de raaklijn evenwijdig aan het (y, z) -vlak; analoge uitspraken gelden indien $\dot{y}(a) = 0$ resp. $\dot{z}(a) = 0$.

Beschouw vervolgens een vlakke kromme gegeven door de parametervoorstelling $\underline{f}(t) = (x(t), y(t))$. Op dezelfde manier vinden we dan: als \underline{f} differentieerbaar is in a met afgeleide $\dot{\underline{f}}(a) = (\dot{x}(a), \dot{y}(a)) \neq \underline{0}$ dan heeft de kromme een raaklijn in het punt $\underline{f}(a)$ met richtingsvector $\dot{\underline{f}}(a)$. Als $\dot{x}(a) = 0$ is de raaklijn evenwijdig aan de y -as, als $\dot{y}(a) = 0$ is de raaklijn evenwijdig aan de x -as.

Zoals eerder opgemerkt is de parametervoorstelling van een vlakke kromme soms te herleiden tot een voorstelling door middel van een reële functie $g(x)$, i.e. de kromme is de grafiek van $y = g(x)$. We geven hierover de volgende stelling.

6.1.5. Stelling. Zij de vectorfunctie $\underline{f}(t) = (x(t), y(t))$ continu differentieerbaar op een omgeving van het punt a en zij $\dot{x}(a) \neq 0$.

Dan bestaat er een omgeving V van a en een reële functie $g(x)$ gedefinieerd in $x(V)$, zodat de kromme $\{\underline{f}(t) \mid t \in V\}$ de grafiek is van $y = g(x)$.

De afgeleide van $g(x)$ in $x(a)$ wordt gegeven door

$$\frac{dy}{dx}(x(a)) = \frac{\dot{y}(a)}{\dot{x}(a)}.$$

Bewijs. De afgeleide $\dot{x}(t)$ is continu op een omgeving van a en $\dot{x}(a) \neq 0$. Daaruit volgt dat er een omgeving V van a bestaat zodat $\dot{x}(t) > 0$ voor alle $t \in V$ of $\dot{x}(t) < 0$ voor alle $t \in V$. Volgens 2.6.8 is $x(t)$ dan strikt monotoon op V . De functie $x(t)$ heeft nu een inverse functie, aan te geven door $t(x)$, met afgeleide $\frac{dt}{dx}(x(a)) = \frac{1}{\dot{x}(a)}$.

Substitueer $t(x)$ in de component $y(t)$, dan ontstaat de samengestelde functie $y(t(x)) =: g(x)$ gedefinieerd in $x(V)$. Het is duidelijk dat de grafiek van $y = g(x)$ juist overeenstemt met de kromme $\{\underline{f}(t) \mid t \in V\}$.

De afgeleide van $g(x) = y(t(x))$ in $x(a)$ is te bepalen met de kettingregel:

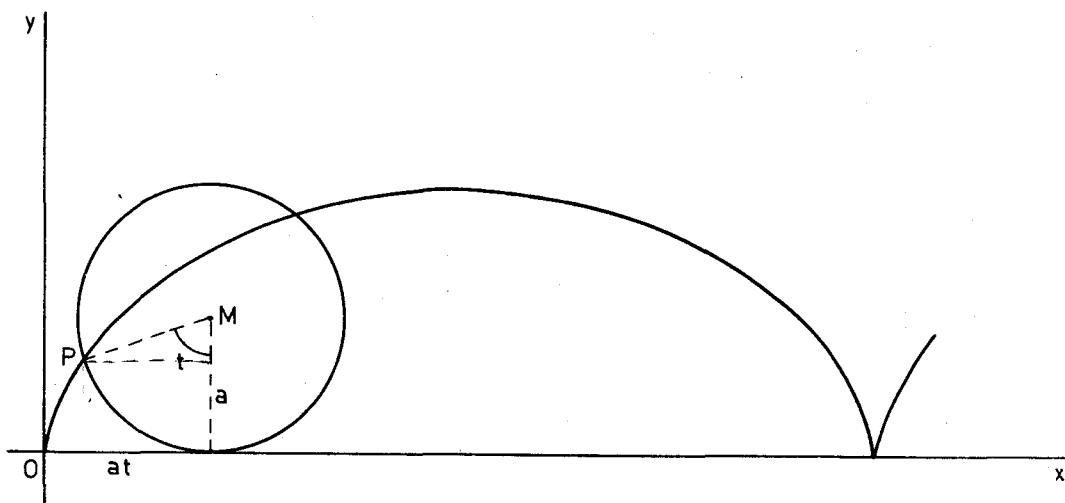
$$\frac{dy}{dx}(x(a)) = \frac{dy}{dt}(a) \frac{dt}{dx}(x(a)) = \frac{\dot{y}(a)}{\dot{x}(a)}.$$

① Opmerking. In vervolg op 6.1.5 zijn ook de tweede en hogere afgeleiden van $g(x)$ in $x(a)$ te bepalen. Zo volgt eenvoudig met behulp van de kettingregel (ga na),

$$\frac{d^2y}{dx^2}(x(a)) = \frac{\ddot{y}(a)\dot{x}(a) - \dot{y}(a)\ddot{x}(a)}{[\dot{x}(a)]^3},$$

waarin $\ddot{x}(a)$, $\ddot{y}(a)$ de tweede afgeleide van $x(t)$, $y(t)$ in a is.

Voorbeeld 6. Cycloïde.



Een cirkel met middelpunt M en straal a rolt over de x -as. Beschouw de kromme beschreven door het punt P dat zich oorspronkelijk in \mathcal{O} bevindt. Deze kromme heet cycloïde.

Neem als parameter de hoek t tussen de voerstraal MP en de verticaal. Uit de figuur is eenvoudig af te lezen dat de coördinaten van P worden gegeven door $(at - a \sin t, a - a \cos t)$. Hiermee hebben we tevens een parameter-voorstelling van de cycloïde gevonden nl. $\underline{f}(t) = a(t - \sin t, 1 - \cos t)$.

② Opgave. Bepaal de punten van de cycloïde waar de raaklijn evenwijdig is aan één van de assen.

23-1-75

6.2. Functies van twee variabelen

Een reële functie f van twee reële variabelen is een afbeelding, waarbij aan elke $\underline{x} = (x, y) \in \text{DOM } f \subset \mathbb{R}^2$ een getal $f(\underline{x}) = f(x, y) \in \mathbb{R}$ wordt toegevoegd. $\text{DOM } f$ is weer de definitieverzameling van f . We noemen f kortweg een functie van twee variabelen.

Vaak wordt een functiewaarde $f(x, y)$ met een aparte letter z aangegeven; men schrijft dan $z = f(x, y)$.

Voorbeelden. 1) $f : \mathbb{R}^2 \rightarrow \mathbb{R}, f(\underline{x}) = f(x, y) = 6 - 3x - 2y$.

2) $f : \mathbb{R}^2 \rightarrow \mathbb{R}, f(x, y) = x^2 + y^2$.

3) $f : \mathbb{R}^2 \rightarrow \mathbb{R}, f(x, y) = x^2 - y^2$.

4) $f : \mathbb{R}^2 \setminus \{0\} \rightarrow \mathbb{R}, f(x, y) = x^2 / (x^2 + y^2)$.

5) $f : \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq 1\} \rightarrow \mathbb{R}, f(x, y) = \sqrt{1 - x^2 - y^2}$.

Grafiek. Kies in \mathbb{R}^3 het gebruikelijke assenstelsel. De grafiek van de functie f is de puntverzameling

$$S := \{(x,y,z) \in \mathbb{R}^3 \mid z = f(x,y)\}.$$

Meetkundig stelt S een oppervlak in \mathbb{R}^3 voor. Men noemt $z = f(x,y)$ de vergelijking van het oppervlak S .

In voorbeeld 1 is de grafiek een plat vlak, in voorbeeld 5 een halve bol. In de voorbeelden 2 en 3 zijn de grafieken resp. een elliptische paraboloid en een hyperbolische paraboloid.

Hoogtekaart. Dikwijls is het moeilijk om een goede grafiek te tekenen; men kan dan gebruik maken van een hoogtekaart. Men tekent in het (x,y) -vlak een aantal niveaulijnen, dat zijn krommen waarop de functie constant is. De niveaulijnen worden ook wel hoogtelijnen genoemd.

In de voorbeelden 1-5 zijn de hoogtelijnen respectievelijk:

- 1) evenwijdige rechten
- 2) concentrische cirkels
- 3) hyperbolen
- 4) rechten door de oorsprong (met uitzondering van de oorsprong)
- 5) concentrische cirkels.

Opgave. Teken hoogtekaarten van de functies

$$f(x,y) = \frac{2y}{1+x^2+y^2}, \quad f(x,y) = \frac{y^2-x}{y^2}.$$

Limieten van functies

- 6.2.1. Definitie. Zij $\underline{a} \in \mathbb{R}^2$ en \underline{a} een verdichtingspunt van $\text{DOM } f$. Dan betekent $\lim_{\underline{x} \rightarrow \underline{a}} f(\underline{x}) = L$, dat er voor iedere omgeving V van L een omgeving U van \underline{a} bestaat zodat voor alle $\underline{x} \in U \cap \text{DOM } f$ met $\underline{x} \neq \underline{a}$ geldt: $f(\underline{x}) \in V$.

Vergelijk deze definitie met 2.4.1.

6.2.2. Stelling. De volgende beweringen zijn gelijkwaardig:

1) $\lim_{\underline{x} \rightarrow \underline{a}} f(\underline{x}) = L.$

2) \underline{a} is een verdichtingspunt van DOM f en

$$\forall \epsilon > 0 \exists \delta > 0 \forall \underline{x} \in \text{DOM } f [0 < |\underline{x} - \underline{a}| < \delta \Rightarrow |f(\underline{x}) - L| < \epsilon].$$

3) $\lim_{\underline{x} \rightarrow \underline{a}} (f(\underline{x}) - L) = 0.$

4) $\lim_{\underline{x} \rightarrow \underline{a}} |f(\underline{x}) - L| = 0.$

Opgave. Bewijs deze stelling.

Voorbeelden. 6) $\lim_{\underline{x} \rightarrow \underline{0}} \frac{xy}{\sqrt{x^2 + y^2}} = 0.$

We gebruiken formulering 2) van 6.2.2. Voor $(x, y) \neq (0, 0)$ geldt:

$$\left| \frac{xy}{\sqrt{x^2 + y^2}} - 0 \right| = \frac{|x||y|}{\sqrt{x^2 + y^2}} = \frac{|x||y|}{|\underline{x}|} \leq \frac{|\underline{x}||\underline{x}|}{|\underline{x}|} = |\underline{x}|.$$

Zij $\epsilon > 0$, dan geldt voor alle \underline{x} met $0 < |\underline{x} - \underline{0}| < \delta := \epsilon$ dat

$$\left| \frac{xy}{\sqrt{x^2 + y^2}} - 0 \right| \leq |\underline{x}| < \epsilon.$$

7) $\lim_{\underline{x} \rightarrow \underline{0}} \frac{xy}{x^2 + y^2}$ bestaat niet.

Dit volgt onmiddellijk uit een hoogtekaart. De rechten $y = \alpha x$ ($x \neq 0$) zijn hoogtelijnen. De waarde van de functie op $y = \alpha x$ is $\alpha/(1 + \alpha^2)$. In elke omgeving van $\underline{0}$ neemt de functie alle waarden uit $[-\frac{1}{2}, \frac{1}{2}]$ aan.

6.2.3. Stelling (bewerkingen met limieten). Als $\lim_{\underline{x} \rightarrow \underline{a}} f(\underline{x}) = L$, $\lim_{\underline{x} \rightarrow \underline{a}} g(\underline{x}) = M$, dan is:

1) $\lim_{\underline{x} \rightarrow \underline{a}} \{f(\underline{x}) + g(\underline{x})\} = L + M;$

2) $\lim_{\underline{x} \rightarrow \underline{a}} \{f(\underline{x})g(\underline{x})\} = LM;$

3) $\lim_{\underline{x} \rightarrow \underline{a}} \frac{f(\underline{x})}{g(\underline{x})} = \frac{L}{M}$ mits $M \neq 0.$

(N.B. Stilzwijgend is ondersteld dat $f(\underline{x})$ en $g(\underline{x})$ dezelfde definitieverzameling hebben.)

Opgave. Bewijs deze stelling.

Continuïteit.

6.2.4. Definitie. Een functie f heet continu in $\underline{a} \in \text{DOM } f$ indien er voor iedere omgeving V van $f(\underline{a})$ een omgeving U van \underline{a} bestaat zó dat voor alle $\underline{x} \in U \cap \text{DOM } f$ geldt: $f(\underline{x}) \in V$.

Vergelijk deze definitie met 2.5.1.

Definitie. Een functie f heet continu op een verzameling $W \subset \mathbb{R}^2$ als f continu is in elk punt van W .

Een functie f heet continu als f continu is op $\text{DOM } f$.

Bewijs zelf de volgende twee stellingen.

Stelling. Een functie f is continu in $\underline{a} \in \text{DOM } f$ dan en slechts dan als

$$\forall \epsilon > 0 \exists \delta > 0 \forall \underline{x} \in \text{DOM } f [|\underline{x} - \underline{a}| < \delta \Rightarrow |f(\underline{x}) - f(\underline{a})| < \epsilon].$$

Stelling. Zij $\underline{a} \in \text{DOM } f$ een verdichtingspunt van $\text{DOM } f$, dan is f continu in \underline{a} dan en slechts dan als $\lim_{\underline{x} \rightarrow \underline{a}} f(\underline{x}) = f(\underline{a})$.

Voorbeelden. 8) $f: \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(\underline{x}) = \begin{cases} \frac{x^3 y}{x^2 + y^2} & \text{als } \underline{x} \neq \underline{0}, \\ 0 & \text{als } \underline{x} = \underline{0}. \end{cases}$

Deze functie is continu in $(0,0)$. Voor $\underline{x} \neq \underline{0}$ geldt namelijk

$$|f(\underline{x}) - f(\underline{0})| = \left| \frac{x^3 y}{x^2 + y^2} \right| = \frac{x^2}{x^2 + y^2} |xy| \leq |x| |y| \leq |\underline{x}|^2.$$

Zij $\epsilon > 0$, dan geldt voor alle \underline{x} met $|\underline{x}| < \delta := \sqrt{\epsilon}$ dat

$$|f(\underline{x}) - f(\underline{0})| < \epsilon.$$

9) $f: \mathbb{R}^2 \rightarrow \mathbb{R}$, $f(\underline{x}) = \begin{cases} \frac{xy}{\sqrt{x^2 + y^2}} & \text{als } \underline{x} \neq \underline{0}, \\ 0 & \text{als } \underline{x} = \underline{0}. \end{cases}$

f is continu in (0,0) omdat

$$\lim_{\underline{x} \rightarrow \underline{0}} \frac{xy}{\sqrt{x^2 + y^2}} = 0 = f(\underline{0}),$$

zie voorbeeld 6.

$$10) f: \mathbb{R}^2 \rightarrow \mathbb{R}, f(\underline{x}) = \begin{cases} \frac{xy}{x^2 + y^2} & \text{als } \underline{x} \neq \underline{0}, \\ 0 & \text{als } \underline{x} = \underline{0}. \end{cases}$$

f is niet continu in (0,0) omdat $\lim_{\underline{x} \rightarrow \underline{0}} f(\underline{x})$ niet bestaat, zie voorbeeld 7.

Met behulp van de volgende stelling kan men meestal snel inzien of een functie continu is.

6.2.5. Stelling. Zijn de functies f en g continu in \underline{a} , dan zijn ook f + g, λf ($\lambda \in \mathbb{R}$), fg en f/g (mits $g(\underline{a}) \neq 0$) continu in \underline{a} . Is $\underline{g}: \mathbb{R} \rightarrow \mathbb{R}^2$ continu in a en $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ continu in $\underline{b} = \underline{g}(a)$, dan is de samengestelde functie $f \circ \underline{g}$ continu in a.

Opgave. Bewijs deze stelling; vergelijk met 2.5.3.

Met 6.2.5 volgt onmiddellijk dat de functie uit voorbeeld 10 continu is op $\mathbb{R}^2 \setminus \{0\}$. De functies uit de voorbeelden 8 en 9 zijn continu op de hele \mathbb{R}^2 .

15.2.78
6.3. Differentieerbaarheid van functies van twee variabelen

6.3.1. Definitie. Zij f gedefinieerd in een omgeving U van $\underline{a} = (a,b)$. De functie f heet differentieerbaar in \underline{a} , indien er getallen $A \in \mathbb{R}$ en $B \in \mathbb{R}$ bestaan en een functie ρ gedefinieerd in een omgeving V van $\underline{0}$, zodanig dat

(i) $f(a+h, b+k) = f(a,b) + Ah + Bk + |\underline{h}| \rho(\underline{h})$ voor alle $\underline{h} = (h,k) \in V$,

(ii) $\lim_{\underline{h} \rightarrow \underline{0}} \rho(\underline{h}) = 0$.

De betekenis van A en B blijkt door in (i) $k = 0$ resp. $h = 0$ te nemen. We vinden dan

$$A = \lim_{h \rightarrow 0} \frac{f(a+h, b) - f(a, b)}{h} \quad \text{resp.} \quad B = \lim_{k \rightarrow 0} \frac{f(a, b+k) - f(a, b)}{k}.$$

A is dus de afgeleide van de functie $f(x,b)$ in $x = a$.

B is de afgeleide van $f(a,y)$ in het punt $y = b$.

We noemen deze afgeleiden, waarbij slechts één van de argumenten varieert, partiële afgeleiden.

Notaties. De partiële afgeleide naar x van $z = f(x,y)$ in (a,b) wordt gegeven door

$$\lim_{h \rightarrow 0} \frac{f(a+h,b) - f(a,b)}{h} =: f_x(a,b) =: \frac{\partial f}{\partial x}(a,b) =: \frac{\partial z}{\partial x}(a,b).$$

De partiële afgeleide naar y van $z = f(x,y)$ in (a,b) wordt gegeven door

$$\lim_{k \rightarrow 0} \frac{f(a,b+k) - f(a,b)}{k} =: f_y(a,b) =: \frac{\partial f}{\partial y}(a,b) =: \frac{\partial z}{\partial y}(a,b).$$

Met deze notatie zijn de getallen A en B uit definitie 6.3.1 voor te stellen door

$$A = f_x(a,b), \quad B = f_y(a,b).$$

De definitie van differentieerbaarheid drukt uit dat f in \underline{a} lineair benaderbaar is, d.w.z. voor $\underline{x} = (x,y)$ in de buurt van $\underline{a} = (a,b)$ is $f(x,y)$ te benaderen door de lineaire functie

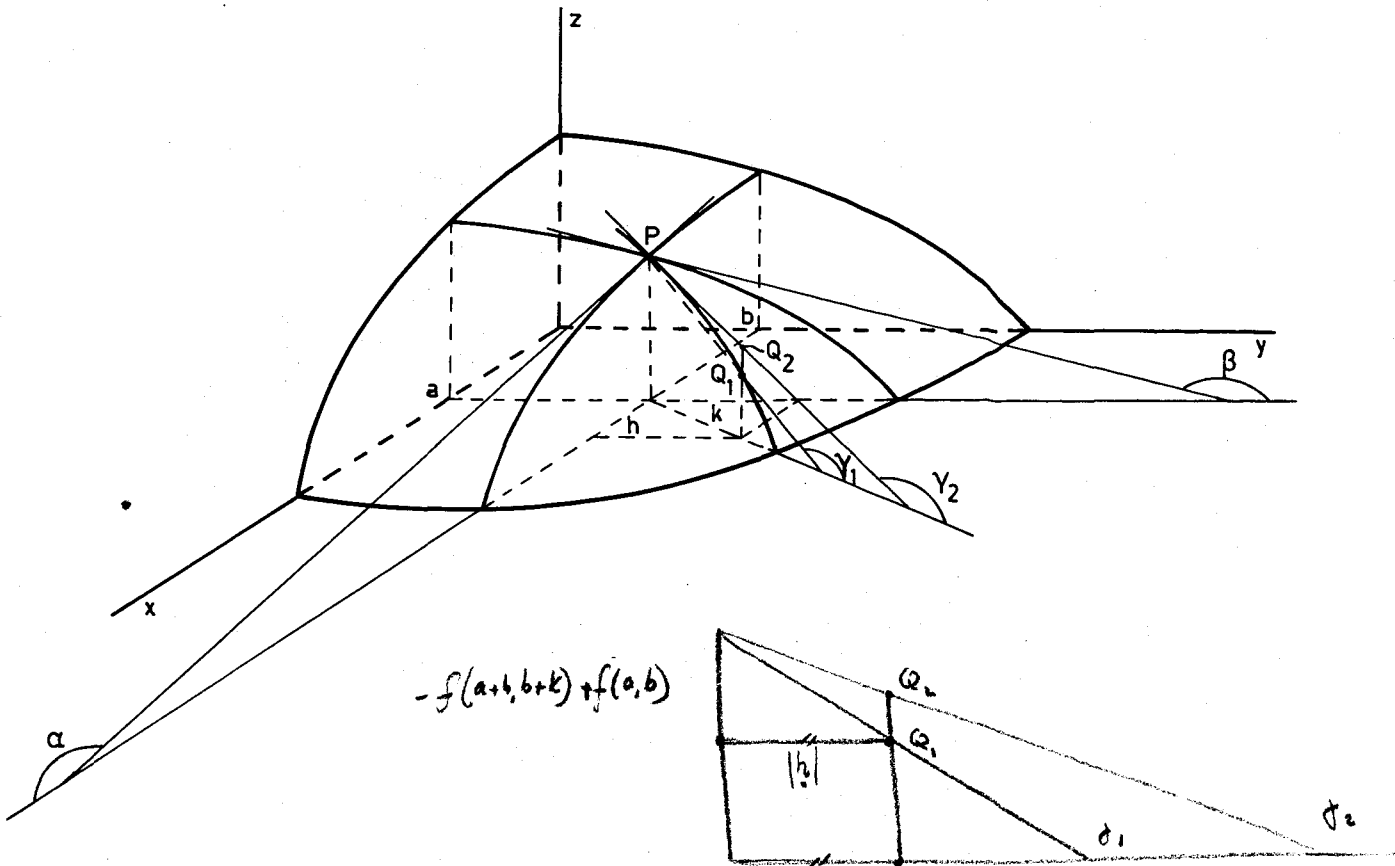
$$l(x,y) := f(a,b) + f_x(a,b)(x-a) + f_y(a,b)(y-b).$$

Bij deze benadering maken we een fout gegeven door $|\underline{x}-\underline{a}| \rho(\underline{x}-\underline{a})$, waarbij $\rho(\underline{x}-\underline{a}) \rightarrow 0$ ($\underline{x} \rightarrow \underline{a}$).

Meetkundige interpretatie. Beschouw de grafiek van f d.i. het oppervlak gegeven door de vergelijking $z = f(x,y)$. Het punt P met coördinaten $(a,b,f(a,b))$ ligt op het oppervlak.

Het vlak $y = b$ snijdt het oppervlak volgens een kromme door P. De raaklijn aan de snijkromme in P ligt in het vlak $y = b$ en heeft de richtingscoëfficiënt $\tan \alpha = f_x(a,b)$, zie figuur.

Evenzo is $\tan \beta = f_y(a,b)$ de richtingscoëfficiënt van de raaklijn in P aan de snijkromme van het vlak $x = a$ met het oppervlak, zie figuur.



Breng nu een vlak aan door de beide raaklijnen in P. De vergelijking van dit vlak luidt

6.3.2.
$$z - f(a,b) = f_x(a,b)(x-a) + f_y(a,b)(y-b).$$

Dit vlak is het raakvlak aan het oppervlak $z = f(x,y)$ in het punt P. Daartoe tonen we aan dat elke rechte door P in het vlak een raaklijn is aan het oppervlak. Beschouw twee punten Q_1 en Q_2 met coördinaten:

$$Q_1 : (a+h, b+k, f(a+h, b+k)),$$

$$Q_2 : (a+h, b+k, f(a,b) + f_x(a,b)h + f_y(a,b)k),$$

dan ligt Q_1 op het oppervlak en Q_2 in het vlak (zie figuur). De rechten PQ_1 en PQ_2 maken hoeken γ_1 en γ_2 met het (x,y) -vlak, gegeven door

$$\tan \gamma_1 = \frac{1}{|h|} [f(a+h, b+k) - f(a,b)],$$

$$\tan \gamma_2 = \frac{1}{|h|} [f_x(a,b)h + f_y(a,b)k].$$

Volgens 6.3.1, punt (i) is nu juist $\tan \gamma_1 - \tan \gamma_2 = \rho(\underline{h})$. Neem nu $\underline{h} \rightarrow 0$ bij vaste h/k , dan is γ_2 eveneens vast en $\gamma_1 \rightarrow \gamma_2$ omdat $\rho(\underline{h}) \rightarrow 0$. De rechte PQ_1 heeft als limietstand de rechte PQ_2 . Derhalve raakt PQ_2 aan het oppervlak. Meetkundig betekent differentieerbaarheid van f in \underline{a} dus dat het oppervlak $z = f(x,y)$ in het punt $(a,b,f(a,b))$ een niet-verticaal raakvlak heeft.

Definitie. Een functie f heet differentieerbaar op een open verzameling $W \subset \mathbb{R}^2$ als f differentieerbaar is in elk punt van W .

Bewijs zelf de volgende stelling:

Stelling. Als f differentieerbaar is in \underline{a} , dan is f continu in \underline{a} .

Voorbeelden. 1) $f(x,y) = xy^2 + y^3$;

de partiële afgeleiden van f zijn

$$f_x(x,y) = y^2, f_y(x,y) = 2xy + 3y^2.$$

Merk op dat f_x en f_y ook weer functies van x en y zijn.

2) $z = \arctan \frac{y}{x}$;

de partiële afgeleiden zijn

$$\frac{\partial z}{\partial x} = \frac{-y}{x^2 + y^2}, \frac{\partial z}{\partial y} = \frac{x}{x^2 + y^2}.$$

In het punt $(1,1)$ is

$$\frac{\partial z}{\partial x}(1,1) = -\frac{1}{2}, \frac{\partial z}{\partial y}(1,1) = \frac{1}{2}.$$

Vooruitlopend op 6.3.3 kan men inzien dat de functie differentieerbaar is in $(1,1)$.

In de buurt van $(1,1)$ is z dan lineair benaderbaar door

$$\frac{\pi}{4} - \frac{1}{2}(x-1) + \frac{1}{2}(y-1).$$

De vergelijking van het raakvlak aan de grafiek in $(1,1, \frac{\pi}{4})$ is

$$z - \frac{\pi}{4} = -\frac{1}{2}(x-1) + \frac{1}{2}(y-1).$$

3) $f(\underline{x}) = \sqrt{x^2 + y^2}$ is continu in $\underline{0}$. Zij nl. $\epsilon > 0$ dan geldt voor alle \underline{x} met $|\underline{x} - \underline{0}| < \delta := \epsilon$ dat

$$|f(\underline{x}) - f(\underline{0})| = |\underline{x}| < \epsilon.$$

De partiële afgeleiden van f in $(0,0)$ bestaan niet, dus de functie is niet differentieerbaar in $(0,0)$.

- 4) Uit de differentieerbaarheid van f volgt het bestaan van de partiële afgeleiden f_x , f_y . Omgekeerd volgt uit het bestaan van de partiële afgeleiden nog niet dat f differentieerbaar is.

De functie

$$f(\underline{x}) = \begin{cases} \frac{xy}{x^2 + y^2} & \text{als } \underline{x} \neq \underline{0}, \\ 0 & \text{als } \underline{x} = \underline{0}, \end{cases}$$

is niet continu in $\underline{0}$ (zie 6.2, voorbeeld 10), dus zeker niet differentieerbaar in $\underline{0}$. Er geldt echter $f_x(0,0) = f_y(0,0) = 0$.

- 6.3.3. Stelling. Zij f gedefinieerd in een omgeving van \underline{a} en laat de partiële afgeleiden f_x en f_y bestaan in deze omgeving. Als f_x en f_y continu zijn in \underline{a} , dan is f differentieerbaar in \underline{a} .

Bewijs. Volgens het gegeven van de stelling is er een omgeving

$U := \{\underline{x} \in \mathbb{R}^2 \mid |\underline{x} - \underline{a}| < \delta\}$ van \underline{a} zodat f_x en f_y bestaan in U . Zij $\underline{a} = (a,b)$, $\underline{h} = (h,k)$ met $|\underline{h}| < \delta$.

Volgens de middelwaardestelling (2.6.7) is er een getal $\theta \in (0,1)$ zodat

$$f(a+h, b+k) - f(a, b+k) = hf_x(a+\theta h, b+k).$$

Omdat f_x continu is in (a,b) , kunnen we schrijven

$$f_x(a+\theta h, b+k) = f_x(a, b) + \rho(h, k)$$

met

$$\lim_{\underline{h} \rightarrow \underline{0}} \rho(h, k) = 0.$$

Uit de voorgaande betrekkingen volgt

$$f(a+h, b+k) - f(a, b+k) = hf_x(a, b) + h\rho(h, k).$$

Evenzo is af te leiden

$$f(a, b+k) - f(a, b) = kf_y(a, b) + k\sigma(k)$$

met

$$\lim_{k \rightarrow 0} \sigma(k) = 0.$$

Nu volgt

$$\begin{aligned} f(a+h, b+k) - f(a, b) &= f(a+h, b+k) - f(a, b+k) + f(a, b+k) - f(a, b) = \\ &= f_x(a, b)h + f_y(a, b)k + |\underline{h}| \tau(\underline{h}), \end{aligned}$$

waarbij $\tau(\underline{h})$ gegeven wordt door

$$\tau(\underline{h}) = \tau(h, k) = \frac{1}{|\underline{h}|} [h\rho(h, k) + k\sigma(k)].$$

Wegens $\lim_{\underline{h} \rightarrow 0} \tau(\underline{h}) = 0$ is f differentieerbaar in \underline{a} . □

Uit 6.3.3 volgt dat functies met continue partiële afgeleiden differentieerbaar zijn. (Functies met continue partiële afgeleiden noemt men continu differentieerbaar.) Hiermee en met de volgende stelling kan men meestal snel vaststellen of een functie differentieerbaar is.

6.3.4. Stelling. Als de functies f, g differentieerbaar zijn in \underline{a} , dan zijn ook $f+g, \lambda f$ ($\lambda \in \mathbb{R}$), fg en f/g (mits $g(\underline{a}) \neq 0$) differentieerbaar in \underline{a} .

Opgave. Bewijs deze stelling.

We geven nu een tweetal kettingregels voor het differentiëren van een samengestelde functie.

6.3.5. Stelling. Als de functie $\underline{g} = (g_1, g_2) : \mathbb{R} \rightarrow \mathbb{R}^2$ differentieerbaar is in a en $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ differentieerbaar is in $\underline{g}(a)$, dan is de samengestelde functie $f \circ \underline{g} : \mathbb{R} \rightarrow \mathbb{R}$ differentieerbaar in a met afgeleide

$$(f \circ \underline{g})'(a) = f_x(\underline{g}(a))g_1'(a) + f_y(\underline{g}(a))g_2'(a).$$

Opmerking. Wat slordiger kan men de kettingregel 6.3.5 ook als volgt formuleren: zijn

$$z = f(x, y), \quad x = g_1(t), \quad y = g_2(t)$$

differentieerbare functies, dan is de samengestelde functie

$$z(t) = f(g_1(t), g_2(t))$$

differentieerbaar met afgeleide

$$\frac{dz}{dt} = \frac{\partial z}{\partial x} \frac{dx}{dt} + \frac{\partial z}{\partial y} \frac{dy}{dt}.$$

Bewijs. Op grond van de differentieerbaarheid van g_1 en g_2 in a geldt:

$$g_1(a+h) = g_1(a) + g_1'(a)h + h\rho_1(h), \lim_{h \rightarrow 0} \rho_1(h) = 0,$$

$$g_2(a+h) = g_2(a) + g_2'(a)h + h\rho_2(h), \lim_{h \rightarrow 0} \rho_2(h) = 0,$$

of in vectornotatie

$$\underline{g}(a+h) = \underline{g}(a) + \underline{g}'(a)h + h\underline{\rho}(h) \quad , \quad \lim_{h \rightarrow 0} |\underline{\rho}(h)| = 0.$$

Uit de differentieerbaarheid van f in $\underline{g}(a)$ volgt

$$\begin{aligned} f(\underline{g}(a+h)) &= f(\underline{g}(a) + \underline{g}'(a)h + h\underline{\rho}(h)) = \\ &= f(\underline{g}(a)) + f_x(\underline{g}(a))[g_1'(a)h + h\rho_1(h)] + f_y(\underline{g}(a)) [g_2'(a)h + h\rho_2(h)] + \\ &+ |\underline{g}'(a)h + h\underline{\rho}(h)| \sigma(\underline{g}'(a)h + h\underline{\rho}(h)) \end{aligned}$$

met $\sigma \rightarrow 0$ voor $h \rightarrow 0$.

Hieruit volgt

$$f(\underline{g}(a+h)) = f(\underline{g}(a)) + [f_x(\underline{g}(a))g_1'(a) + f_y(\underline{g}(a))g_2'(a)]h + h\tau(h),$$

waarbij $\tau(h)$ gegeven wordt door

$$\tau(h) = f_x(\underline{g}(a))\rho_1(h) + f_y(\underline{g}(a))\rho_2(h) + \frac{|h|}{h} |\underline{g}'(a) + \underline{\rho}(h)| \sigma.$$

Het is duidelijk dat $\lim_{h \rightarrow 0} \tau(h) = 0$. Daarmee is bewezen dat $f \circ g$ differentieerbaar is in a met afgeleide $f_x(\underline{g}(a))g_1'(a) + f_y(\underline{g}(a))g_2'(a)$. □

6.3.6. Stelling. Als $g : \mathbb{R}^2 \rightarrow \mathbb{R}$ differentieerbaar is in \underline{a} en $f : \mathbb{R} \rightarrow \mathbb{R}$ differentieerbaar is in $g(\underline{a})$, dan is de samengestelde functie $f \circ g : \mathbb{R}^2 \rightarrow \mathbb{R}$ differentieerbaar in \underline{a} met partiële afgeleiden

$$\frac{\partial (f \circ g)}{\partial x}(\underline{a}) = f'(g(\underline{a}))g_x(\underline{a}), \quad \frac{\partial (f \circ g)}{\partial y}(\underline{a}) = f'(g(\underline{a}))g_y(\underline{a}).$$

Bewijs. Het bewijs is analoog aan dat van 6.3.5 en wordt als opgave aan de lezer overgelaten.

□

15.2.78.

Voorbeelden. 5) $z = xe^y$ met $x = \sin t$, $y = t^2$. Dan is

$$\frac{dz}{dt} = e^y \cos t + xe^y 2t = e^{t^2} \cos t + 2te^{t^2} \sin t.$$

6) Men dient duidelijk onderscheid te maken tussen de notaties $\frac{\partial z}{\partial x}$ en $\frac{dz}{dx}$.
Door $z = x + y^2$, $y = \log x$ is z gegeven als functie van x en y , maar via $y = \log x$ ook als functie van x alleen (nl. $z = x + (\log x)^2$).

Er geldt nu

$$\frac{\partial z}{\partial x} = 1 \quad \text{en} \quad \frac{dz}{dx} = 1 + \frac{2 \log x}{x}.$$

In dit voorbeeld speelt x de rol van de variabele t in de stelling 6.3.5.

Hogere partiële afgeleiden. De partiële afgeleiden f_x en f_y van een functie f zijn zelf ook weer functies van x en y . Partieel differentiëren (indien mogelijk) van f_x en f_y geeft de partiële afgeleiden van de tweede orde van f , die als volgt genoteerd worden:

$$\frac{\partial}{\partial x} \left(\frac{\partial f}{\partial x} \right) =: \frac{\partial^2 f}{\partial x^2} =: f_{xx}, \quad \frac{\partial}{\partial y} \left(\frac{\partial f}{\partial x} \right) =: \frac{\partial^2 f}{\partial y \partial x} =: f_{xy},$$

$$\frac{\partial}{\partial x} \left(\frac{\partial f}{\partial y} \right) =: \frac{\partial^2 f}{\partial x \partial y} =: f_{yx}, \quad \frac{\partial}{\partial y} \left(\frac{\partial f}{\partial y} \right) =: \frac{\partial^2 f}{\partial y^2} =: f_{yy}.$$

N.B. Let op de volgorde van differentiatie; f_{xy} wordt verkregen door eerst naar x en dan naar y te differentiëren; f_{yx} wordt verkregen door eerst naar y en dan naar x te differentiëren.

Voorbeelden. 7) $f(x,y) = x^2 y^3$;

de partiële afgeleiden van de eerste en tweede orde van f zijn

$$f_x = 2xy^3, \quad f_y = 3x^2 y^2,$$

$$f_{xx} = 2y^3, \quad f_{xy} = 6xy^2, \quad f_{yx} = 6xy^2, \quad f_{yy} = 6x^2 y.$$

Merk op dat $f_{xy} = f_{yx}$. In het volgende voorbeeld is dit niet het geval.

8) $f(x,y) = xy \frac{x^2 - y^2}{x^2 + y^2}$ als $(x,y) \neq (0,0)$ en $f(0,0) = 0$.

Bereken eerst de partiële afgeleiden van de eerste orde

$$f_x(0,k) = \lim_{h \rightarrow 0} \frac{f(h,k) - f(0,k)}{h} = \lim_{h \rightarrow 0} k \frac{h^2 - k^2}{h^2 + k^2} = -k,$$

$$f_y(h,0) = \lim_{k \rightarrow 0} \frac{f(h,k) - f(h,0)}{k} = \lim_{k \rightarrow 0} h \frac{h^2 - k^2}{h^2 + k^2} = h.$$

Bereken vervolgens f_{xy} en f_{yx} in $(0,0)$,

$$f_{xy}(0,0) = \lim_{k \rightarrow 0} \frac{f_x(0,k) - f_x(0,0)}{k} = \lim_{k \rightarrow 0} \frac{-k - 0}{k} = -1,$$

$$f_{yx}(0,0) = \lim_{h \rightarrow 0} \frac{f_y(h,0) - f_y(0,0)}{h} = \lim_{h \rightarrow 0} \frac{h - 0}{h} = 1,$$

dus $f_{xy}(0,0) \neq f_{yx}(0,0)$.

Uit de volgende stelling blijkt dat het resultaat van tweemaal partieel differentiëren onafhankelijk is van de volgorde van differentiatie, indien de resulterende afgeleiden van de tweede orde continu zijn.

6.3.7. Stelling. Als f_{xy} en f_{yx} bestaan in een omgeving van \underline{a} en continu zijn in \underline{a} , dan geldt: $f_{xy}(\underline{a}) = f_{yx}(\underline{a})$.

Bewijs. Laat f_{xy} en f_{yx} bestaan in een omgeving $U := \{\underline{x} \in \mathbb{R}^2 \mid |\underline{x} - \underline{a}| < \delta\}$ van \underline{a} . Zij $\underline{a} = (a,b)$, $\underline{h} = (h,k)$ met $|\underline{h}| < \delta$.

Beschouw de uitdrukking

$$E := f(a+h, b+k) - f(a+h, b) - f(a, b+k) + f(a, b).$$

Zij

$$\varphi(x) := f(x, b+k) - f(x, b),$$

dan is

$$E = \varphi(a+h) - \varphi(a).$$

Pas de middelwaardestelling (2.6.7) toe, dan is

$$E = h\varphi'(a+\theta_1 h) = h[f_x(a+\theta_1 h, b+k) - f_x(a+\theta_1 h, b)]$$

voor zekere $\theta_1 \in (0,1)$.

Nogmaals toepassen van de middelwaardestelling geeft

$$E = hk f_{xy}(a+\theta_1 h, b+\theta_2 k)$$

voor zekere $\theta_2 \in (0,1)$.

Zij vervolgens

$$\psi(y) := f(a+h, y) - f(a, y) ,$$

dan is

$$E = \psi(b+k) - \psi(b).$$

Door opnieuw tweemaal toepassen van de middelwaardestelling volgt

$$\begin{aligned} E &= k\psi'(b+\theta_3 k) = k[f_y(a+h, b+\theta_3 k) - f_y(a, b+\theta_3 k)] = \\ &= hk f_{yx}(a+\theta_4 h, b+\theta_3 k) \end{aligned}$$

voor zekere $\theta_3 \in (0,1)$, $\theta_4 \in (0,1)$.

Combineer de voorgaande resultaten dan is

$$f_{xy}(a+\theta_1 h, b+\theta_2 k) = f_{yx}(a+\theta_4 h, b+\theta_3 k) .$$

Laat nu $\underline{h} = (h, k)$ naar $\underline{0}$ naderen, dan volgt op grond van de continuïteit van f_{xy} en f_{yx} in \underline{a} dat $f_{xy}(\underline{a}) = f_{yx}(\underline{a})$. □

Men kan nu ook partiële afgeleiden van de derde en hogere orde definiëren. Er zijn 2^n partiële afgeleiden van de orde n . Als deze afgeleiden continu zijn, dan kan men weer bewijzen dat het resultaat van partieel differentiëren onafhankelijk is van de volgorde van differentiatie, zodat er slechts $n+1$ verschillende partiële afgeleiden van de orde n zijn. Wij zullen in het vervolg aannemen dat van de te beschouwen functies de gemengde partiële afgeleiden die door verwisseling van volgorde uit elkaar ontstaan, inderdaad gelijk zijn.

Er geldt dan bijv. $f_{xxy} = f_{xyx} = f_{yxx}$ oftewel $\frac{\partial^3 f}{\partial y \partial x^2} = \frac{\partial^3 f}{\partial x \partial y \partial x} = \frac{\partial^3 f}{\partial x^2 \partial y}$.

Opmerking. Kettingregels voor hogere partiële afgeleiden kunnen in elk afzonderlijk geval worden afgeleid door toepassing van de gewone kettingregel (6.3.5), waarbij te bedenken is dat partiële afgeleiden van x en van y afhangen.

Voorbeeld. Zijn de functies

$$z = f(x,y), \quad x = x(t), \quad y = y(t)$$

tweemaal differentieerbaar, dan is de samengestelde functie

$$z = f(x(t), y(t))$$

tweemaal differentieerbaar met afgeleiden

$$\frac{dz}{dt} = \frac{\partial z}{\partial x} \frac{dx}{dt} + \frac{\partial z}{\partial y} \frac{dy}{dt},$$

$$\begin{aligned} \frac{d^2 z}{dt^2} &= \frac{\partial z}{\partial x} \frac{d^2 x}{dt^2} + \left[\frac{\partial^2 z}{\partial x^2} \frac{dx}{dt} + \frac{\partial^2 z}{\partial y \partial x} \frac{dy}{dt} \right] \frac{dx}{dt} + \\ &+ \frac{\partial z}{\partial y} \frac{d^2 y}{dt^2} + \left[\frac{\partial^2 z}{\partial x \partial y} \frac{dx}{dt} + \frac{\partial^2 z}{\partial y^2} \frac{dy}{dt} \right] \frac{dy}{dt}. \end{aligned}$$

Formule van Taylor. We zullen de formule van Taylor (2.6.10) generaliseren tot functies van twee variabelen. We voeren de volgende notatie in:

$$\left(h \frac{\partial}{\partial x} + k \frac{\partial}{\partial y} \right)^m f(x,y) := \sum_{j=0}^m \binom{m}{j} h^{m-j} k^j \frac{\partial^m f}{\partial x^{m-j} \partial y^j} (x,y),$$

voor $m \in \mathbb{N}$; vergelijk met het binomium van Newton (1.3.7).

Hierbij is verondersteld dat f continue partiële afgeleiden van de orde m heeft.

6.3.8. Stelling (Taylor). Laat de functie f continue partiële afgeleiden van de orde $n+1$ ($n \geq 0$) bezitten in een omgeving $U := \{ \underline{x} \in \mathbb{R}^2 \mid |\underline{x} - \underline{a}| < \delta \}$ van \underline{a} . Zij $\underline{a} + \underline{h} \in U$, $\underline{h} = (h,k)$, dan is er een getal $\theta \in (0,1)$, zodat

$$f(\underline{a} + \underline{h}) = f(\underline{a}) + \sum_{m=1}^n \frac{1}{m!} \left(h \frac{\partial}{\partial x} + k \frac{\partial}{\partial y} \right)^m f(\underline{a}) + R_n,$$

waarin

$$R_n = \frac{1}{(n+1)!} \left(h \frac{\partial}{\partial x} + k \frac{\partial}{\partial y} \right)^{n+1} f(\underline{a} + \theta \underline{h}).$$

Bewijs. Met behulp van $\underline{g}(t) := \underline{a} + t\underline{h}$ introduceren we de samengestelde functie

$$F(t) := f(\underline{g}(t)) = f(\underline{a} + t\underline{h}).$$

Volgens de kettingregel (6.3.5) is F differentieerbaar op $[0,1]$ met afgeleide

$$F'(t) = f_x(\underline{a}+t\underline{h})h + f_y(\underline{a}+t\underline{h})k = \left(h \frac{\partial}{\partial x} + k \frac{\partial}{\partial y}\right) f(\underline{a}+t\underline{h}).$$

Voor $n \geq 1$ zijn f_x en f_y opnieuw differentieerbaar. Nogmaals toepassen van de kettingregel levert voor $t \in [0,1]$,

$$\begin{aligned} F''(t) &= f_{xx}(\underline{a}+t\underline{h})h^2 + 2f_{xy}(\underline{a}+t\underline{h})hk + f_{yy}(\underline{a}+t\underline{h})k^2 = \\ &= \left(h \frac{\partial}{\partial x} + k \frac{\partial}{\partial y}\right)^2 f(\underline{a}+t\underline{h}). \end{aligned}$$

Algemeen is door volledige inductie te bewijzen (ga dit na):

$$F^{(m)}(t) = \left(h \frac{\partial}{\partial x} + k \frac{\partial}{\partial y}\right)^m f(\underline{a}+t\underline{h})$$

voor $m = 1, 2, \dots, n+1$ en $t \in [0,1]$.

Pas nu de formule van Taylor (2.6.10) toe op de functie F , dan is

$$F(1) = F(0) + \sum_{m=1}^n \frac{1}{m!} F^{(m)}(0) + \frac{1}{(n+1)!} F^{(n+1)}(\theta)$$

voor zekere $\theta \in (0,1)$.

Na substitutie van $F(1)$, $F^{(m)}(0)$ enz. als boven gegeven volgt het gewenste resultaat. □

De in 6.3.8 gegeven betrekking voor $f(\underline{a}+h)$ heet de formule van Taylor voor functies van twee variabelen; R_n heet de restterm.

Voorbeeld. Zij $f(x,y) = \sin(xe^y)$, $\underline{a} = (0,0)$, dan is

$$\begin{aligned} f(\underline{h}) &= \sin(he^k) = f(\underline{0}) + [hf_x(\underline{0}) + kf_y(\underline{0})] + \\ &+ \frac{1}{2!} [h^2 f_{xx}(\underline{0}) + 2hkf_{xy}(\underline{0}) + k^2 f_{yy}(\underline{0})] + R_2 \end{aligned}$$

met

$$R_2 = \frac{1}{3!} \left(h \frac{\partial}{\partial x} + k \frac{\partial}{\partial y}\right)^3 f(\theta \underline{h}), \theta \in (0,1).$$

college 30.1

Na uitwerking van de partiële afgeleiden vinden we

$$\sin(h e^k) = h + hk + R_2.$$

Voor kleine waarden van x en y geldt dus de benadering

$$\sin(x e^y) \approx x + xy.$$

Als bijzonder geval ($n = 0$) van 6.3.8 vermelden we nog de volgende stelling die een generalisering is van de middelwaardestelling (2.6.7).

6.3.9. Stelling (middelwaardestelling voor functies van twee variabelen). Laat de functie f continue partiële afgeleiden f_x en f_y bezitten in een omgeving $U := \{\underline{x} \in \mathbb{R}^2 \mid |\underline{x} - \underline{a}| < \delta\}$ van \underline{a} . Zij $\underline{a} + \underline{h} \in U$, $\underline{h} = (h, k)$, dan is er een getal $\theta \in (0, 1)$, zodat

$$f(\underline{a} + \underline{h}) - f(\underline{a}) = hf_x(\underline{a} + \theta \underline{h}) + kf_y(\underline{a} + \theta \underline{h}).$$

Deze stelling is ook rechtstreeks te bewijzen door toepassing van de middelwaardestelling (2.6.7) op de functie $f(\underline{a} + \theta \underline{h})$.

6.4. Functies van drie variabelen

Een reële functie f van drie reële variabelen is een afbeelding, waarbij aan elke $\underline{x} = (x, y, z) \in \text{DOM } f \subset \mathbb{R}^3$ een getal $f(\underline{x}) = f(x, y, z) \in \mathbb{R}$ wordt toegevoegd. We noemen f kortweg een functie van drie variabelen.

De theorie uit 6.2 en 6.3 is zonder meer uit te breiden tot functies van drie variabelen (en ook tot functies van meer dan drie variabelen). In de meeste definities, stellingen en bewijzen behoeven slechts voor de hand liggende wijzigingen aangebracht te worden.

Een eenvoudige meetkundige voorstelling van de functie $f(x, y, z)$, analoog aan de grafiek van een functie van twee variabelen, is er niet. Inzicht in de functie kunnen we verkrijgen door niveaувlakken te beschouwen, dat zijn oppervlakken in \mathbb{R}^3 gegeven door de vergelijking $f(x, y, z) = C$ met C constant. Niveaувlakken zijn het analogon van hoogtelijnen in \mathbb{R}^2 .

Voorbeeld. De functie

$$f(x, y, z) = x^2 + y^2 + z^2$$

heeft als niveaувlakken een stelsel bollen met middelpunt $\underline{0}$.

We herhalen de definitie van differentieerbaarheid en de kettingregel voor functies van drie variabelen.

6.4.1. Definitie. Zij f gedefinieerd in een omgeving U van $\underline{a} = (a,b,c)$. De functie f heet differentieerbaar in \underline{a} , indien er reële getallen A , B en C bestaan en een functie ρ gedefinieerd in een omgeving V van $\underline{0}$, zodanig dat

$$(i) \quad f(a+h, b+k, c+l) = f(a, b, c) + Ah + Bk + Cl + |\underline{h}| \rho(\underline{h})$$

voor alle $\underline{h} = (h, k, l) \in V$,

$$(ii) \quad \lim_{\underline{h} \rightarrow \underline{0}} \rho(\underline{h}) = 0.$$

Het blijkt weer dat A , B en C partiële afgeleiden zijn, nl.

$$A = f_x(\underline{a}), \quad B = f_y(\underline{a}), \quad C = f_z(\underline{a}).$$

6.4.2. Stelling (kettingregel). Zij de functie $\underline{g} = (g_1, g_2, g_3) : \mathbb{R} \rightarrow \mathbb{R}^3$ differentieerbaar in a en zij $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ differentieerbaar in $\underline{g}(a)$, dan is de samengestelde functie $f \circ \underline{g} : \mathbb{R} \rightarrow \mathbb{R}$ differentieerbaar in a met afgeleide

$$(f \circ \underline{g})'(a) = f_x(\underline{g}(a))g_1'(a) + f_y(\underline{g}(a))g_2'(a) + f_z(\underline{g}(a))g_3'(a).$$

Wat slordiger kan men 6.4.2 ook als volgt formuleren: zijn

$$w = f(x, y, z), \quad x = g_1(t), \quad y = g_2(t), \quad z = g_3(t)$$

differentieerbare functies, dan is de samengestelde functie

$$w(t) = f(g_1(t), g_2(t), g_3(t))$$

differentieerbaar met afgeleide

$$\frac{dw}{dt} = \frac{\partial w}{\partial x} \frac{dx}{dt} + \frac{\partial w}{\partial y} \frac{dy}{dt} + \frac{\partial w}{\partial z} \frac{dz}{dt}.$$

Opgave. Formuleer een kettingregel voor de afgeleide van $f \circ g$ als $f : \mathbb{R} \rightarrow \mathbb{R}$ en $g : \mathbb{R}^3 \rightarrow \mathbb{R}$; vergelijk met 6.3.6.

6.5. Impliciete functies

Uit een vergelijking $f(x,y) = 0$ kan men soms y oplossen en uitdrukken in x , d.w.z. we kunnen y als functie van x beschouwen: $y = g(x)$. We zeggen dat y door $f(x,y) = 0$ als functie van x impliciet gegeven is.

Voorbeeld 1. Door de vergelijking $x^2 + y^2 = 1$ (eenheidscirkel) worden op $[-1,1]$ twee functies impliciet gegeven, nl.

$$y = g_1(x) = \sqrt{1-x^2} \quad \text{en} \quad y = g_2(x) = -\sqrt{1-x^2}.$$

Beide functies zijn differentieerbaar op $(-1,1)$ met afgeleiden

$$g_1'(x) = \frac{-x}{\sqrt{1-x^2}} = -\frac{x}{y} \quad \text{en} \quad g_2'(x) = \frac{x}{\sqrt{1-x^2}} = -\frac{x}{y}.$$

Hoewel in het algemeen door de vergelijking $f(x,y) = 0$ een aantal functies $g_1(x), g_2(x), \dots$ bepaald wordt, is het meestal niet gemakkelijk g_1, g_2, \dots ook expliciet in x uit te drukken, zoals het voorbeeld $f(x,y) = x^3 + y^3 - 3xy = 0$ leert.

Om in deze situatie toch de afgeleide van y eenvoudig te bepalen, denken we ons $y = y(x)$ althans in principe opgelost uit $f(x,y) = 0$. Na substitutie van $y = y(x)$ in de vergelijking volgt

$$f(x, y(x)) = 0.$$

$$z = f(x, y) ; x = x ; y = y(x)$$

Differentiatie met behulp van de kettingregel levert nu

$$f_x + f_y y' = 0,$$

$$\frac{dz}{dx} = \frac{\partial z}{\partial x} \cdot \frac{dx}{dx} + \frac{\partial z}{\partial y} \cdot \frac{dy}{dx} \neq 0 = z_x \cdot 1 + z_y \cdot y'$$

waaruit y' opgelost kan worden: $y' = -f_x/f_y$ mits $f_y \neq 0$.

Het spreekt niet vanzelf dat door $f(x,y) = 0$ een differentieerbare functie $y(x)$ bepaald wordt. Preciese voorwaarden hiervoor worden gegeven in een stelling te behandelen in het college Wiskunde 30, die we als volgt citeren: Laat f continue partiële afgeleiden f_x en f_y bezitten in een omgeving van (a,b) en zij $f(a,b) = 0$, $f_y(a,b) \neq 0$, dan bestaat er een functie $y(x)$ gedefinieerd in een omgeving V van a met de eigenschappen $y(a) = b$, $f(x, y(x)) = 0$ voor $x \in V$; voorts is $y(x)$ differentieerbaar op V met afgeleide $y' = -f_x/f_y$.

Voorbeelden. 2) Uit de vergelijking $x^2 + y^2 = 1$ volgt door impliciet differentiëren:

$$2x + 2yy' = 0, y' = -\frac{x}{y} \text{ mits } y \neq 0,$$

in overeenstemming met voorbeeld 1.

De tweede afgeleide van $y(x)$ volgt door de vergelijking $x + yy' = 0$ opnieuw impliciet te differentiëren:

$$1 + (y')^2 + yy'' = 0,$$

waaruit volgt

$$y'' = -\frac{1 + (y')^2}{y} = -\frac{1}{y^3} \text{ mits } y \neq 0.$$

3) Beschouw de vergelijking $x^3 + y^3 - 3xy = 0$.

Impliciet differentiëren naar x geeft

$$3x^2 + 3y^2y' - 3y - 3xy' = 0,$$

waaruit volgt

$$y' = \frac{y - x^2}{y^2 - x} \text{ mits } y^2 - x \neq 0.$$

$$\begin{aligned} x^3 + x^6 - 3x^2 &= 0 \\ x^2(x^2 - 2) &= 0 \end{aligned}$$

We onderzoeken nu de kromme beschreven door de vergelijking $x^3 + y^3 - 3xy = 0$. Horizontale raaklijnen kunnen optreden als $y - x^2 = 0$. Gecombineerd met de oorspronkelijke vergelijking levert dit de punten $(x,y) = (0,0)$ en $(x,y) = (\sqrt[3]{2}, \sqrt[3]{4})$. Het punt $(0,0)$ is een bijzonder punt omdat daar ook $y^2 - x = 0$.

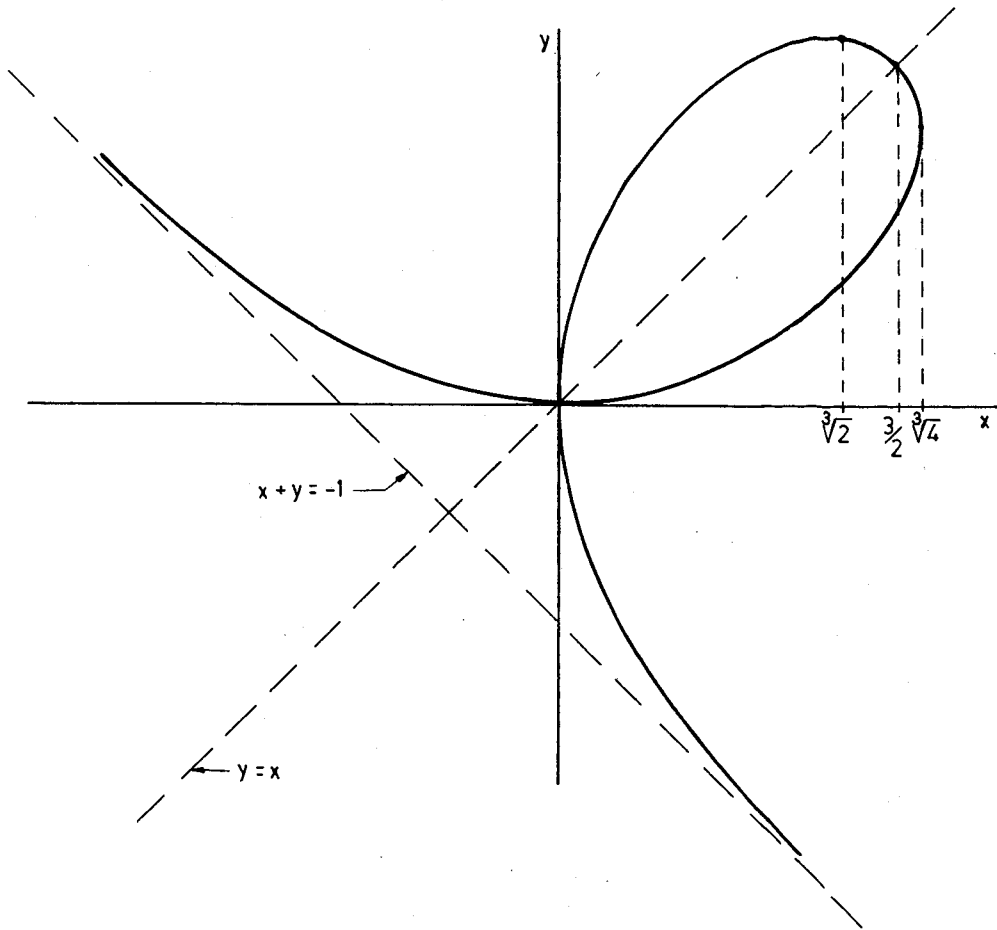
Verticale raaklijnen kunnen optreden in de punten $(0,0)$ en $(\sqrt[3]{4}, \sqrt[3]{2})$, zoals volgt uit $y^2 - x = 0$ en de oorspronkelijke vergelijking.

De kromme is symmetrisch t.o.v. de rechte $y = x$. We snijden de kromme met de rechte $x + y = p$. De x -coördinaten van de snijpunten worden bepaald door de vergelijking

$$3(p+1)x^2 - 3p(p+1)x + p^3 = 0.$$

Voor $p = -1$ is er geen snijpunt; voor $p = 0$ is $(0,0)$ een dubbel snijpunt; uit de discriminant $3p^2(p+1)(3-p)$ blijkt dat er twee snijpunten zijn voor $-1 < p \leq 3$ en geen snijpunten als $p < -1$ of $p > 3$.

De kromme heet het folium van Descartes; zie figuur. De kromme heeft een scheve asymptoot $x + y = -1$.



De methode van het differentiëren van impliciet gegeven functies kan uitgebreid worden tot meer variabelen en tot meer vergelijkingen (zie ook college Wiskunde 30).

De vergelijking $F(x,y,z) = 0$ beschrijft een oppervlak in \mathbb{R}^3 . Door deze vergelijking is z impliciet gegeven als functie van x en y : $z = z(x,y)$.

De partiële afgeleiden $\frac{\partial z}{\partial x}$ en $\frac{\partial z}{\partial y}$ kunnen we vinden door partiële differentiatie van

$$F(x,y,z(x,y)) = 0$$

onder toepassing van de kettingregel:

$$\frac{\partial F}{\partial x} + \frac{\partial F}{\partial z} \frac{\partial z}{\partial x} = 0, \quad \frac{\partial z}{\partial x} = - \frac{\partial F / \partial x}{\partial F / \partial z},$$

$$\frac{\partial F}{\partial y} + \frac{\partial F}{\partial z} \frac{\partial z}{\partial y} = 0, \quad \frac{\partial z}{\partial y} = - \frac{\partial F / \partial y}{\partial F / \partial z},$$

mits $\frac{\partial F}{\partial z} \neq 0$.

Voorbeeld 4. Beschouw de vergelijking $x - e^y \sin z = 0$.

Partieel differentiëren naar x en naar y geeft

$$1 - e^y \cos z \frac{\partial z}{\partial x} = 0,$$

$$-e^y \sin z - e^y \cos z \frac{\partial z}{\partial y} = 0,$$

waaruit volgt

$$\frac{\partial z}{\partial x} = \frac{1}{e^y \cos z}, \quad \frac{\partial z}{\partial y} = -\frac{\sin z}{\cos z} = -\tan z,$$

mits $\cos z \neq 0$.

Het stelsel vergelijkingen

$$\begin{cases} F(x,y,z) = 0 \\ G(x,y,z) = 0 \end{cases}$$

beschrijft een kromme in \mathbb{R}^3 .

Men kan y en z als impliciet gegeven functies van x beschouwen en vragen naar de afgeleiden $y'(x)$ en $z'(x)$.

Differentiatie naar x levert

$$\begin{cases} \frac{\partial F}{\partial x} + \frac{\partial F}{\partial y} \frac{dy}{dx} + \frac{\partial F}{\partial z} \frac{dz}{dx} = 0, \\ \frac{\partial G}{\partial x} + \frac{\partial G}{\partial y} \frac{dy}{dx} + \frac{\partial G}{\partial z} \frac{dz}{dx} = 0, \end{cases}$$

waaruit y' en z' zijn op te lossen mits de coëfficiëntendeterminant $\neq 0$ is.

Voorbeeld 5. Beschouw het stelsel vergelijkingen

$$\begin{cases} x^2 + y^2 + z^2 = 5, \\ xy + yz = 2. \end{cases}$$

Bereken $y'(x)$, $z'(x)$ en $z''(x)$ in $(x,y,z) = (0,1,2)$.

Differentiatie naar x levert

$$2x + 2yy' + 2zz' = 0 \quad \text{en} \quad y + (x+z)y' + yz' = 0.$$

In (0,1,2) wordt dit

$$2y'(0) + 4z'(0) = 0 \quad \text{en} \quad 1 + 2y'(0) + z'(0) = 0,$$

met oplossing $y'(0) = -2/3$ en $z'(0) = 1/3$.

Nogmaals differentiëren geeft

$$\begin{cases} 1 + (y')^2 + yy'' + (z')^2 + zz'' = 0, \\ y' + (x+z)y'' + (1+z')y' + y'z' + yz'' = 0. \end{cases}$$

Invullen van $(x,y,z) = (0,1,2)$ en van de bekende waarden van $y'(0)$ en $z'(0)$ geeft twee vergelijkingen voor $y''(0)$ en $z''(0)$, waaruit volgt $z''(0) = -44/27$.

6.6. Richtingsafgeleide en gradiënt

We brengen de definitie van het inwendig product $(\underline{a}, \underline{b})$ van twee vectoren \underline{a} en \underline{b} in herinnering. In \mathbb{R}^2 is $(\underline{a}, \underline{b}) = |\underline{a}| |\underline{b}| \cos \varphi = a_1 b_1 + a_2 b_2$, in \mathbb{R}^3 is $(\underline{a}, \underline{b}) = |\underline{a}| |\underline{b}| \cos \varphi = a_1 b_1 + a_2 b_2 + a_3 b_3$, waarbij φ de hoek tussen de vectoren is en $\underline{a} = (a_1, a_2)$ resp. (a_1, a_2, a_3) , $\underline{b} = (b_1, b_2)$ resp. (b_1, b_2, b_3) . Twee vectoren \underline{a} en \underline{b} staan loodrecht op elkaar dan en slechts dan als $(\underline{a}, \underline{b}) = 0$.

In \mathbb{R}^3 is $n_1(x-a) + n_2(y-b) + n_3(z-c) = 0$ de vergelijking van een vlak V door (a,b,c) met $\underline{n} = (n_1, n_2, n_3)$ als normaalvector, d.w.z. \underline{n} staat loodrecht op het vlak V . Dit laatste is in te zien door de vergelijking in de vorm $(\underline{n}, \underline{x}-\underline{a}) = 0$ te schrijven met $\underline{x} = (x,y,z)$ en $\underline{a} = (a,b,c)$.

In \mathbb{R}^2 is analoog $n_1(x-a) + n_2(y-b) = 0$ de vergelijking van een rechte ℓ door (a,b) met $\underline{n} = (n_1, n_2)$ als normaalvector, d.w.z. \underline{n} staat loodrecht op de rechte ℓ .

6.6.1. Gradiënt. Zij $f: \text{DOM } f \rightarrow \mathbb{R}$, $\text{DOM } f \subset \mathbb{R}^n$, een reële functie van n variabelen, differentieerbaar in $\underline{a} \in \mathbb{R}^n$ (zie 6.4), dan verstaan we onder $\text{grad } f(\underline{a})$ de rijvector

$$\left(\frac{\partial f}{\partial x_1}(\underline{a}), \frac{\partial f}{\partial x_2}(\underline{a}), \dots, \frac{\partial f}{\partial x_n}(\underline{a}) \right).$$

Opmerking. In plaats van $\text{grad } f$ gebruikt met ook de notatie ∇f .

Rekenregels. $\text{grad}(c f(\underline{a})) = c \text{ grad } f(\underline{a})$ (c is constante)

$$\text{grad}(f(\underline{a}) + g(\underline{a})) = \text{grad } f(\underline{a}) + \text{grad } g(\underline{a})$$

$$\text{grad}(f(\underline{a})g(\underline{a})) = f(\underline{a}) \text{ grad } g(\underline{a}) + g(\underline{a}) \text{ grad } f(\underline{a}).$$

6.6.2. Richtingsafgeleide. Zij f een reële functie van n variabelen, zij \underline{a} een inwendig punt van $\text{DOM } f$ en $\underline{x} = \underline{a} + t\underline{v}$ een parametervoorstelling van een rechte door \underline{a} met richtingsvector \underline{v} . Kies $|\underline{v}| = 1$, dan is $|t|$ de afstand van \underline{x} tot \underline{a} . We beschouwen de samengestelde functie $g(t) = f(\underline{a} + t\underline{v})$. Als g een afgeleide heeft in $t = 0$, m.a.w. als

$$g'(0) = \lim_{t \rightarrow 0} \frac{g(t) - g(0)}{t} = \lim_{t \rightarrow 0} \frac{f(\underline{a} + t\underline{v}) - f(\underline{a})}{t}$$

bestaat, dan noemen we $g'(0)$ de richtingsafgeleide van f in \underline{a} in de richting van \underline{v} .

Merk op dat $f_x(\underline{a})$, $f_y(\underline{a})$ en $f_z(\underline{a})$ richtingsafgeleiden zijn in de richting van de coördinaatassen.

6.6.3. Stelling. Zij f differentieerbaar in \underline{a} dan bestaat de richtingsafgeleide van f in \underline{a} in elke richting \underline{v} ($|\underline{v}| = 1$) en is gelijk aan $(\text{grad } f(\underline{a}), \underline{v})$.

Bewijs. Dit is een onmiddellijk gevolg van de kettingregel. □

Opgave. $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ zij gedefinieerd door

$$f(\underline{x}) := \begin{cases} \frac{xy}{x^2 + y^2} & \text{voor } \underline{x} \neq \underline{0}, \\ 0 & \text{voor } \underline{x} = \underline{0}. \end{cases}$$

a) Bewijs dat f_x en f_y overal bestaan.

b) Heeft f , behalve f_x en f_y , nog andere richtingsafgeleiden in $(0,0)$?

6.6.4. Meetkundige interpretatie. We beschouwen de differentieerbare functie $f(x,y,z)$.

Door het punt $\underline{a} = (a,b,c)$ gaat het niveauvlak met vergelijking

$$f(x,y,z) = f(a,b,c). \tag{1}$$

De vergelijking van het raakvlak aan het niveauvlak in het punt \underline{a} kunnen we opstellen door op te merken dat door de vergelijking (1) z impliciet als functie van x en y is gegeven. (We nemen aan dat $f_z(\underline{a}) \neq 0$ is; zie 6.5.) Volgens 6.3.2 is de vergelijking van het raakvlak

$$z - c = \frac{\partial z}{\partial x} (a,b)(x-a) + \frac{\partial z}{\partial y} (a,b)(y-b) . \quad (2)$$

Uit (1) bepalen we door impliciet differentiëren $\frac{\partial z}{\partial x}$ en $\frac{\partial z}{\partial y}$. Dit geeft

$$\frac{\partial z}{\partial x} (a,b) = - \frac{f_x(\underline{a})}{f_z(\underline{a})} , \quad \frac{\partial z}{\partial y} (a,b) = - \frac{f_y(\underline{a})}{f_z(\underline{a})} .$$

Invullen in (2) geeft tenslotte voor het raakvlak de vergelijking:

$$f_x(\underline{a})(x-a) + f_y(\underline{a})(y-b) + f_z(\underline{a})(z-c) = 0 .$$

Uit de vergelijking volgt dat de vector $\text{grad } f(\underline{a})$ loodrecht op het niveauvlak in \underline{a} staat. De richtingsafgeleide in \underline{a} in de richting \underline{v} is $(\text{grad } f(\underline{a}), \underline{v}) = |\text{grad } f(\underline{a})| \cos \varphi$, als φ de hoek is tussen \underline{v} en $\text{grad } f(\underline{a})$. De richtingsafgeleide is dus maximaal als $\varphi = 0$, d.w.z. in de richting van $\text{grad } f(\underline{a})$. De vector $\text{grad } f(\underline{a})$ geeft dus aan in welke richting f het sterkst toeneemt. Loodrecht op $\text{grad } f(\underline{a})$ is de richtingsafgeleide nul; dit geldt dus voor richtingen in het raakvlak aan het niveauvlak.

Niveauvlakken snijden elkaar uiteraard niet.

Krommen loodrecht op alle niveauvlakken heten orthogonale trajectoriën; de raaklijn aan de orthogonale trajectorie door \underline{a} heeft richtingsvector $\text{grad } f(\underline{a})$.

Voor $n = 2$ geldt een analoge meetkundige interpretatie. We hebben dan te maken met niveaulijnen in \mathbb{R}^2 in plaats van met niveauvlakken. De raaklijn in $\underline{a} = (a,b)$ aan de niveaulijn $f(x,y) = f(a,b)$ wordt gegeven door de vergelijking

$$f_x(\underline{a})(x-a) + f_y(\underline{a})(y-b) = 0 .$$

Voorbeeld. $f(x,y) = xy$. Niveaulijnen zijn orthogonale hyperbolen met x -as en y -as als asymptoten.

Wegens $\text{grad } f(x,y) = (y,x)$ geldt voor de orthogonale trajectoriën:

$$\left. \begin{array}{l} \frac{dx}{dt} = y \\ \frac{dy}{dt} = x \end{array} \right\} \frac{dy}{dx} = \frac{x}{y}$$

met als oplossingen $x^2 - y^2 = C$; dat zijn orthogonale hyperbolen met $y = x$ en $y = -x$ als asymptoten.

6.7. Extrema Δ conditie?

We beschouwen functies van twee of drie variabelen. Locale resp. globale extrema worden gedefinieerd als bij functies van één variabele (zie 2.5.6).

Een belangrijke rol speelt de volgende stelling, die we niet zullen bewijzen (zie 2.5.7).

6.7.1. Stelling (Weierstrass). Zij $V \subset \mathbb{R}^n$ een begrensde gesloten verzameling en zij $f : V \rightarrow \mathbb{R}$ continu op V , dan heeft f op V een globaal maximum en een globaal minimum.

Inwendige punten van $\text{DOM } f$ waar f differentieerbaar is en $\text{grad } f = \underline{0}$ heten stationaire punten.

6.7.2. Stelling. Als f een lokaal extremum heeft in een inwendig punt $\underline{a} \in \text{DOM } f$ en als f differentieerbaar is in \underline{a} , dan is $\text{grad } f(\underline{a}) = \underline{0}$, d.w.z. \underline{a} is een stationair punt.

Bewijs. Kies een richting \underline{v} . De samengestelde functie $g(t) := f(\underline{a} + t\underline{v})$ heeft in $t = 0$ een lokaal extremum en is daar differentieerbaar met $g'(0) = (\text{grad } f(\underline{a}), \underline{v})$. Volgens 2.6.5 is dan $(\text{grad } f(\underline{a}), \underline{v}) = 0$. Omdat dit geldt voor elke richting \underline{v} , is $\text{grad } f(\underline{a}) = \underline{0}$. □

Voorbeelden. 1) $f(x,y) = x^2 + y^2$; $\text{grad } f(x,y) = (2x, 2y)$.

Het enige stationaire punt is $(0,0)$. In dit punt heeft f een globaal minimum met functiewaarde 0.

2) $f(x,y) = \sqrt{x^2 + y^2}$ heeft geen stationaire punten, maar wel een globaal minimum in $(0,0)$. In $(0,0)$ is f niet differentieerbaar.

3) $f(x,y) = xy$; $\text{grad } f(x,y) = (y,x)$.

Het enige stationaire punt is $(0,0)$, maar f heeft in dat punt géén extremum. Immers $f(0,0) = 0$ en $f(x,y) > 0$ in het eerste en derde kwadrant en $f(x,y) < 0$ in het tweede en vierde kwadrant.

We noemen het punt $(0,0)$ een zadelpunt.

Samenvattend kan men zeggen dat een functie f op een verzameling V extrema kan hebben:

- a) in inwendige punten van V die stationaire punten van f zijn;
- b) in inwendige punten van V waar f niet differentieerbaar is;
- c) in randpunten van V .

Voorbeeld 4. Bepaal de extrema van $f(x,y) = 1 - x^2 - y^2$ op de verzameling $V := \{(x,y) \mid x^2 + y^2 \leq 1\}$.

V is begrensd en gesloten en f is continu, dus volgens 6.7.1 moet f een globaal maximum en een globaal minimum hebben op V . Inwendige punten zijn de punten (x,y) met $x^2 + y^2 < 1$. Omdat $\text{grad } f(x,y) = (-2x, -2y)$, is $(0,0)$ het enige stationaire punt met $f(0,0) = 1$. Voor de randpunten van V (de punten (x,y) met $x^2 + y^2 = 1$) is $f(x,y) = 0$ en voor de inwendige punten is $f(x,y) > 0$.

Conclusie. In elk punt van de rand heeft f een globaal minimum (waarde 0); er moet ook een globaal maximum zijn dat niet op de rand wordt aangenomen; er is slechts één stationair punt nl. $(0,0)$, dus in $(0,0)$ heeft f een globaal maximum (waarde 1).

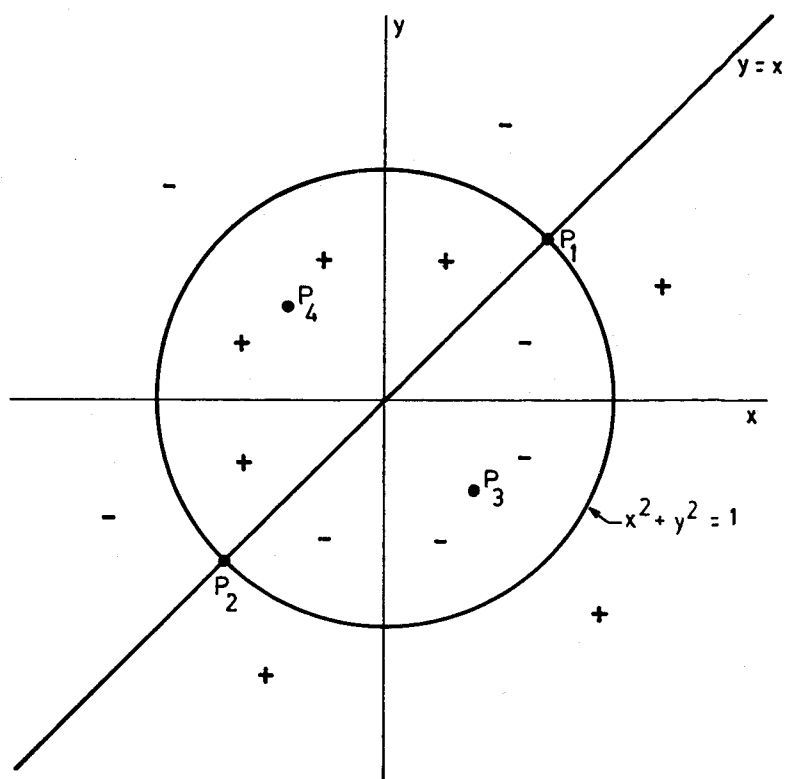
Voorbeeld 5. Bepaal de extrema van $f(x,y) = (x-y)(x^2+y^2-1)$ in \mathbb{R}^2 .

Nu is $\text{grad } f(x,y) = (3x^2 - 2xy + y^2 - 1, -x^2 + 2xy - 3y^2 + 1)$.

Er zijn vier stationaire punten:

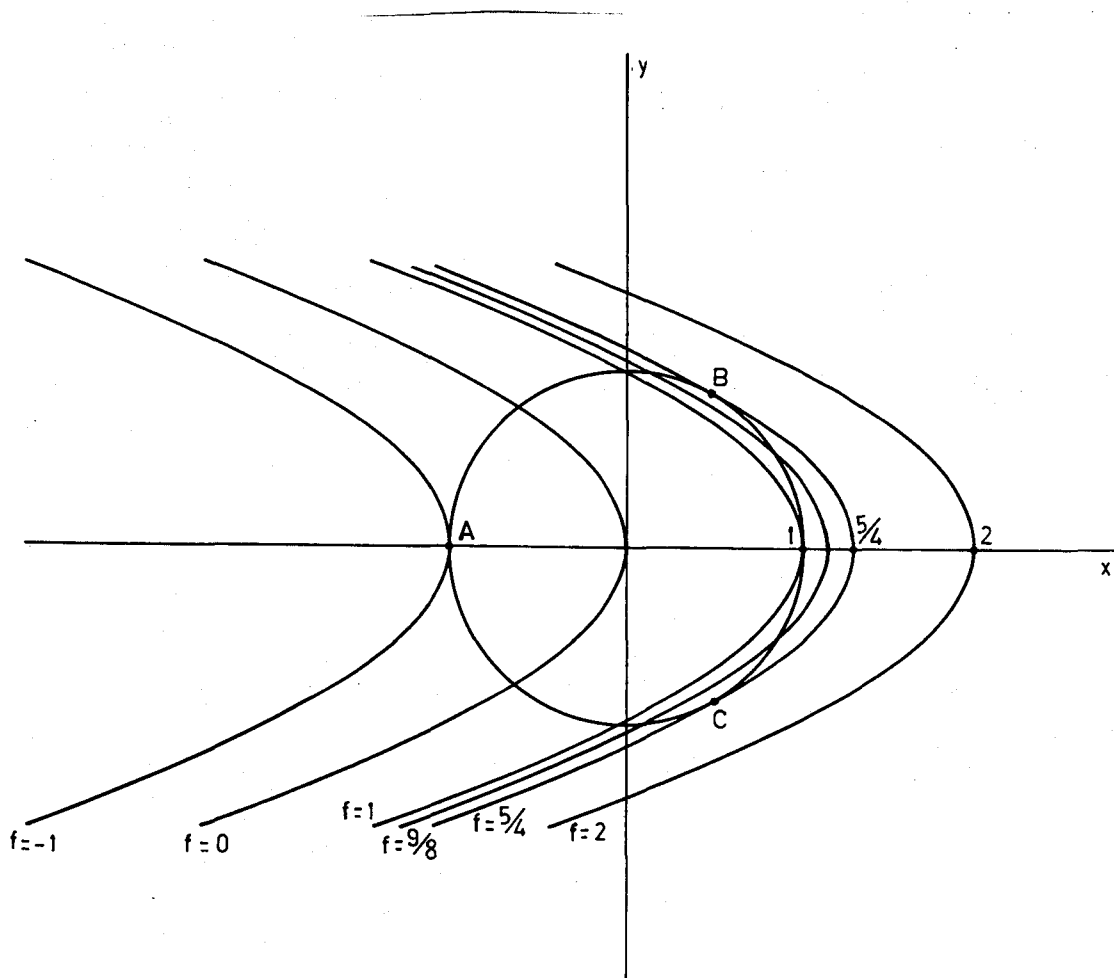
$$P_1 = \left(\frac{1}{2}\sqrt{2}, \frac{1}{2}\sqrt{2}\right), \text{ functiewaarde } 0; P_2 = \left(-\frac{1}{2}\sqrt{2}, -\frac{1}{2}\sqrt{2}\right), \text{ functiewaarde } 0;$$

$$P_3 = \left(\frac{1}{6}\sqrt{6}, -\frac{1}{6}\sqrt{6}\right), \text{ functiewaarde } -\frac{2}{9}\sqrt{6}; P_4 = \left(-\frac{1}{6}\sqrt{6}, \frac{1}{6}\sqrt{6}\right), \text{ functiewaarde } \frac{2}{9}\sqrt{6}.$$



Van de hoogtekaart zijn de nul-niveaulijnen (nullijnen) getekend, die het vlak in vier gebieden verdelen waar f tekenvast is. De tekenverdeling is ook aangegeven in de hoogtekaart. Hieruit blijkt onmiddellijk dat P_1 en P_2 zadelpunten zijn. Beschouw $V := \{(x,y) \mid x \leq y \wedge x^2 + y^2 \leq 1\}$. Op V is $f(x,y) \geq 0$ en op de rand van V is $f(x,y) = 0$. Aangezien V begrensd en gesloten is moet f dus een maximum hebben in het inwendige van V ; de enige kandidaat is P_4 , dus in P_4 heeft f een globaal maximum t.o.v. V . Ten opzichte van de gehele \mathbb{R}^2 heeft f in P_4 slechts een lokaal maximum. Evenzo heeft f in P_3 een lokaal minimum.

Voorbeeld 6. Bepaal de extrema van $f(x,y) = x + y^2$ op de verzameling $V := \{(x,y) \mid x^2 + y^2 \leq 1\}$. Beschouw eerst de inwendige punten van V , dat zijn punten (x,y) met $x^2 + y^2 < 1$. Daar $\text{grad } f(x,y) = (1, 2y) \neq (0,0)$ zijn er géén stationaire punten in het inwendige van V . Aangezien V begrensd en gesloten is en f continu, moeten zowel het globale maximum als het globale minimum op de rand gezocht worden. We tekenen de hoogtekaart.



We zien direkt dat het globale minimum van f de waarde -1 heeft en aangenomen wordt in $(-1,0)$. Om het globale maximum te bepalen zoeken we een getal c zó dat de parabool $x+y^2 = c$ raakt aan de cirkel $x^2+y^2 = 1$. De voorwaarde waaraan c moet voldoen is dan dat de vierkantsvergelijking $x^2 - x + c = 1$ een tweevoudige wortel heeft. Met de discriminant volgt dan $1 - 4(c - 1) = 0$; $c = 5/4$. De tweevoudige wortel is $x = \frac{1}{2}$. Het globale maximum wordt dus aangenomen in de punten $(\frac{1}{2}, \pm\frac{1}{2}\sqrt{3})$; de waarde is $5/4$.

We merken op dat het probleem om de extrema te bepalen van $f(x,y) = x+y^2$ op de verzameling $W := \{(x,y) \mid x^2+y^2 = 1\}$, dus op de rand van V , een andere oplossing heeft. Op W geldt $y^2 = 1 - x^2$ en $-1 \leq x \leq 1$; substitutie in $f(x,y)$ geeft $f(x,y(x)) = -x^2 + x + 1$ te beschouwen voor $-1 \leq x \leq 1$. Deze functie heeft voor $x = \frac{1}{2}$ een globaal maximum met waarde $5/4$. Er zijn randminima voor $x = -1$, waarde -1 , en ook voor $x = 1$, waarde 1 . De functie $f(x,y) = x+y^2$ heeft dus op W de volgende extrema: globale minima met waarde $5/4$ in $(\frac{1}{2}, \frac{1}{2}\sqrt{3})$ en $(\frac{1}{2}, -\frac{1}{2}\sqrt{3})$;

*Samengevat. had je
in product gedaan
orthogonale proj. hebbe zo wel gedaan
(te maken met D.V.)*

een globaal minimum met waarde -1 in $(-1,0)$; een lokaal minimum met waarde $+1$ in $(1,0)$. Als we de extrema van $f(x,y)$ op V bepalen dan voldoet het lokale minimum niet; op de doorsnede van de x -as en V , d.w.z. op $\{(x,0) \mid -1 \leq x \leq 1\}$, heeft f in $(1,0)$ een globaal maximum want $f(x,0) = x \leq 1 = f(1,0)$.

Als we dus van een functie de extrema op de rand van V bepaald hebben (bijv. door eliminatie zoals in dit voorbeeld of met de methode uit 6.8), dan kunnen we niet blindelings deze extrema ook tot extrema van f op V verklaren.

6.8. Extrema onder nevenvoorwaarden

Beschouw het volgende probleem. Bepaal de extrema van $f(x,y)$ op de verzameling van punten (x,y) die voldoen aan $g(x,y) = 0$.

De volgende stelling geeft een nodige voorwaarde voor een extremum.

Stelling. Als f een extremum heeft onder de nevenvoorwaarde $g(x,y) = 0$ in een punt \underline{a} waar f en g differentieerbaar zijn en waar $\text{grad } g(\underline{a}) \neq \underline{0}$, dan is er een getal λ zodat geldt

$$\text{grad } f(\underline{a}) + \lambda \text{ grad } g(\underline{a}) = \underline{0}.$$

We bewijzen de stelling niet.

Meetkundig is de stelling als volgt plausibel te maken. De nevenvoorwaarde $g(x,y) = 0$ is de vergelijking van een kromme K in \mathbb{R}^2 . Heeft f in het punt $\underline{a} \in K$ een extremum, dan raakt de niveaulijn van f in \underline{a} aan de kromme K . De raaklijnen vallen dan samen evenals de normalen en dat betekent dat de vector $\text{grad } f(\underline{a})$ een veelvoud van $\text{grad } g(\underline{a})$ is.

Men gebruikt de stelling aldus.

Los x , y en λ op uit de drie vergelijkingen

$$\begin{cases} g(x,y) = 0, \\ f_x(x,y) + \lambda g_x(x,y) = 0, \\ f_y(x,y) + \lambda g_y(x,y) = 0. \end{cases}$$

Is (x_0, y_0, λ_0) een oplossing van dit stelsel vergelijkingen, dan kan er in (x_0, y_0) een extremum optreden. In punten waar $\text{grad } g = \underline{0}$ is de methode niet toepasbaar; zulke punten komen dus ook voor extrema in aanmerking en dienen apart onderzocht te worden.

De methode heet multiplicatorenmethode van Lagrange (λ heet de multiplicator).

Voorbeeld. Bepaal de extrema van $f(x,y) = x + y^2$ onder de nevenvoorwaarde $x^2 + y^2 = 1$.

Los op

$$\begin{cases} x^2 + y^2 - 1 = 0 \\ 1 + 2\lambda x = 0 \\ 2y + 2\lambda y = 0 \end{cases} \rightarrow 2y(\lambda + 1) = 0 \begin{cases} \lambda = -1 \rightarrow x = \frac{1}{2} \rightarrow y = \pm \frac{\sqrt{3}}{2} \\ y = 0 \rightarrow x = \pm 1 \rightarrow \lambda = \pm \frac{1}{2} \end{cases}$$

Oplissing:

$$\lambda = -\frac{1}{2}; (x,y) = (1,0); f(1,0) = 1$$

$$\lambda = \frac{1}{2}; (x,y) = (-1,0); f(-1,0) = -1$$

$$\lambda = -1; (x,y) = (\frac{1}{2}, \pm \frac{\sqrt{3}}{2}); f(\frac{1}{2}, \pm \frac{\sqrt{3}}{2}) = 5/4.$$

Met behulp van de stelling van Weierstrass concluderen we tot globale maxima in $(\frac{1}{2}, \pm \frac{\sqrt{3}}{2})$, een globaal minimum in $(-1,0)$ en een lokaal minimum in $(1,0)$.

Men kan de multiplicatorenmethode uitbreiden tot functies van meer variabelen en meer nevenvoorwaarden.

Voorbeeld. Bepaal de extrema van $f(x,y,z)$ onder de nevenvoorwaarden $g(x,y,z) = 0$ en $h(x,y,z) = 0$.

Dit leidt tot het oplossen van x, y, z, λ_1 en λ_2 uit het stelsel van vijf vergelijkingen

$$\begin{cases} g = 0, h = 0, \\ f_x + \lambda_1 g_x + \lambda_2 h_x = 0, \\ f_y + \lambda_1 g_y + \lambda_2 h_y = 0, \\ f_z + \lambda_1 g_z + \lambda_2 h_z = 0. \end{cases}$$

Hoofdstuk 7. Lineaire algebra

7.1. Voorwoord

Lineaire algebra is het deel van de wiskunde dat gegroeid is uit de theorie van de stelsels van eerstegraads vergelijkingen in meerdere onbekenden. Wiskundige beschrijvingen met behulp van eerstegraads (ook wel genoemd lineaire) betrekkingen worden in zeer veel gebieden van onderzoek gebruikt. Om die reden worden methoden en resultaten van de lineaire algebra toegepast in talrijke uiteenlopende gebieden van wetenschap. De grondslag van het succes van de lineaire algebra is de synthese van twee geheel verschillende benaderingswijzen, nl. de rekenkundige en de (abstract) meetkundige. In de eerste benadering is de lineaire algebra een techniek ontwikkeld om doelmatig en systematisch te rekenen met stelsels lineaire betrekkingen; in de meetkundige benadering worden de objecten uit de lineaire algebra voorgesteld als meetkundige grootheden, waardoor ze een aanschouwelijke inhoud krijgen. In dit hoofdstuk komen beide gezichtspunten naar voren; eerst het rekenkundige en geleidelijk aan ook het abstract meetkundige. Beschouwingen over stelsels vergelijkingen lopen als een draad door het hele hoofdstuk.

7.2. Bewerkingen met matrices

7.2.1. Inleiding

Een matrix is een rechthoekig schema van getallen.

Voorbeelden. $\begin{bmatrix} 1 & 4 & 4 & x \\ 2 & 1 & 2 & a \\ 3 & 0 & 0 & 6 \end{bmatrix}$; $\begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}$; $[3,5,0,4,7]$.

We spreken van rijen, kolommen, elementen. De volgende notaties voor een matrix met r rijen en k kolommen zijn gebruikelijk:

$$A = A^{(r \times k)} = [a_{ij}] = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1k} \\ a_{21} & a_{22} & \dots & a_{2k} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ a_{r1} & a_{r2} & \dots & a_{rk} \end{bmatrix} .$$

Is A een matrix dan geeft men soms het element in de i-de rij en j-de kolom aan met A_{ij} . Als de duidelijkheid er om vraagt gebruike men komma's:

$[3,5,0,4,7]$; $a_{i+1,j+1}$. Is A een matrix met r rijen en k kolommen dan zeggen we dat A het formaat r bij k heeft, of dat A een $(r \times k)$ -matrix is.

Vanuit het rekenkundige standpunt zijn matrices niet meer dan een hulpmiddel om ingewikkelde algebraïsche en numerieke berekeningen te systematiseren. We zullen voor matrices bewerkingen definiëren die generalisaties zijn van de bewerkingen met reële getallen (een reëel getal is op te vatten als een matrix met één rij en één kolom).

De inhoud van deze paragraaf is nu als volgt samen te vatten: Eerst definiëren we de som en het verschil van twee matrices (7.2.2), het product van een matrix en een reëel getal (7.2.2) en het product van een matrix en een matrix (7.2.3). In 7.2.4 worden deze bewerkingen geïllustreerd aan een boekhoudkundig voorbeeld. De nummers 7.2.5 en 7.2.6 geven enkele begrippen uit de matrixrekening. In de daarna komende passages 7.2.7 en 7.2.8 wordt aandacht geschonken aan matrices van speciale formaten. Tenslotte geven we in 7.2.9 een eerste reeks beschouwingen over stelsels lineaire vergelijkingen.

7.2.2. Optelling, aftrekking en scalaire vermenigvuldiging

Definitie. Als A en B matrices zijn van hetzelfde formaat dan zijn de som $A+B$ en het verschil $A-B$ gedefinieerd door

$$(A + B)_{ij} := A_{ij} + B_{ij} , \quad (A - B)_{ij} := A_{ij} - B_{ij} .$$

Als A een matrix is en $p \in \mathbb{R}$ dan is de matrix pA gedefinieerd door:

$$(pA)_{ij} := pA_{ij} .$$

Voorbeelden.

$$\begin{bmatrix} 1 & 0 & x & 1 \\ 2 & 0 & y & 3 \\ 3 & 0 & z & 1 \end{bmatrix} + \begin{bmatrix} 4 & 1 & 0 & 1 \\ 5 & 2 & 3 & 1 \\ 0 & a & 1 & 0 \end{bmatrix} = \begin{bmatrix} 5 & 1 & x & 2 \\ 7 & 2 & y+3 & 4 \\ 3 & a & z+1 & 1 \end{bmatrix},$$

$$7 \begin{bmatrix} 1 & 0 & x \\ 2 & 4 & y \\ 1 & 0 & z \end{bmatrix} = \begin{bmatrix} 7 & 0 & 7x \\ 14 & 28 & 7y \\ 7 & 0 & 7z \end{bmatrix}.$$

We schrijven $-A$ voor $-1A$. Duidt \emptyset een matrix aan waarvan alle elementen nul zijn (een zgn. nulmatrix) dan is

$$0A^{(r \times k)} = \emptyset = \emptyset^{(r \times k)}.$$

Eigenschappen. Bij matrixbewerkingen wordt steeds verondersteld dat de formaten zo zijn dat de bewerkingen gedefinieerd zijn.

$$(A + B) + C = [A_{ij} + B_{ij} + C_{ij}] = A + (B + C)$$

$$A + B = B + A$$

$$A + \emptyset = \emptyset + A = A$$

$$A - A = \emptyset$$

$$(p + q)A = pA + qA$$

$$p(A + B) = pA + pB$$

$$p(qA) = (pq)A.$$

7.2.3. Het matrixproduct

Van twee matrices A en B zullen we het product in de volgorde AB alleen definiëren indien het aantal kolommen van A gelijk is aan het aantal rijen van B .

Is $A^{(m \times n)} = [a_{ij}]$, $B^{(n \times p)} = [b_{ij}]$, dan is $AB := C = [c_{ij}]$ een $(m \times p)$ -matrix gedefinieerd door

$$c_{ik} := \sum_{j=1}^n a_{ij} b_{jk}, \quad i = 1, \dots, m; k = 1, \dots, p.$$

Voorbeeld.

$$\begin{matrix}
 \begin{bmatrix} * & * & * & * & * & * \\ * & * & * & * & * & * \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} & a_{36} \\ * & * & * & * & * & * \end{bmatrix} &
 \begin{bmatrix} * & b_{12} & * \\ * & b_{22} & * \\ * & b_{32} & * \\ * & b_{42} & * \\ * & b_{52} & * \\ * & b_{62} & * \end{bmatrix} &
 = &
 \begin{bmatrix} * & * & * \\ * & * & * \\ * & c_{32} & * \\ * & * & * \end{bmatrix} \\
 A^{(4 \times 6)} & B^{(6 \times 3)} & & C^{(4 \times 3)}
 \end{matrix}$$

$$c_{32} = a_{31}b_{12} + a_{32}b_{22} + a_{33}b_{32} + a_{34}b_{42} + a_{35}b_{52} + a_{36}b_{62} .$$

Nog een voorbeeld:

$$[a_1, a_2] \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} = [a_1 b_1 + a_2 b_2] = a_1 b_1 + a_2 b_2$$

(bij een (1×1) -matrix, d.i. een getal, laten we de haken weg);

$$\begin{bmatrix} b_1 \\ b_2 \end{bmatrix} [a_1, a_2] = \begin{bmatrix} b_1 a_1 & b_1 a_2 \\ b_2 a_1 & b_2 a_2 \end{bmatrix} .$$

Men ziet dat de factoren in een matrixproduct niet verwisseld mogen worden. Bij $A^{(4 \times 6)}$ en $B^{(6 \times 3)}$ is AB wel maar BA niet gedefinieerd. Is A vierkant (d.w.z. heeft A evenveel rijen als kolommen) dan is AA gedefinieerd; we schrijven $AA =: A^2$.

Opmerkingen.

$$1) \quad \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 2 \\ 0 & -1 \end{bmatrix} = \mathcal{O} = \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 2 & -4 \\ -1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 2 & -4 \\ -1 & 2 \end{bmatrix} .$$

Uit $PA = PB$ volgt niet $A = B$. Evenmin volgt uit $CQ = DQ$ dat $C = D$. Uit $QR = \mathcal{O}$ volgt niet $Q = \mathcal{O}$ of $R = \mathcal{O}$.

2) Als A en B beide $(p \times p)$ -matrices zijn, dan is

$$(A + B)(A - B) = A^2 - AB + BA - B^2$$

en omdat AB in het algemeen niet gelijk is aan BA , is dit niet verder te vereenvoudigen.

Eigenschappen.

$$\left. \begin{aligned} A(B + C) &= AB + AC \\ (A + B)C &= AC + BC \end{aligned} \right\} \text{(distributiviteit)}$$

$$(AB)C = A(BC) . \quad \text{(associativiteit)}$$

7.2.4. Voorbeeld. Een steenfabriek fabriceert stenen in vijf kwaliteiten. De fabriek heeft vier magazijnen met verkooppunten. Voorraadgegevens worden nu handig geadmistreerd met (4×5) -matrices waarbij de rijen overeenkomen met de magazijnen, de kolommen met de kwaliteiten.

Laat A de aanwezige voorraad op 1 januari 1974 voorstellen:

$$\begin{array}{l} \text{magazijn 1} \rightarrow \\ 2 \rightarrow \\ 3 \rightarrow \\ 4 \rightarrow \end{array} \left[\begin{array}{ccccc} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} \end{array} \right]$$

$$\begin{array}{c} \uparrow \\ \uparrow \\ \uparrow \\ \uparrow \\ \uparrow \end{array} \begin{array}{c} \text{kwaliteit 1} \\ 2 \\ 3 \\ 4 \\ 5 \end{array}$$

(a_{34} is het aantal op 1-1-1974 aanwezige stenen in magazijn 3 van kwaliteit 4). Als $B^{(4 \times 5)}$ de door de fabriek aan de verkooppunten geleverde hoeveelheden en $C^{(4 \times 5)}$ de verkoop gedurende 1974 voorstelt, dan is $A + B - C$ de voorraadmatrix per 1-1-1975.

Laten we aannemen dat grondstofkosten, fabricagekosten en verkoopprijs afhangen van de kwaliteit. Laat g_1, g_2, \dots, g_5 de grondstofkosten, f_1, f_2, \dots, f_5 de fabricagekosten en v_1, v_2, \dots, v_5 de verkoopprijs per steen zijn. Laat

$$\begin{bmatrix} a_{11} & a_{12} & * & * & a_{15} \\ * & * & * & * & * \\ * & * & * & * & * \\ a_{41} & * & * & * & a_{45} \end{bmatrix} \begin{bmatrix} g_1 & f_1 & v_1 \\ g_2 & f_2 & v_2 \\ g_3 & f_3 & v_3 \\ g_4 & f_4 & v_4 \\ g_5 & f_5 & v_5 \end{bmatrix} =: \begin{bmatrix} p_{11} & p_{12} & p_{13} \\ * & * & * \\ * & * & * \\ p_{41} & * & p_{43} \end{bmatrix}$$

$$A \quad \quad \quad V \quad \quad =: \quad \quad P .$$

Dan stelt de k-de rij van P voor achtereenvolgens de in de voorraad van magazijn k geïnvesteerde grondstofkosten (= p_{k1}), de geïnvesteerde fabricagekosten (= p_{k2}), de verkoopwaarde (= p_{k3}).

Merk op dat de kolommen van A en de rijen van V corresponderen met het aspect kwaliteit en dat de elementen van $P = AV$ sommen zijn over alle kwaliteiten. Iets dergelijks treft men steeds aan in boekhoudkundig gebruik van het matrixproduct.

7.2.5. Blokmatrices

Vaak is het handig een matrix opgebouwd te denken uit kleinere matrices; bijvoorbeeld

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} & \cdot & \cdot & a_{17} \\ a_{21} & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{31} & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{41} & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ a_{51} & \cdot & \cdot & \cdot & \cdot & \cdot & a_{57} \end{bmatrix} =: \begin{bmatrix} P & Q & R \\ S & T & U \end{bmatrix},$$

waarbij

$$P := \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \dots, U := \begin{bmatrix} a_{34} & a_{35} & a_{36} & a_{37} \\ a_{44} & \cdot & \cdot & \cdot \\ a_{54} & \cdot & \cdot & a_{57} \end{bmatrix}.$$

Als alle formaten van overeenkomstige blokken overeenstemmen kan men matrices optellen door de blokken op te tellen:

$$\begin{bmatrix} A_1 & B_1 \\ C_1 & D_1 \end{bmatrix} + \begin{bmatrix} A_2 & B_2 \\ C_2 & D_2 \end{bmatrix} = \begin{bmatrix} A_1 + A_2 & B_1 + B_2 \\ C_1 + C_2 & D_1 + D_2 \end{bmatrix}.$$

Een analoge formule geldt voor vermenigvuldiging van blokmatrices.

7.2.6. De getransponeerde matrix

Is A een $(r \times k)$ -matrix dan is de getransponeerde A^T een $(k \times r)$ -matrix die uit A ontstaat door rijen en kolommen te verwisselen. Dus $(A^T)_{ij} = A_{ji}$.

Is

$$A = \begin{bmatrix} 1 & 6 \\ 2 & 7 \\ 3 & 8 \\ 4 & 9 \\ 5 & 0 \end{bmatrix}$$

dan is

$$A^T = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 6 & 7 & 8 & 9 & 0 \end{bmatrix} .$$

Eigenschappen.

$$(A + B)^T = A^T + B^T ; \quad (AB)^T = B^T A^T .$$

7.2.7. Vectoren

Matrices met slechts één rij of slechts één kolom noemt men vectoren, (rijvectoren, kolomvectoren). We noteren kolomvectoren met onderstreepte kleine letters: a. Een rijvector kan men dan aangeven met a^T.

In dit hoofdstuk is \mathbb{R}^n de verzameling van alle kolomvectoren met n elementen. Anders dan in andere delen van de wiskunde (en van deze syllabus) is het in de lineaire algebra van belang het onderscheid tussen rijen en kolommen strikt te handhaven.

In \mathbb{R}^n is door de definities uit 7.2.2 een optelling u + v en een scalaire vermenigvuldiging pu gedefinieerd:

$$\begin{bmatrix} u_1 \\ u_2 \\ \cdot \\ \cdot \\ u_n \end{bmatrix} + \begin{bmatrix} v_1 \\ v_2 \\ \cdot \\ \cdot \\ v_n \end{bmatrix} = \begin{bmatrix} u_1 + v_1 \\ u_2 + v_2 \\ \cdot \\ \cdot \\ u_n + v_n \end{bmatrix} ; \quad p \begin{bmatrix} u_1 \\ u_2 \\ \cdot \\ \cdot \\ u_n \end{bmatrix} = \begin{bmatrix} pu_1 \\ pu_2 \\ \cdot \\ \cdot \\ pu_n \end{bmatrix} .$$

Een belangrijk begrip in \mathbb{R}^n is het inwendig product van twee vectoren, gedefinieerd door:

$$(\underline{u}, \underline{v}) := \underline{u}^T \underline{v} = u_1 v_1 + u_2 v_2 + \dots + u_n v_n = \sum_{i=1}^n u_i v_i .$$

Eenvoudig ziet men in dat:

$$(\underline{u}, \underline{v}) = (\underline{v}, \underline{u}) ;$$

$$(\underline{u}, \underline{v} + \underline{w}) = (\underline{u}, \underline{v}) + (\underline{u}, \underline{w}) ;$$

$$(p\underline{u}, \underline{v}) = p(\underline{u}, \underline{v}) \quad (p \in \mathbb{R}) ;$$

$(\underline{u}, \underline{u}) \geq 0$ en uit $(\underline{u}, \underline{u}) = 0$ volgt $\underline{u} = \underline{0}$, waarin

$$\underline{0} := \begin{bmatrix} 0 \\ 0 \\ \cdot \\ \cdot \\ 0 \end{bmatrix},$$

de nulvector is.

Het inwendig product zal nog op vele plaatsen opduiken (7.2.9, 7.4.3, 7.4.5, 7.8.5 enz.); in \mathbb{R}^2 en \mathbb{R}^3 kennen we het al.

Is \underline{a}_i^T de i -de rij van een matrix A , \underline{b}_j de j -de kolom van een matrix B dan is $(AB)_{ij} = (\underline{a}_i, \underline{b}_j)$.

Als A een $(r \times k)$ -matrix is, $\underline{x} \in \mathbb{R}^k$, $\underline{y} \in \mathbb{R}^r$, dan geldt:

$$(\underline{A}\underline{x}, \underline{y}) = (\underline{A}\underline{x})^T \underline{y} = \underline{x}^T \underline{A}^T \underline{y} = \underline{x}^T (\underline{A}^T \underline{y}) = (\underline{x}, \underline{A}^T \underline{y}).$$

7.2.8. Enkele soorten vierkante matrices

Hoewel 7.2.8 op dit moment de functie heeft een aantal voorbeelden van matrixvermenigvuldiging te laten zien, zullen we de geïntroduceerde begrippen: eenheidsmatrix, inverse matrix, diagonaalmatrix en permutatiematrix nog vaak tegenkomen.

7.2.8.1. Eenheidsmatrix

$$I^{(n \times n)} := \begin{bmatrix} 1 & 0 & 0 & \cdot & \cdot & 0 \\ 0 & 1 & 0 & \cdot & \cdot & 0 \\ 0 & 0 & 1 & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \cdot & \cdot & 1 \end{bmatrix}$$

d.w.z.

$$I_{ij} := \delta_{ij} := \begin{cases} 1 & \text{als } i = j, \\ 0 & \text{als } i \neq j. \end{cases}$$

(δ_{ij} heet het symbool van Kronecker; Kronecker-delta).

Er geldt:

$$A^{(m \times n)} I^{(n \times n)} = A^{(m \times n)},$$

$$I^{(n \times n)} B^{(n \times p)} = B^{(n \times p)}.$$

7.2.8.2. Inverse van een vierkante matrix

Laat A een $(n \times n)$ -matrix zijn. Als er een matrix B bestaat zodat $BA = I^{(n \times n)}$, dan heet B de inverse van A, notatie A^{-1} . Later zullen we bewijzen dat dan ook $AA^{-1} = I$ en dat een vierkante matrix ten hoogste één inverse heeft.

Eén voorbeeld: als we de inverse willen bepalen van

$$\begin{bmatrix} 3 & 2 \\ 7 & 5 \end{bmatrix}$$

moeten we x, y, u en v vinden zó dat

$$\begin{bmatrix} x & y \\ u & v \end{bmatrix} \begin{bmatrix} 3 & 2 \\ 7 & 5 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Of:

$$\begin{cases} 3x + 7y = 1 \\ 2x + 5y = 0 \end{cases} \quad \begin{cases} 3u + 7v = 0 \\ 2u + 5v = 1 \end{cases}.$$

We vinden: $x = 5$, $y = -2$, $u = -7$, $v = 3$, en dus

$$\begin{bmatrix} 3 & 2 \\ 7 & 5 \end{bmatrix}^{-1} = \begin{bmatrix} 5 & -2 \\ -7 & 3 \end{bmatrix}.$$

Merk op dat inderdaad

$$\begin{bmatrix} 3 & 2 \\ 7 & 5 \end{bmatrix} \begin{bmatrix} 5 & -2 \\ -7 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}.$$

Als een matrix A een inverse heeft dan heet A inverteerbaar.

7.2.8.3. Diagonaalmatrix

$D^{(n \times n)} = [d_{ij}]$ heet diagonaalmatrix indien $d_{ij} = 0$ als $i \neq j$. Rechtsvermenigvuldiging met een diagonaalmatrix betekent vermenigvuldiging van de kolommen, linksvermenigvuldiging betekent vermenigvuldiging van de rijen:

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} p & 0 & 0 \\ 0 & q & 0 \\ 0 & 0 & r \end{bmatrix} = \begin{bmatrix} pa_{11} & qa_{12} & ra_{13} \\ pa_{21} & qa_{22} & ra_{23} \\ pa_{31} & qa_{32} & ra_{33} \end{bmatrix},$$

$$\begin{bmatrix} p & 0 & 0 \\ 0 & q & 0 \\ 0 & 0 & r \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \begin{bmatrix} pa_{11} & pa_{12} & pa_{13} \\ qa_{21} & qa_{22} & qa_{23} \\ ra_{31} & ra_{32} & ra_{33} \end{bmatrix}.$$

Als van een $(n \times n)$ -diagonaalmatrix D alle diagonaalelementen $D_{ii} \neq 0$ zijn dan is D inverteerbaar en D^{-1} is een diagonaalmatrix met

$$(D^{-1})_{ii} = \frac{1}{D_{ii}} \quad (i = 1, \dots, n).$$

7.2.8.4. Permutatiematrices

Veranderingen van de volgorde van de rijen of kolommen van een matrix kan men verkrijgen door links- of rechtsvermenigvuldiging met zgn. permutatiematrices. We geven de algemene definitie maar beperken ons verder tot een voorbeeld van een (3×3) -matrix. Zij f een bijectie van $\{1, 2, \dots, n\}$ in $\{1, 2, \dots, n\}$ (f heet ook wel een permutatie van $\{1, 2, \dots, n\}$, zie 1.3, dan is de permutatiematrix P behorende bij f gedefinieerd door:

$$P_{ij}^{(n \times n)} := \begin{cases} 1 & \text{als } i = f(j), \\ 0 & \text{als } i \neq f(j), \end{cases} \quad \text{of} \quad P_{ij}^{(n \times n)} = \delta_{i, f(j)} \quad (i, j = 1, \dots, n);$$

in elke rij en kolom van P staat dus precies één 1.

Laat nu een permutatie f van $(1, 2, 3)$ gegeven zijn door $f(1) = 3$, $f(2) = 1$, $f(3) = 2$, dan is de permutatiematrix

$$P = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}.$$

Het is duidelijk dat $\underline{x} \mapsto A\underline{x}$ een afbeelding $\mathbb{R}^n \rightarrow \mathbb{R}^m$ is. Het oplossen van het stelsel $A\underline{x} = \underline{b}$ betekent het bepalen van alle vectoren, die door deze afbeelding op \underline{b} worden afgebeeld.

Op de volgende uiterst belangrijke eigenschappen van de afbeelding $\underline{x} \mapsto A\underline{x}$ zullen we vanuit verschillende gezichtspunten nog meermalen terugkomen.

✓ 1) Is A een $(m \times n)$ -matrix dan volgt uit $\underline{x} \mapsto A\underline{x}$, $\underline{y} \mapsto A\underline{y}$, $p \in \mathbb{R}$ dat:

$$\underline{x} + \underline{y} \mapsto A\underline{x} + A\underline{y} ,$$

$$p\underline{x} \mapsto pA\underline{x} .$$

X 2) Is A een $(m \times n)$ -matrix en B een $(p \times m)$ -matrix, dan geldt voor de samengestelde afbeelding van de $(\mathbb{R}^n \rightarrow \mathbb{R}^m)$ -afbeelding $\underline{x} \mapsto A\underline{x}$ en de $(\mathbb{R}^m \rightarrow \mathbb{R}^p)$ afbeelding $\underline{y} \mapsto B\underline{y}$:

$$\underline{x} \mapsto B(A\underline{x}) = (BA)\underline{x} .$$

We keren terug tot vergelijkingen en beschouwen

$$\{\underline{x} \in \mathbb{R}^n \mid A\underline{x} = \underline{0} \in \mathbb{R}^m\} =: N(A) ,$$

d.i. de verzameling van de oplossingen van het zgn. homogene stelsel vergelijkingen $A\underline{x} = \underline{0}$. $N(A)$ heet de nulruimte van A en ook wel de oplossingsruimte van $A\underline{x} = \underline{0}$. We zien dat $N(A) \neq \emptyset$ is, omdat $\underline{0} \in N(A)$. Uit 1) volgt: als $\underline{u} \in N(A)$ en $\underline{v} \in N(A)$, $p \in \mathbb{R}$, dan is ook $\underline{u} + \underline{v} \in N(A)$, $p\underline{u} \in N(A)$.

Voor een inhomogeen stelsel vergelijkingen $A\underline{x} = \underline{b}$ ($\underline{b} \in \mathbb{R}^m$, $\underline{b} \neq \underline{0}$) geldt nu: als $\underline{a} \in \mathbb{R}^n$ een oplossing is (d.w.z. $A\underline{a} = \underline{b}$) dan is de algemene oplossing

$$\underline{a} + N(A) := \{\underline{a} + \underline{x} \mid \underline{x} \in N(A)\} .$$

De analogie met hetgeen in 5.3.1, 5.4.2.1 over differentiaalvergelijkingen gezegd is ligt erg voor de hand.

Als A $(n \times n)$ een inverse heeft, dan kunnen we uit $A\underline{a} = \underline{b}$ concluderen: $A^{-1}A\underline{a} = A^{-1}\underline{b}$, $I_n\underline{a} = A^{-1}\underline{b}$, $\underline{a} = A^{-1}\underline{b}$. Is A vierkant en inverteerbaar, dan heeft het stelsel $A\underline{x} = \underline{b}$ voor elke $\underline{b} \in \mathbb{R}^n$ precies één oplossing, nl. $A^{-1}\underline{b}$; i.h.b. is dan $N(A) = \{\underline{0}\}$.

Met behulp van de begrippen uit deze paragraaf kunnen we het stelsel (1) nog op twee andere manieren schrijven. Elk van deze schrijfwijzen zal later een eigen meetkundige interpretatie krijgen.

Zij

$$\underline{a}_1 := \begin{bmatrix} a_{11} \\ a_{21} \\ \cdot \\ \cdot \\ a_{m1} \end{bmatrix}, \quad \underline{a}_2 := \begin{bmatrix} a_{12} \\ a_{22} \\ \cdot \\ \cdot \\ a_{m2} \end{bmatrix}, \quad \dots, \quad \underline{a}_n := \begin{bmatrix} a_{1n} \\ a_{2n} \\ \cdot \\ \cdot \\ a_{mn} \end{bmatrix},$$

dan kunnen we (1) ook schrijven als

$$x_1 \underline{a}_1 + x_2 \underline{a}_2 + \dots + x_n \underline{a}_n = \underline{b}.$$

Zij

$$\underline{c}_1^T := [a_{11}, a_{12}, \dots, a_{1n}], \quad \underline{c}_2^T := [a_{21}, a_{22}, \dots, a_{2n}], \quad \dots, \\ \underline{c}_m^T := [a_{m1}, a_{m2}, \dots, a_{mn}],$$

dan is (1) te schrijven als:

$$\begin{cases} (\underline{c}_1, \underline{x}) = b_1 \\ (\underline{c}_2, \underline{x}) = b_2 \\ \vdots \\ (\underline{c}_m, \underline{x}) = b_m. \end{cases}$$

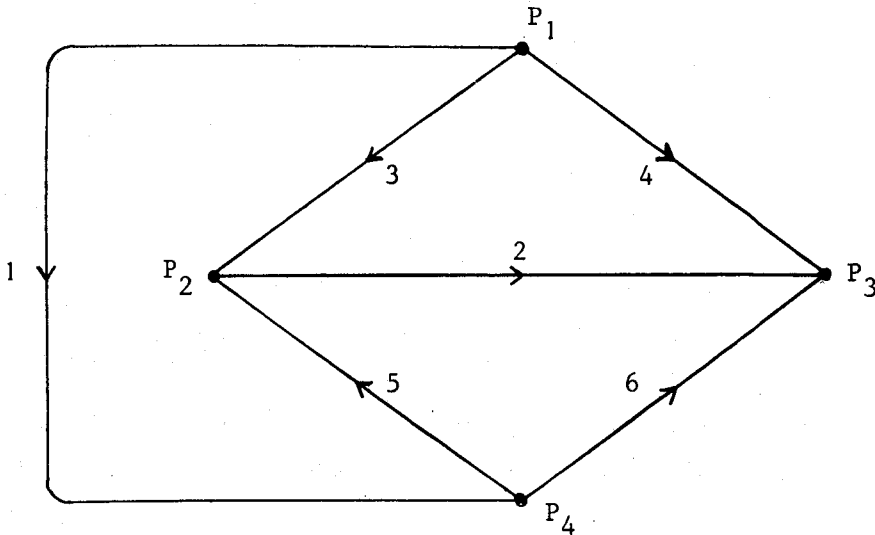
7.3. Voorbeelden en toepassingen

7.3.1. Inleiding

In deze paragraaf worden voorbeelden gegeven van het gebruik van matrixrekening in een aantal toepassingsgebieden. Deze voorbeelden dienen als motivering voor de begrippen die in 7.2 zijn ingevoerd en voor beschouwingen uit latere paragrafen. Er wordt echter niet naar terugverwezen zodat men een of meerdere van deze voorbeelden desnoods kan overslaan.

7.3.2. Een transportnetwerk

Vier steden P_1 , P_2 , P_3 en P_4 bezitten ieder een gasfabriek. Er liggen zes pijpleidingen, 1 t/m 6, volgens onderstaand plaatje, die we willekeurig van een richting voorzien.



(Een dergelijk systeem bestaande uit eindig veel punten met gerichte verbindingen heet een gerichte graaf.)

De prijs van een m^3 gas in P_k is w_k . Uit P_k wordt per tijdseenheid u_k m^3 gas uitgevoerd ($u_k < 0$ betekent invoer); het prijsverschil per m^3 tussen eindpunt en beginpunt van pijp l is v_l , bijv. $v_6 = w_3 - w_4$. Langs pijpleiding l wordt per tijdseenheid t_l m^3 gas vervoerd ($t_l < 0$ betekent dat het transport tegen de pijlrichting in geschiedt). We nemen aan dat geen gas verloren gaat.

Zij

$$\underline{w} := \begin{bmatrix} w_1 \\ w_2 \\ w_3 \\ w_4 \end{bmatrix}, \quad \underline{u} := \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{bmatrix}, \quad \underline{t} := \begin{bmatrix} t_1 \\ t_2 \\ t_3 \\ t_4 \\ t_5 \\ t_6 \end{bmatrix}, \quad \underline{v} := \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \\ v_6 \end{bmatrix}.$$

De onderlinge ligging van de steden en pijpleidingen wordt beschreven door een matrix $A^{(4 \times 6)} := [a_{ij}]$ waarvan

$$a_{ij} := \begin{cases} 1 & \text{als } P_i \text{ beginpunt is van pijp } j; \\ -1 & \text{als } P_i \text{ eindpunt is van pijp } j; \\ 0 & \text{anders.} \end{cases}$$

Dus

$$A = \begin{bmatrix} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & -1 & 0 & -1 & 0 \\ 0 & -1 & 0 & -1 & 0 & -1 \\ -1 & 0 & 0 & 0 & 1 & 1 \end{bmatrix}$$

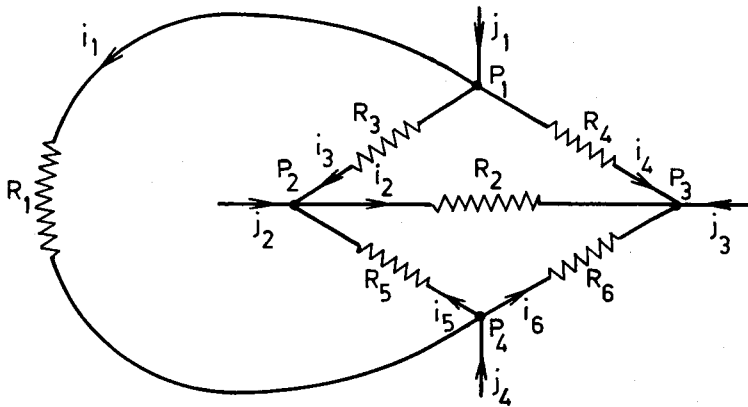
In iedere kolom komt een +1 en een -1 voor, immers pijp j heeft een beginpunt en een eindpunt.

De samenhang tussen \underline{u} , \underline{t} , \underline{v} en \underline{w} wordt nu gegeven door

$$\underline{u} = A\underline{t} ; \quad \underline{v} = -A^T \underline{w} .$$

7.3.3. Een elektrisch netwerk

Beschouw het volgende netwerk.



We nummeren de knopen P_1, P_2, P_3, P_4 ; de takken $1, 2, \dots, 6$. Laat stromen i_1, \dots, i_6 volgens pijlrichting door de draden lopen, waarbij volgens de gebruikelijke tekenafspraken $i_t < 0$ betekent dat de stroom in tak t tegen de pijlrichting in loopt.

De meetkundige structuur van het netwerk wordt beschreven door de matrix $A^{(4 \times 6)} := [a_{kt}]$ waarbij

$$a_{kt} := \begin{cases} 1 & \text{als knoop } k \text{ beginpunt is van tak } t, \\ -1 & \text{als knoop } k \text{ eindpunt is van tak } t, \\ 0 & \text{anders.} \end{cases}$$

Deze matrix is dezelfde als in het vorige voorbeeld.

Laat

$$\underline{w} := \begin{bmatrix} w_1 \\ w_2 \\ w_3 \\ w_4 \end{bmatrix} \quad \text{en} \quad \underline{j} := \begin{bmatrix} j_1 \\ j_2 \\ j_3 \\ j_4 \end{bmatrix},$$

waarbij w_k de spanning in P_k en j_k de in P_k van buiten binnenkomende stroom is. Zij voorts:

$$\underline{v} := \begin{bmatrix} v_1 \\ v_2 \\ \cdot \\ \cdot \\ v_6 \end{bmatrix}, \quad \underline{i} := \begin{bmatrix} i_1 \\ i_2 \\ \cdot \\ \cdot \\ i_6 \end{bmatrix}, \quad R := \begin{bmatrix} R_1 & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & R_2 & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & \cdot & \cdot & R_6 \end{bmatrix},$$

waarbij v_t het spanningsverschil over tak t , R_t de weerstand in tak t voorstelt.

De wet van Ohm in matrixvorm luidt nu: $\underline{v} = R\underline{i}$.

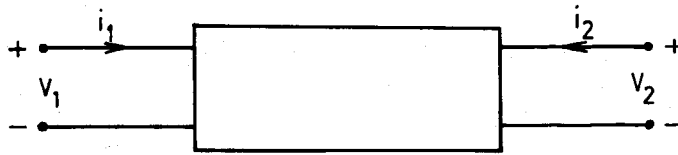
Voorts is de stroomwet van Kirchhoff in matrixvorm: $\underline{j} = A\underline{i}$.

Voor de spanningen geldt: $\underline{v} = A^T \underline{w}$. We vinden $\underline{j} = AR^{-1}A^T \underline{w}$.

Vergelijk dit voorbeeld met 7.3.2; let vooral op de tekenconventies in beide voorbeelden.

7.3.4. Kettingmatrices van een tweepoort

Beschouw een tweepoort dat is een elektrisch netwerk met twee ingangen en twee uitgangen, waarbij het verband tussen de ingangsgrootheden i_1 en V_1 en de uitgangsgrootheden i_2 en V_2 lineair is.



Er zijn verschillende methoden om de afhankelijkheden tussen de stromen i_1 en i_2 en de spanningsverschillen V_1 en V_2 uit te drukken:

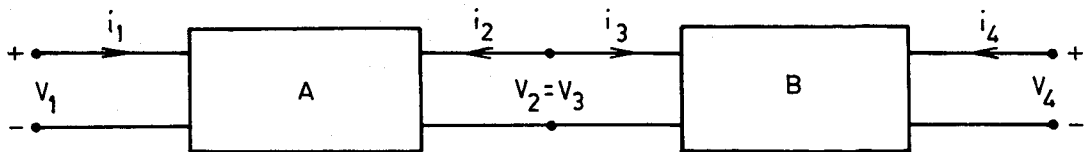
$$\begin{bmatrix} V_1 \\ V_2 \end{bmatrix} = \begin{bmatrix} Z_{11} & Z_{12} \\ Z_{21} & Z_{22} \end{bmatrix} \begin{bmatrix} i_1 \\ i_2 \end{bmatrix} \quad \underline{\text{impedantiematrix}}$$

$$\begin{bmatrix} i_1 \\ i_2 \end{bmatrix} = \begin{bmatrix} Y_{11} & Y_{12} \\ Y_{21} & Y_{22} \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \end{bmatrix} \quad \underline{\text{admittantiematrix}}$$

Merk op dat de admittantiematrix de inverse is van de impedantiematrix. We zullen in dit voorbeeld gebruik maken van zgn. kettingmatrices:

$$\begin{bmatrix} V_1 \\ i_1 \end{bmatrix} = \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} V_2 \\ -i_2 \end{bmatrix} .$$

Deze kettingmatrices hebben de volgende fraaie eigenschap:



Schakelt men een tweepoort met kettingmatrix B achter een tweepoort met kettingmatrix A, dan is het resultaat weer een tweepoort nu met kettingmatrix AB. Immers $V_2 = V_3$ en $-i_2 = i_3$ zodat

$$\begin{bmatrix} V_1 \\ i_1 \end{bmatrix} = A \begin{bmatrix} V_2 \\ -i_2 \end{bmatrix} = A \begin{bmatrix} V_3 \\ i_3 \end{bmatrix} = AB \begin{bmatrix} V_4 \\ -i_4 \end{bmatrix} .$$

$$\begin{bmatrix} V_3 \\ i_3 \end{bmatrix} = B \begin{bmatrix} V_4 \\ -i_4 \end{bmatrix}$$

$$\begin{bmatrix} V_1 \\ i_1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ R_1^{-1} & 1 \end{bmatrix} \begin{bmatrix} V_2 \\ -i_2 \end{bmatrix} \quad \text{of} \quad \begin{matrix} V_1 = V_2 \\ i_1 = V_2 R_1^{-1} - i_2 \end{matrix} \quad \text{of} \quad \begin{matrix} i_1 + i_2 = V_2 R_1^{-1} \\ (i_1 + i_2) R_1 = V_2 \end{matrix}$$

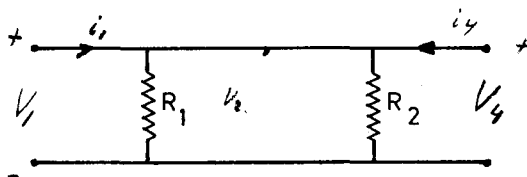
- 53 -

Uit het feit dat

$$\begin{bmatrix} 1 & 0 \\ R_1^{-1} & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ R_2^{-1} & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ R_1^{-1} + R_2^{-1} & 1 \end{bmatrix}$$

de kettingmatrix is van de tweepoort bestaande uit twee parallel geschakelde weerstanden volgt de bekende formule voor de substitutieweerstand

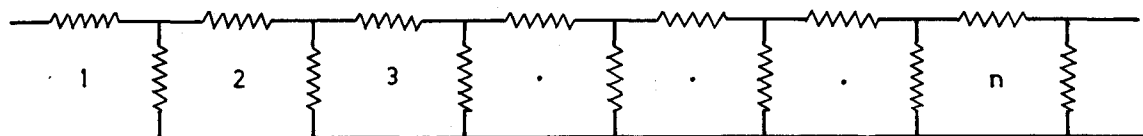
$$\frac{1}{R} = \frac{1}{R_1} + \frac{1}{R_2} .$$



De kettingmatrix van onderstaand laddernetwerk met n mazen en alle weerstanden \$1\Omega\$ is

$$\begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}^n .$$

$$\begin{bmatrix} V_1 \\ i_1 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} V_2 \\ -i_2 \end{bmatrix} \quad \text{of} \quad \begin{matrix} V_1 = 2V_2 - i_2 \\ i_1 = V_2 - i_2 \end{matrix}$$



Door volledige inductie bewijst men

$$\begin{bmatrix} 2 & 1 \\ 1 & 1 \end{bmatrix}^n = \begin{bmatrix} d_{2n+1} & d_{2n} \\ d_{2n} & d_{2n-1} \end{bmatrix} \quad n = 1, 2, \dots ,$$

waarbij \$d_k\$ de in 1.2 ingevoerde getallen van Fibonacci voorstellen.

7.3.5. Evenwicht van de prijsbalans

Een n -tal producenten P_1, \dots, P_n produceren elk één product g_1, \dots, g_n . P_i koopt en verbruikt per jaar een deel, nl. a_{ij} , van de jaarproductie van het product g_j . We nemen aan dat het systeem van producenten en producten gesloten is, d.w.z. dat de gehele productie aan g_1, \dots, g_n door P_1, \dots, P_n wordt verbruikt en dat niets van buiten het systeem wordt toegevoegd. (Deze aanname is iets minder streng dan op het eerste gezicht lijkt omdat we altijd kunnen veronderstellen dat een overproductie aan g_i door P_i wordt opgenomen en dus verrekend wordt in a_{ii} .)

Voor de matrix $A = [a_{ij}]$ geldt nu:

$$a_{ij} \geq 0 \quad (i, j = 1, \dots, n) ; \quad (1)$$

$$\sum_{i=1}^n a_{ij} = 1 \quad (j = 1, \dots, n) \quad (2)$$

(P_1, \dots, P_n gebruiken per jaar de hele geproduceerde hoeveelheid g_j).

Een matrix A die voldoet aan (1) en (2) heet ruilmatrix. Laat w_j de prijs van de jaarproductie aan g_j zijn. De uitgave van P_i is dan

$$\sum_{j=1}^n a_{ij} w_j ;$$

de inkomst van P_i is w_i (interpreteer dit goed als $a_{ii} > 0$).

Beschouw de prijsvector $\underline{w} = [w_1, w_2, \dots, w_n]^T$. De notatie $\underline{w} \geq \underline{0}$ betekent $w_i \geq 0$ ($i = 1, \dots, n$). Het productiesysteem met ruilmatrix A heet in evenwicht als er een prijsvector $\underline{w} \geq \underline{0}$, $\underline{w} \neq \underline{0}$ bestaat zó dat geen van de producenten P_1, \dots, P_n genoodzaakt is meer uit te geven dan hij ontvangt. Evenwicht betekent dus dat er een $\underline{w} \geq \underline{0}$, $\underline{w} \neq \underline{0}$ bestaat zó dat $\underline{w} - A\underline{w} \geq \underline{0}$ of:

$$A\underline{w} \leq \underline{w} . \quad (3)$$

Een vector \underline{w} die aan (3) voldoet heet een stabiele prijsvector van A . Het lijkt alsof dit een ongelijkheid is; dit is echter niet zo, want voor iedere \underline{w} die aan $A\underline{w} \leq \underline{w}$ voldoet is $A\underline{w} = \underline{w}$. Stel nl.

$$A\underline{w} \leq \underline{w} \quad \text{en} \quad w_k > \sum_{j=1}^n a_{kj} w_j \quad \text{voor zekere } k \in \{1, 2, \dots, n\} ,$$

dan is

$$\sum_{i=1}^n w_i > \sum_{i=1}^n \left(\sum_{j=1}^n a_{ij} w_j \right) = \sum_{j=1}^n w_j \sum_{i=1}^n a_{ij} = \sum_{j=1}^n w_j,$$

hetgeen onmogelijk is. Een veel kortere manier om het bovenstaande bewijs weer te geven is de volgende: zij $\underline{j} = [1, 1, \dots, 1]^T$; de voorwaarde (2) kan nu geschreven worden als $\underline{j}^T A = \underline{j}^T$. Als nu $A\underline{w} \leq \underline{w}$, $A\underline{w} \neq \underline{w}$ zou zijn, dan zou $\underline{j}^T A\underline{w} < \underline{j}^T \underline{w}$ zijn; evenwel $\underline{j}^T A\underline{w} = \underline{j}^T \underline{w}$, tegenspraak.

De stabiele prijsvectoren zijn dus oplossingen $\underline{w} \geq \underline{0}$, $\underline{w} \neq \underline{0}$ van $A\underline{w} = \underline{w}$ of $(I - A)\underline{w} = \underline{0}$. (Een vector $\underline{w} \neq \underline{0}$ waarvoor $A\underline{w} = \underline{w}$ heet een eigenvector van A bij eigenwaarde 1, zie 7.8.3.) Als \underline{w} een oplossing is en $r \geq 0$, dan is ook $r\underline{w}$ een oplossing; voor evenwicht is alleen de onderlinge verhouding der prijzen van belang. Men kan bewijzen dat er bij iedere A die aan (1) en (2) voldoet, een $\underline{w} \geq \underline{0}$, $\underline{w} \neq \underline{0}$ bestaat met $A\underline{w} = \underline{w}$.

7.3.6. Markov-processen, een bevolkingsprobleem

In een land trekt elk jaar $2\frac{1}{2}$ % van de plattelandsbevolking naar de stad en vestigt zich 1% van de stadsbevolking op het platteland. In het jaar $t = 0$ woont 60% van de bevolking in de stad, 40% op het land. Hoe is de situatie 1 jaar later en 10 jaar later en op den duur, aangenomen dat zowel op het platteland als in de stad de bevolking niet door andere oorzaken verandert (de geboortes en sterftes compenseren elkaar).

Zij

$$\underline{x}_t := \begin{bmatrix} s_t \\ \ell_t \end{bmatrix}$$

waarbij s_t het gedeelte van de totale bevolking, N, is dat in jaar t in de stad woont en ℓ_t het gedeelte dat op het land woont. Dan is:

$$s_1 N = 0.99 s_0 N + 0.025 \ell_0 N$$

$$\ell_1 N = 0.01 s_0 N + 0.975 \ell_0 N.$$

Met

$$P := \begin{bmatrix} 0.99 & 0.025 \\ 0.01 & 0.975 \end{bmatrix}$$

staat hier: $\underline{x}_1 = P\underline{x}_0$.

Hieruit volgt $\underline{x}_{10} = P^{10} \underline{x}_0$. De matrix P heet de overgangsmatrix van het proces. (Merk op dat een dergelijke matrix in 7.3.5 ruilmatrix genoemd werd.) Als we een oplossing $\underline{x} \neq \underline{0}$ van $P\underline{x} = \underline{x}$ (eigenvector bij eigenwaarde 1) kunnen vinden, dan beschrijft deze een stationaire verdeling. Schrijf

$$P := \begin{bmatrix} 1-p & q \\ p & 1-q \end{bmatrix}$$

en noem de gezochte oplossing van $P\underline{x} = \underline{x}$: $\underline{x} = [s, \ell]^T$ met $s + \ell = 1$, dan vinden we $\underline{x} = \frac{1}{p+q} [q, p]^T$. In de stationaire toestand woont dus het $\frac{q}{p+q} = \frac{5}{7}$ -de deel van de bevolking in de stad.

7.3.7. Een productieproces

Een drinkwaterfabriek maakt ruw grondwater geschikt voor levering aan een waterleidingbedrijf door het viermaal een periode van biologische zuivering te laten ondergaan. Deze vier perioden zijn even lang. Het bedrijf beschikt over vier bassins, genummerd 1,2,3,4, waarin achtereenvolgens de eerste, tweede, derde en vierde zuivering plaatsvindt. Alle vier bassins zijn gelijktijdig in bedrijf. Na afloop van elke zuiveringsperiode wordt water overgepompt van bassin 4 naar het waterleidingbedrijf, van bassin 3 naar 4, van 2 naar 3 en van 1 naar 2. Bij het overpompen van water uit bassin i ($i = 1,2,3,4$) blijft een vast deel, d , van de inhoud in bassin i achter, een vast deel, b , komt in de volgende productiefase, en de rest, $1-b-d$, gaat bij het overpompen verloren. Tenslotte wordt voordat de nieuwe periode ingaat, $N \text{ m}^3$ onbewerkt grondwater in bassin 1 gebracht. In de vector $\underline{s}(t) = [s_1(t), s_2(t), s_3(t), s_4(t)]^T$ stelt $s_i(t)$ het aantal m^3 water voor, dat zich gedurende periode t in bassin i bevindt. Aan het einde van periode t wordt dan $bs_4(t) \text{ m}^3$ aan het waterleidingbedrijf geleverd.

Zij

$$\underline{v} := \begin{bmatrix} N \\ 0 \\ 0 \\ 0 \end{bmatrix}, \quad A := \begin{bmatrix} d & 0 & 0 & 0 \\ b & d & 0 & 0 \\ 0 & b & d & 0 \\ 0 & 0 & b & d \end{bmatrix},$$

dan is $\underline{s}(t+1) = \underline{v} + A\underline{s}(t)$.

We willen berekenen hoe groot de inhoud van de bassins moet zijn opdat deze elke periode dezelfde is. Dit betekent dat we een $\underline{s} \in \mathbb{R}^4$ dienen te bepalen die voldoet aan $\underline{s} = \underline{v} + A\underline{s}$. Anders geschreven: $(I - A)\underline{s} = \underline{v}$. We kunnen dit stelsel vergelijkingen direkt oplossen en vinden dan

$$\underline{s} = N \left[\frac{1}{1-d}, \frac{b}{(1-d)^2}, \frac{b^2}{(1-d)^3}, \frac{b^3}{(1-d)^4} \right]^T.$$

We zullen dit resultaat nog langs andere weg afleiden. Als we over de inverse van $I - A$ zouden beschikken, zou de oplossing zijn: $\underline{s} = (I - A)^{-1} \underline{v}$. Nu is

$$(I - A)(I + A + A^2 + \dots + A^n) = I - A^{n+1}.$$

Deze betrekking doet aan machtreeksen denken. We definiëren een limietbegrip voor matrices aldus: $\lim_{n \rightarrow \infty} B_n = B$ betekent dat alle B_n en B hetzelfde formaat hebben en dat $\lim_{n \rightarrow \infty} (B_n)_{ij} = B_{ij}$ voor alle i en j .

We zullen nu bewijzen dat de matrix-machtreeks $I + A + A^2 + A^3 + \dots$ convergeert. Het is dan duidelijk dat de som de inverse is van $I - A$.

Zij

$$Z := \begin{bmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix},$$

dan is

$$Z^2 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}, \quad Z^3 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

en $Z^4 = \mathcal{O}$. Nu geldt $A = dI + bZ$; $A^2 = d^2I + 2dbZ + b^2Z^2$ (bedenk dat $IZ = ZI = Z$); $A^3 = d^3I + 3d^2bZ + 3db^2Z^2 + b^3Z^3$. Omdat $Z^k = \mathcal{O}$ voor $k \geq 4$, is algemeen volgens het binomium van Newton (1.3.7):

$$A^n = d^n I + \binom{n}{1} d^{n-1} b Z + \binom{n}{2} d^{n-2} b^2 Z^2 + \binom{n}{3} d^{n-3} b^3 Z^3 \quad \text{voor } n \geq 3.$$

Hieruit volgt:

$$\sum_{n=0}^{\infty} A^n = \left(\sum_{n=0}^{\infty} d^n \right) I + \left(\sum_{n=1}^{\infty} \binom{n}{1} d^{n-1} \right) bZ + \left(\sum_{n=2}^{\infty} \binom{n}{2} d^{n-2} \right) b^2 Z^2 + \\ + \left(\sum_{n=3}^{\infty} \binom{n}{3} d^{n-3} \right) b^3 Z^3 ,$$

waarbij $A^0 := I$. Omdat $0 < d < 1$ zijn de machtreeks $\sum_{n=0}^{\infty} d^n$ en haar afgeleiden naar d convergent (zie 3.4.3 en haar gevolg) en er geldt

$$\sum_{n=0}^{\infty} d^n = (1-d)^{-1} , \quad \sum_{n=k}^{\infty} \binom{n}{k} d^{n-k} = (1-d)^k .$$

Hieruit volgt:

$$\sum_{n=0}^{\infty} A^n = (1-d)^{-1} I + (1-d)^{-2} bZ + (1-d)^{-3} b^2 Z^2 + (1-d)^{-4} b^3 Z^3 = \\ = \begin{bmatrix} (1-d)^{-1} & 0 & 0 & 0 \\ (1-d)^{-2} b & (1-d)^{-1} & 0 & 0 \\ (1-d)^{-3} b^2 & (1-d)^{-2} b & (1-d)^{-1} & 0 \\ (1-d)^{-4} b^3 & (1-d)^{-3} b^2 & (1-d)^{-2} b & (1-d)^{-1} \end{bmatrix} = (I - A)^{-1} .$$

7.4. Analytische meetkunde in \mathbb{R}^3 en meetkundige terminologie in \mathbb{R}^n met $n > 3$

7.4.1. Inleiding

In deze paragraaf worden toepassingen van de matrixrekening in de analytische meetkunde besproken. Hoewel we de elementen van \mathbb{R}^2 en \mathbb{R}^3 in dit hoofdstuk als kolommen noteren handhaven we de opvatting dat het de cartesische coördinaten van punten in vlak en ruimte zijn, zie 1.5. Bovendien vatten we a niet alleen op als een punt maar ook als de verbindingspijl van 0 naar dit punt. We herhalen deels aan de hand van uitgewerkte voorbeelden, een gedeelte van de analytische meetkunde in de ruimte. Men herhale voor zichzelf de eenvoudigere resultaten in het vlak.

In 7.4.5 voeren we ook voor \mathbb{R}^n met $n > 3$ een meetkundige terminologie in. Het zal blijken dat deze n-dimensionale meetkunde een nuttige intuïtieve steun is bij vele formele algebraïsche bewerkingen en resultaten. Deze paragraaf is de directe aanleiding tot de abstract meetkundige beschouwingen waaraan het grootste deel van dit hoofdstuk gewijd zal zijn.

7.4.2. Lijnen en vlakken

Bij gegeven vectoren $\underline{v} \neq \underline{0}$ en \underline{a} in \mathbb{R}^3 is $\underline{x} = [x,y,z]^T = \underline{a} + \lambda \underline{v}$, een parametervoorstelling van een rechte door \underline{a} (steunvector) met richting \underline{v} (richtingsvector).

Zijn \underline{v} en \underline{w} twee vectoren zó dat de punten $\underline{0}$ (de oorsprong), \underline{v} en \underline{w} niet op een rechte liggen, dan is

$$\underline{x} = \underline{a} + \lambda \underline{v} + \mu \underline{w}$$

een parametervoorstelling van een vlak door \underline{a} evenwijdig aan het vlak door $\underline{0}$, \underline{v} en \underline{w} .

Een vlak wordt ook gekarakteriseerd door een vergelijking $ax + by + cz = d$; uit een parametervoorstelling vindt men de vergelijking door λ en μ te elimineren. Merk op dat $\underline{x} = \underline{a} + \lambda \underline{v} + \mu \underline{w}$ drie lineaire betrekkingen zijn, nl.

$$\begin{cases} x = a_1 + \lambda v_1 + \mu w_1 \\ y = a_2 + \lambda v_2 + \mu w_2 \\ z = a_3 + \lambda v_3 + \mu w_3 \end{cases} \quad \text{als} \quad \underline{a} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}, \quad \underline{v} = \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}, \quad \underline{w} = \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix}.$$

Het werken met parametervoorstellingen wordt nu geïllustreerd aan zes voorbeelden.

Voorbeeld 1. Beschouw de rechte $\underline{x} = [1,1,1]^T + [0,1,2]^T$.

Het punt $[1,3,5]^T$ ligt op de rechte (neem $\lambda = 2$). Het snijpunt met het (x,y)-vlak (vergelijking $z = 0$) vinden we uit: $0 = z = 1 + 2\lambda$, $\lambda = -\frac{1}{2}$; snijpunt $[1, \frac{1}{2}, 0]^T$. De rechte snijdt het (y,z)-vlak ($x=0$) niet, omdat (de verbindingslijn van $\underline{0}$ en) $[0,1,2]^T$ evenwijdig is aan het vlak $x = 0$.

Voorbeeld 2. Beschouw het vlak

$$\underline{x} = [1,2,3]^T + \lambda [0,1,1]^T + \mu [1,0,-2]^T.$$

Het snijpunt met de y-as vinden we uit:

$$\begin{cases} 0 = x = 1 + \mu & \mu = -1 \\ 0 = z = 3 + \lambda - 2\mu & \lambda = -5 \end{cases}; \quad \text{snijpunt } [0, -3, 0]^T.$$

De vergelijking van het vlak is $2x - y + z = 3$.

De snijlijn met het (x,y)-vlak wordt gegeven door het paar vergelijkingen:

$$\begin{cases} z = 0 \\ 2x - y + z = 3 \end{cases}$$

of ook door

$$\begin{cases} z = 0 \\ 2x - y = 3 . \end{cases}$$

Deze snijlijn heeft bijv. als parametervoorstelling (neem $x = \lambda$, dan is $y = -3 + 2\lambda$, $z = 0$);

$$\underline{x} = [0, -3, 0]^T + \lambda [1, 2, 0]^T .$$

Voorbeeld 3. Het snijpunt S van de rechte $\underline{x} = [1, 1, 1]^T + \lambda [0, 1, 2]^T$ met het vlak $2x - y + z = 3$ vinden we door λ op te lossen uit

$$2(1) - (1 + \lambda) + (1 + 2\lambda) = 3 .$$

We vinden $\lambda = 1$, $S = [1, 2, 3]^T$.

Voorbeeld 4. Om een parametervoorstelling te vinden van het vlak met vergelijking $2x - 4y + 3z = 12$ stellen we bijv. $y = \lambda$, $z = 2\mu$, dan is $x = 6 + 2\lambda - 3\mu$, zodat we hebben verkregen

$$\underline{x} = [6, 0, 0]^T + \lambda [2, 1, 0]^T + \mu [-3, 0, 2]^T .$$

Voorbeeld 5. Om een parametervoorstelling te vinden van de snijlijn van de vlakken $x + y = 1$, $2x - y + z = 3$ stellen we bijv. $x = \lambda$; dan is $y = 1 - \lambda$, $z = 4 - 3\lambda$, zodat we vinden:

$$\underline{x} = [0, 1, 4]^T + \lambda [1, -1, -3]^T .$$

Voorbeeld 6. Het vlak door de punten

$$\underline{a} = [a_1, a_2, a_3]^T , \quad \underline{b} = [b_1, b_2, b_3]^T \quad \text{en} \quad \underline{c} = [c_1, c_2, c_3]^T$$

die verondersteld worden niet op een rechte te liggen, heeft als parametervoorstelling

$$\underline{x} = \underline{c} + \lambda(\underline{b} - \underline{a}) + \mu(\underline{c} - \underline{a}) .$$

7.4.3. Meetkundige betekenis van het inwendig product

Is $\underline{a} = [a_1, a_2, a_3]^T$, $\underline{b} = [b_1, b_2, b_3]^T$ dan definiëerden we (zie 7.2.7)

$$(\underline{a}, \underline{b}) = a_1 b_1 + a_2 b_2 + a_3 b_3 .$$

Volgens de stelling van Pythagoras is $(\underline{a}, \underline{a})$ dus gelijk aan het kwadraat van de lengte van het verbindingslijnstuk van $\underline{0}$ en \underline{a} . We noemen dit de lengte van \underline{a} , notatie $|\underline{a}|$, $|\underline{a}| = \sqrt{(\underline{a}, \underline{a})}$. Dan is (wederom op grond van de stelling van Pythagoras)

$$|\underline{a} - \underline{b}| = \sqrt{(\underline{a} - \underline{b}, \underline{a} - \underline{b})} = [(a_1 - b_1)^2 + (a_2 - b_2)^2 + (a_3 - b_3)^2]^{\frac{1}{2}}$$

de afstand van \underline{a} tot \underline{b} .

Zij φ de hoek tussen de rechte door $\underline{0}$ en \underline{a} en de rechte door $\underline{0}$ en \underline{b} dan is volgens de cosinusregel:

$$|\underline{a}|^2 + |\underline{b}|^2 - 2|\underline{a}||\underline{b}|\cos \varphi = |\underline{a} - \underline{b}|^2 .$$

Wegens

$$|\underline{a} - \underline{b}|^2 = (\underline{a} - \underline{b}, \underline{a} - \underline{b}) = (\underline{a}, \underline{a}) - 2(\underline{a}, \underline{b}) + (\underline{b}, \underline{b})$$

volgt dan:

$$(\underline{a}, \underline{b}) = |\underline{a}||\underline{b}|\cos \varphi .$$

Is i.h.b. $\cos \varphi = 0$ dus $\underline{a} \perp \underline{b}$, dan is ook $(\underline{a}, \underline{b}) = 0$. Is $|\underline{a}| \neq 0 \neq |\underline{b}|$ dan volgt omgekeerd uit $(\underline{a}, \underline{b}) = 0$ dat $\underline{a} \perp \underline{b}$. We zullen van nu af $(\underline{a}, \underline{b}) = 0$ vaak aangeven als $\underline{a} \perp \underline{b}$ (ook al is een van beide misschien $\underline{0}$). We illustreren het gebruik van het inwendig product aan een aantal voorbeelden.

Voorbeeld 1. De rechten $\underline{x} = [1, 1, 6]^T + \lambda[1, -2, -1]^T$ en $\underline{y} = [3, -1, 2]^T + \mu[-1, 2, -5]^T$ zijn loodrecht omdat het inwendig produkt van hun richtingsvectoren nul is.

Voorbeeld 2. De vergelijking van een vlak $\perp \underline{v}$ is: $(\underline{x}, \underline{v}) = C$. Zij gegeven een vlak met vergelijking $ax + by + cz = d$, dan is de rechte door $\underline{0}$ en $[a, b, c]^T$ loodrecht op dit vlak.

Voorbeeld 3. De vergelijking van de bol met middelpunt \underline{a} en straal r is:
 $(\underline{x} - \underline{a}, \underline{x} - \underline{a}) = r^2$. Het poolvlak van \underline{p} t.o.v. deze bol heeft als vergelijking: $(\underline{x} - \underline{a}, \underline{p} - \underline{a}) = r^2$. Dit vlak staat dus loodrecht op $\underline{p} - \underline{a}$. Als \underline{p} zelf op de bol ligt (dus $(\underline{p} - \underline{a}, \underline{p} - \underline{a}) = r^2$), dan is het poolvlak raakvlak aan de bol in \underline{p} .

Voorbeeld 4. Voor de hoek φ , die $\underline{a} = [2, 2, 1]^T$ maakt met de y -as, geldt
 $\cos \varphi = 2/3$.

Voorbeeld 5. Bepaal de afstand van $\underline{a} = [-1, 2, 3]^T$ tot de rechte

$$\ell : \underline{x} = [-1, 14, -3] + \lambda[-1, 2, -2]$$

Het vlak door \underline{a} loodrecht op ℓ heeft als vergelijking:

$$-x + 2y - 2z = -1$$

De lijn ℓ snijdt dit vlak in $\underline{b} = [3, 6, 5]^T$.

De afstand van \underline{a} tot \underline{b} is de gevraagde afstand; deze is $(4^2 + 4^2 + 2^2)^{\frac{1}{2}} = 6$.

Voorbeeld 6. Om de afstand te bepalen van de rechten

$$\ell : \underline{x} = [0, 1, 0]^T + \lambda[2, 1, 1]^T$$

en

$$m : \underline{x} = [0, -1, 3]^T + \mu[1, 2, 1]^T$$

nemen we een willekeurig punt op $\ell : [2\lambda, \lambda + 1, \lambda]^T$, en een willekeurig punt op $m : [\mu, 2\mu - 1, \mu + 3]^T$. De voorwaarde dat de verbindingsvector van beide punten met $[2, 1, 1]^T$ en met $[1, 2, 1]^T$ inwendig product 0 heeft, leidt tot twee vergelijkingen in λ en μ :

$$\begin{cases} 2(2\lambda - \mu) + (\lambda - 2\mu + 2) + (\lambda - \mu - 3) = 0, \\ (2\lambda - \mu) + 2(\lambda - 2\mu + 2) + (\lambda - \mu - 3) = 0. \end{cases}$$

De oplossing $\lambda = 1, \mu = 1$ geeft de punten op ℓ en m waartussen de afstand van de rechten wordt aangenomen. Deze punten zijn $[2, 2, 1]^T$ op ℓ en $[1, 1, 4]^T$ op m ; de afstand is $\sqrt{11}$.

Voorbeeld 7. Alle vectoren $\underline{x} = [x, y, z]^T \neq \underline{0}$ die gelijke hoeken maken met $\underline{a} = [2, 2, 1]^T$ en $\underline{b} = [4, 0, 3]^T$ voldoen aan:

$$\frac{(\underline{a}, \underline{x})}{|\underline{a}| |\underline{x}|} = \frac{(\underline{b}, \underline{x})}{|\underline{b}| |\underline{x}|} .$$

Uitwerking levert $x - 5y + 2z = 0$, d.i. de vergelijking van het bissectricevlak in \underline{a} en \underline{b} .

Voorbeeld 8. Als van een viervlak twee paren overstaande ribben elkaar loodrecht kruisen, is ook het derde paar onderling loodrecht.

Bewijs. Laat de hoekpunten van het viervlak zijn $\underline{0}$, \underline{a} , \underline{b} en \underline{c} , en laat gegeven zijn dat ribbe $(\underline{0}, \underline{a}) \perp$ ribbe $(\underline{b}, \underline{c})$ en ribbe $(\underline{0}, \underline{b}) \perp$ ribbe $(\underline{a}, \underline{c})$. Met gebruik van het inwendig product kan men deze gegevens uitdrukken als

$(\underline{a}, \underline{b} - \underline{c}) = 0$, $(\underline{b}, \underline{a} - \underline{c}) = 0$, waaruit direct volgt $(\underline{c}, \underline{a} - \underline{b}) = 0$ ofwel ribbe $(\underline{0}, \underline{c}) \perp$ ribbe $(\underline{a}, \underline{b})$. □

7.4.4. Het uitwendig product

Anders dan het behandelde in 7.4.2 en 7.4.3 zal het nu in te voeren uitwendig product niet van \mathbb{R}^3 tot \mathbb{R}^n gegeneraliseerd worden. Het speelt in de lineaire algebra verder geen rol; we introduceren het hier slechts omdat het in de analytische meetkunde soms handig is.

Zij $\underline{a} = [a_1, a_2, a_3]^T$ en $\underline{b} = [b_1, b_2, b_3]^T$, dan definiëren we het uitwendig product (of vectorproduct) $\underline{a} \times \underline{b}$ door:

$$\underline{a} \times \underline{b} := [a_2 b_3 - a_3 b_2, a_3 b_1 - a_1 b_3, a_1 b_2 - a_2 b_1]^T .$$

Men kan eenvoudig inzien dat

$$\underline{a} \times \underline{b} \perp \underline{a} , \quad \underline{a} \times \underline{b} \perp \underline{b} .$$

Voorbeeld. De vergelijking van het vlak door $\underline{0}$, \underline{a} en \underline{b} is: $(\underline{x}, \underline{a} \times \underline{b}) = 0$.

7.4.5. Meetkundige terminologie in \mathbb{R}^n ($n > 3$)

Ook voor \mathbb{R}^n met $n > 3$ voeren we een meetkundige terminologie in. Neem bijv. $n = 4$. Het punt $\underline{0} = [0,0,0,0]^T$ heet de oorsprong; $\{[a,0,0,0]^T \mid a \in \mathbb{R}\}$ heet de x -as, enz. De verzameling $\{\underline{x} \in \mathbb{R}^4 \mid \underline{x} = \underline{a} + \lambda \underline{v}, \lambda \in \mathbb{R}\}$ heet een rechte met \underline{a} als steunvector en \underline{v} als richtingsvector ($\underline{v} \neq \underline{0}$).

Evenzo zeggen we als $(\underline{x}, \underline{y}) = 0$: $\underline{x} \perp \underline{y}$, \underline{x} loodrecht op \underline{y} .

We noemen $\sqrt{(\underline{x}-\underline{y}, \underline{x}-\underline{y})} =: |\underline{x}-\underline{y}|$ de afstand van \underline{x} en \underline{y} ; de afstand van \underline{x} tot $\underline{0}$, $|\underline{x}| = \sqrt{(\underline{x}, \underline{x})}$, heet ook de lengte van \underline{x} .

Het is duidelijk dat het afstandsbe­grip in \mathbb{R}^n voldoet aan:

$$\begin{aligned} |\underline{x} - \underline{x}| &= |\underline{0}| = 0 ; \\ |\underline{x} - \underline{y}| &> 0 \text{ als } \underline{x} \neq \underline{y} ; \\ |\underline{x} - \underline{y}| &= |\underline{y} - \underline{x}| . \end{aligned}$$

Een andere eigenschap:

$$|\underline{x} - \underline{y}| \leq |\underline{x} - \underline{z}| + |\underline{z} - \underline{y}| ,$$

de driehoeksongelijkheid voor afstanden in \mathbb{R}^n , volgt uit de ongelijkheid van Cauchy-Schwarz:

Stelling (ongelijkheid van Cauchy-Schwarz). Voor alle $\underline{x}, \underline{y} \in \mathbb{R}^n$ geldt:

$$|(\underline{x}, \underline{y})| \leq |\underline{x}| |\underline{y}| .$$

Bewijs. We moeten aantonen dat $(\underline{x}, \underline{y})^2 \leq (\underline{x}, \underline{x})(\underline{y}, \underline{y})$.

Als $\underline{x} = \underline{0}$ is de bewering triviaal. Als $\underline{x} \neq \underline{0}$ is voor alle $\lambda \in \mathbb{R}$ de kwadratische functie

$$(\lambda \underline{x} + \underline{y}, \lambda \underline{x} + \underline{y}) = \lambda^2 (\underline{x}, \underline{x}) + 2\lambda (\underline{x}, \underline{y}) + (\underline{y}, \underline{y}) \geq 0 .$$

Voor de discriminant geldt derhalve $4(\underline{x}, \underline{y})^2 - 4(\underline{x}, \underline{x})(\underline{y}, \underline{y}) \leq 0$. □

Opmerkingen. 1) Als $\underline{x} \neq \underline{0}$ dan volgt uit $|(\underline{x}, \underline{y})| = |\underline{x}| |\underline{y}|$ dat er een $\lambda \in \mathbb{R}$ bestaat waarvoor $\lambda \underline{x} + \underline{y} = \underline{0}$.

2) De ongelijkheid van Cauchy-Schwarz maakt het mogelijk om in \mathbb{R}^n de hoek φ tussen \underline{x} en \underline{y} ($\underline{x} \neq \underline{0} \neq \underline{y}$) te definiëren door

$$\cos \varphi := \frac{(\underline{x}, \underline{y})}{|\underline{x}| |\underline{y}|} .$$

De driehoeksongelijkheid is nu als volgt af te leiden:

$$\begin{aligned} |\underline{x} - \underline{y}|^2 &= |(\underline{x} - \underline{z}) + (\underline{z} - \underline{y})|^2 = |\underline{x} - \underline{z}|^2 + 2(\underline{x} - \underline{z}, \underline{z} - \underline{y}) + |\underline{z} - \underline{y}|^2 \leq \\ &\leq |\underline{x} - \underline{z}|^2 + 2|\underline{x} - \underline{z}| |\underline{z} - \underline{y}| + |\underline{z} - \underline{y}|^2 = (|\underline{x} - \underline{z}| + |\underline{z} - \underline{y}|)^2 . \end{aligned}$$

7.5. Vectorruimten

7.5.1. Inleiding

We beginnen nu met wat men de abstract-meetekundige behandeling van de lineaire algebra zou kunnen noemen. In 7.5.2 definiëren we een vectorruimte als een verzameling, waarvan het er niet toe doet wat de elementen voor objecten zijn, maar waarin een structuur aanwezig is die erg lijkt op die van \mathbb{R}^n . Onze meetkundige intuïtie kan dan een leidraad zijn bij de bestudering van dergelijke abstracte vectorruimten. In dit college, dat toch een eerste kennismaking is, zullen we met het bestuderen van de abstracte situatie niet erg ver gaan; de meerderheid van de beschouwingen zal zich in \mathbb{R}^n afspelen.

Na de definitie en eenvoudigste eigenschappen van vectorruimten in 7.5.2, is de rest van deze paragraaf gewijd aan voorbeelden. Deze worden gegeven in twee series: eerst vindt men in 7.5.3 een aantal voorbeelden die geen andere functie hebben dan de wijdheid van het begrip vectorruimte te illustreren; daarna komen in 7.5.4, 7.5.5 en 7.5.6 de voor dit hoofdstuk essentiële voorbeelden van vectorruimten. Zoals te verwachten hebben deze te maken met stelsels lineaire vergelijkingen.

7.5.2. Definitie van vectorruimten en deelruimten

Definitie. Een reële vectorruimte (of kortweg vectorruimte of lineaire ruimte) is een verzameling, L , van elementen, genaamd vectoren (genoteerd $\underline{a}, \underline{b}, \dots$) waarin twee bewerkingen gedefinieerd zijn, namelijk een optelling: bij elk tweetal vectoren \underline{a} en \underline{b} is er één vector die we $\underline{a} + \underline{b}$ noemen, en een scalaire vermenigvuldiging: bij elk reëel getal r en elke vector \underline{a} is er één vector die we $r\underline{a}$ noemen, en wel zo dat deze optelling en scalaire vermenigvuldiging voldoen aan de onderstaande regels 1 t/m 8; hierin zijn $\underline{a}, \underline{b}, \underline{c} \in L$, $r, s \in \mathbb{R}$.

- 1) $(\underline{a} + \underline{b}) + \underline{c} = \underline{a} + (\underline{b} + \underline{c})$ (associatieve eigenschap);
- 2) $\underline{a} + \underline{b} = \underline{b} + \underline{a}$ (commutatieve eigenschap);
- 3) er is precies één vector, $\underline{0}$ (genaamd nulvector), zodat $\underline{a} + \underline{0} = \underline{a}$ voor elke $\underline{a} \in L$;
- 4) bij elke $\underline{a} \in L$ bestaat er precies één vector, genoteerd $\underline{-a}$, waarvoor $\underline{a} + (\underline{-a}) = \underline{0}$;
- 5) $1\underline{a} = \underline{a}$;

*het ja, betrekking hier;
Elke groep*

- 6) $r(\underline{sa}) = (rs)\underline{a}$;
 - 7) $(r + s)\underline{a} = r\underline{a} + s\underline{a}$;
 - 8) $r(\underline{a} + \underline{b}) = r\underline{a} + r\underline{b}$.
- (distributieve eigenschappen).

Opmerkingen. 1) Men ziet onmiddellijk dat \mathbb{R}^n met de optelling en scalaire vermenigvuldiging uit 7.2.7 een vectorruimte is. Voor het inzicht zijn \mathbb{R}^2 en \mathbb{R}^3 zeer belangrijke voorbeelden.

2) Men zou ook complexe vectorruimten kunnen invoeren door overal in de definitie reëel door complex en \mathbb{R} door \mathbb{C} te vervangen. We doen dit niet.

Eigenschappen. 1) $\forall \underline{a} \in L \quad \forall \underline{b} \in L \quad \exists! \underline{x} \in L \quad [\underline{a} + \underline{x} = \underline{b}]$; in woorden: in L heeft de vergelijking $\underline{a} + \underline{x} = \underline{b}$ voor elke \underline{a} en \underline{b} een eënduidige oplossing.

- 2) $\forall \underline{a} \in L \quad [0\underline{a} = \underline{0}]$
- 3) $\forall r \in \mathbb{R} \quad [r\underline{0} = \underline{0}]$
- 4) $\forall \underline{a} \in L \quad [-\underline{a} = (-1)\underline{a}]$.

$$D(0,1) \subset C(0,1)$$

Definitie. Een deelverzameling V van een vectorruimte L heet een (lineaire) deelruimte indien V met de aanwezige optelling en scalaire vermenigvuldiging zelf een vectorruimte is.

Voorbeelden van deelruimten in \mathbb{R}^2 zijn: $\{\underline{0}\}$, \mathbb{R}^2 zelf, rechten door $\underline{0}$. In \mathbb{R}^3 hebben we als deelruimten: $\{\underline{0}\}$, rechten door $\underline{0}$, vlakken door $\underline{0}$, \mathbb{R}^3 zelf.

Eigenschap. 5) Zij L een vectorruimte, $V \subset L$, $V \neq \emptyset$ dan is V een deelruimte dan en slechts dan als voor alle $r \in \mathbb{R}$ en alle $\underline{x}, \underline{y} \in V$ geldt:

$$\underline{x} + \underline{y} \in V, \quad r\underline{x} \in V.$$

Definitie. Is V een deelruimte van een vectorruimte L dan heet een verzameling van de vorm

$$\underline{a} + V := \{ \underline{a} + \underline{x} \mid \underline{x} \in V \}$$

een nevenruimte (van V).

Rechten in \mathbb{R}^2 niet door $\underline{0}$, vlakken en rechten niet door $\underline{0}$ in \mathbb{R}^3 zijn voorbeelden van nevenruimten die niet zelf deelruimten zijn. Uiteraard is elke deelruimte van een vectorruimte ook een nevenruimte want $\underline{0} + V = V$.

Eigenschap. 6) Als V en W deelruimten zijn van een vectorruimte L en $\underline{a}, \underline{b} \in L$, dan is $\underline{a} + V = \underline{b} + W$ dan en slechts dan als $V = W$ en $\underline{a} - \underline{b} \in V$.

n x 1 matrix

7.5.3. Voorbeelden van vectorruimten en deelruimten

- 1) De verzameling van de $(m \times n)$ -matrices vormt met de bewerkingen uit 7.2.2 een vectorruimte. Het nulelement is $\mathcal{O}^{(m \times n)}$.
- 2) De verzameling van alle rijen $\underline{a} := (a_1, a_2, a_3, \dots)$ vormt een vectorruimte als $\underline{a} + \underline{b} := (a_1 + b_1, a_2 + b_2, \dots)$ en $r\underline{a} := (ra_1, ra_2, \dots)$ ($r \in \mathbb{R}$). Dit voorbeeld is een voor de hand liggende uitbreiding van \mathbb{R}^n .
- 3) De verzameling, H , van alle rijen $\underline{a} = (a_1, a_2, \dots)$ waarvoor $\sum_{n=1}^{\infty} a_n^2$ convergeert vormt een deelruimte van de vectorruimte uit voorbeeld 2. Het is nu niet meteen duidelijk dat uit $\underline{a} \in H$, $\underline{b} \in H$ volgt $\underline{a} + \underline{b} \in H$. We zullen echter aantonen dat uit de convergentie van $\sum_{n=1}^{\infty} a_n^2$ en $\sum_{n=1}^{\infty} b_n^2$ volgt dat $\sum_{n=1}^{\infty} (a_n + b_n)^2$ convergeert.

De reeks $\sum_{n=1}^{\infty} (a_n + b_n)^2$ heeft niet negatieve termen. Uit $(a_n + b_n)^2 \leq 2a_n^2 + 2b_n^2$ (ga na!) volgt voor de partiële sommen:

$$\sum_{n=1}^N (a_n + b_n)^2 \leq 2 \sum_{n=1}^N a_n^2 + 2 \sum_{n=1}^N b_n^2 \leq 2 \left(\sum_{n=1}^{\infty} a_n^2 + \sum_{n=1}^{\infty} b_n^2 \right).$$

Een reeks met niet negatieve termen waarvan de partiële sommen begrensd zijn is convergent (3.2.1).

De vectorruimte H heet een Hilbertruimte.

- 4) Zij F de verzameling van alle reële functies op \mathbb{R} . Voor $f \in F$, $g \in F$ en $r \in \mathbb{R}$ definiëren we de functies $f + g$ en rf door:

$$\begin{cases} (f + g)(t) := f(t) + g(t) \\ (rf)(t) := rf(t) \end{cases} \quad (t \in \mathbb{R}).$$

Met deze definities is F een vectorruimte.

- 5) De verzameling C van alle continue functies op \mathbb{R} vormt een deelruimte van F uit voorbeeld 4. De verzameling $\text{Pol } n$ bestaande uit het nulpolynoom en alle polynomen met graad $\leq n$ ($n \in \mathbb{N}$) vormt een deelruimte van C en dus ook van F .

7.5.4. Voortgebrachte lineaire deelruimten

Zij $\underline{v}_1, \underline{v}_2, \dots, \underline{v}_k$ een stelsel vectoren uit een vectorruimte L (bijv. uit \mathbb{R}^n). Een vector

$$r_1 \underline{v}_1 + r_2 \underline{v}_2 + \dots + r_k \underline{v}_k$$

met $r_1, r_2, \dots, r_k \in \mathbb{R}$ heet een lineaire combinatie van $\underline{v}_1, \underline{v}_2, \dots, \underline{v}_k$.

De verzameling van alle lineaire combinaties van $\underline{v}_1, \dots, \underline{v}_k$ dus $\{r_1 \underline{v}_1 + \dots + r_k \underline{v}_k \mid r_1, \dots, r_k \in \mathbb{R}\}$ vormt een deelruimte van L ; notatie: $\langle \underline{v}_1, \dots, \underline{v}_k \rangle$.

De deelruimte $\langle \underline{v}_1, \dots, \underline{v}_k \rangle \subset L$ heet de door $\underline{v}_1, \dots, \underline{v}_k$ voortgebrachte of opgespannen lineaire ruimte. We breiden het begrip iets uit: is M een willekeurige deelverzameling van L , dan is ook de verzameling van alle lineaire combinaties van eindig veel elementen uit M een deelruimte; notatie: $\langle M \rangle$.

Voorbeeld 1. Beschouw in \mathbb{R}^4 : $\underline{a} = [1, 1, 2, 0]^T$, $\underline{b} = [0, 2, 0, 1]^T$ en $\underline{c} = [0, 0, 1, 3]^T$. De vraag of $\underline{d} = [d_1, d_2, d_3, d_4]^T$ element van $\langle \underline{a}, \underline{b}, \underline{c} \rangle$ is, is in feite de vraag naar het bestaan van reële getallen x_1, x_2, x_3 zó dat $\underline{d} = x_1 \underline{a} + x_2 \underline{b} + x_3 \underline{c}$. Dit voert tot een stelsel lineaire vergelijkingen in x_1, x_2, x_3 , nl.

$$\begin{cases} d_1 = x_1 \\ d_2 = x_1 + 2x_2 \\ d_3 = 2x_1 + x_3 \\ d_4 = x_2 + 3x_3 \end{cases}$$

Zo is

$$[1, 3, 4, 7]^T \in \langle \underline{a}, \underline{b}, \underline{c} \rangle$$

en

$$[1, 3, 4, 6]^T \notin \langle \underline{a}, \underline{b}, \underline{c} \rangle .$$

Voorbeeld 2. Zij $\underline{e}_1 = [1, 0, \dots, 0]^T$, $\underline{e}_2 = [0, 1, \dots, 0]^T$, ..., $\underline{e}_n = [0, \dots, 0, 1]^T$ in \mathbb{R}^n dan is $\mathbb{R}^n = \langle \underline{e}_1, \underline{e}_2, \dots, \underline{e}_n \rangle$.

Opmerkingen. 1) Voor iedere vectorruimte L geldt $\langle L \rangle = L$.

2) In het bijzonder is $\langle \langle \underline{v}_1, \dots, \underline{v}_k \rangle \rangle = \langle \underline{v}_1, \dots, \underline{v}_k \rangle$; immers een lineaire combinatie van lineaire combinaties van $\underline{v}_1, \dots, \underline{v}_k$ is zelf een lineaire combinatie van $\underline{v}_1, \dots, \underline{v}_k$.

7.5.5. Rijenruimte, kolommenruimte, oplossingsruimte

Zij $A = [a_{ij}]$ een $(m \times n)$ -matrix; laat $\underline{a}_1, \underline{a}_2, \dots, \underline{a}_m$, ($\underline{a}_i \in \mathbb{R}^n$, $i = 1, \dots, m$) de rijen van A zijn en laat $\underline{b}_1, \underline{b}_2, \dots, \underline{b}_n \in \mathbb{R}^m$ de kolommen van A zijn. Met de bloknootatie uit 7.2.5 betekent dit:

$$A = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{bmatrix} = \begin{bmatrix} \underline{a}_1^T \\ \vdots \\ \underline{a}_m^T \end{bmatrix} = [\underline{b}_1, \dots, \underline{b}_n] .$$

De deelruimte $\langle \underline{a}_1, \dots, \underline{a}_m \rangle \subset \mathbb{R}^n$ heet de rijenruimte van de matrix. De deelruimte $\langle \underline{b}_1, \dots, \underline{b}_n \rangle \subset \mathbb{R}^m$ heet de kolommenruimte van A .

Merk op dat $\langle \underline{b}_1, \dots, \underline{b}_n \rangle = A\mathbb{R}^n := \{ \underline{Ax} \mid \underline{x} \in \mathbb{R}^n \}$; omdat $A\mathbb{R}^n$ het beeld is van \mathbb{R}^n onder de afbeelding $\underline{x} \mapsto \underline{Ax}$, heet de kolommenruimte van A ook wel de beeldruimte van deze afbeelding.

Uit 7.2.9 volgt dat $N(A)$, de oplossingsruimte van het stelsel $\underline{Ax} = \underline{0}$, een deelruimte van \mathbb{R}^n is.

Ook uit 7.2.9 volgt dat als het inhomogene stelsel $\underline{Ax} = \underline{b}$ een oplossing heeft (dat is als $\underline{b} \in A\mathbb{R}^n$), de verzameling van alle oplossingen een nevenruimte van $N(A)$ is.

7.5.6. Orthogonaal complement

Is M een deelverzameling van \mathbb{R}^n dan definiëren we

$$M^\perp := \{ \underline{x} \in \mathbb{R}^n \mid (\underline{x}, \underline{y}) = 0 \text{ voor alle } \underline{y} \in M \} .$$

M^\perp is de verzameling van alle vectoren \underline{x} die loodrecht staan op alle elementen van M . M^\perp is een deelruimte van \mathbb{R}^n (ga na) en heet het orthogonaal complement of orthoplement van M .

Voorbeeld. $N(A)$ is het orthoplement van de rijenruimte van A .

Eigenschappen. 1) Is $N \subset M$ dan is $N^\perp \supset M^\perp$.

2) Als $(\underline{x}, \underline{y}) = 0$ voor alle $\underline{y} \in \mathbb{R}^n$, dan is $\underline{x} = \underline{0}$, m.a.w. $(\mathbb{R}^n)^\perp = \{\underline{0}\}$.

Stelling. Is $\underline{v}_1, \dots, \underline{v}_k$ een stelsel vectoren uit \mathbb{R}^n , dan is

$$\{\underline{v}_1, \dots, \underline{v}_k\}^\perp = \langle \underline{v}_1, \dots, \underline{v}_k \rangle^\perp.$$

In woorden: het orthoplement van een eindige verzameling vectoren is gelijk aan het orthoplement van de door die vectoren voortgebrachte lineaire deelruimte.

Bewijs. Voer in de afkortingen: $M := \{\underline{v}_1, \dots, \underline{v}_k\}$, $L := \langle M \rangle$.

Als $\underline{x} \in M^\perp$, dan is $(\underline{x}, \underline{v}_i) = 0$ ($i = 1, \dots, k$). Voor elke vector $r_1 \underline{v}_1 + \dots + r_k \underline{v}_k \in L$ geldt dan:

$$(\underline{x}, r_1 \underline{v}_1 + \dots + r_k \underline{v}_k) = r_1 (\underline{x}, \underline{v}_1) + \dots + r_k (\underline{x}, \underline{v}_k) = 0 ;$$

dus $\underline{x} \in L^\perp$. Conclusie $M^\perp \subset L^\perp$.

Omdat $M \subset L$, geldt ook $M^\perp \supset L^\perp$, zodat $M^\perp = L^\perp$. □

7.6. Afhankelijkheid en onafhankelijkheid, bases

7.6.1. Inleiding

Deze paragraaf vormt met de volgende (7.7) het centrale deel van hoofdstuk 7. Grofweg kunnen we het probleem dat in 7.6 aan de orde komt zo omschrijven: gegeven een vectorruimte L (bijv. \mathbb{R}^n) en een deelruimte V , kunnen we nu V op zodanige wijze voorstellen door $V = \langle \underline{v}_1, \dots, \underline{v}_k \rangle$ dat voor willekeurige $\underline{a} \in L$ in een oogopslag is vast te stellen of $\underline{a} \in V$ dan wel $\underline{a} \notin V$. De paragraaf valt uiteen in vier grote delen.

7.6.2. Afhankelijkheid en onafhankelijkheid

Vb (1, 2, 3, 4)

Definitie. Zij L een vectorruimte. Het stelsel vectoren v_1, v_2, \dots, v_k uit L heet afhankelijk (of lineair afhankelijk) als er reële getallen r_1, r_2, \dots, r_k bestaan, die niet alle nul zijn, zodanig dat

$$r_1 v_1 + r_2 v_2 + \dots + r_k v_k = \underline{0}.$$

Anders gezegd: de vectoren v_1, \dots, v_k zijn afhankelijk als $\underline{0}$ op meer dan één manier als lineaire combinatie van v_1, \dots, v_k te schrijven is; (steeds is: $\underline{0} = 0v_1 + \dots + 0v_k$).

Voorbeeld 1. De vectoren $\underline{a} = [1, 1, 2, 3]^T$, $\underline{b} = [0, 1, 0, 1]^T$ en $\underline{c} = [1, 3, 2, 5]^T$ uit \mathbb{R}^4 zijn afhankelijk want

$$\underline{0} = \underline{a} + 2\underline{b} - \underline{c}.$$

Het homogene stelsel vergelijkingen

$$\underline{0} = x\underline{a} + y\underline{b} + z\underline{c}$$

of

$$\begin{cases} 0 = x & + z \\ 0 = x + y + 3z \\ 0 = 2x & + 2z \\ 0 = 3x + y + 5z \end{cases}$$

heeft een van $x = 0, y = 0, z = 0$ (de nuloplossing) verschillende oplossing, bijv.: $x = 1, y = 2, z = -1$.

Definitie. Het stelsel vectoren v_1, v_2, \dots, v_k uit L heet onafhankelijk (of lineair onafhankelijk) als het stelsel niet afhankelijk is.

Anders gezegd: v_1, \dots, v_k zijn onafhankelijk als uit $r_1 v_1 + \dots + r_k v_k = \underline{0}$ volgt $r_1 = \dots = r_k = 0$.

Voorbeeld 2. De vectoren $\underline{a} = [1, 1, 2, 3]^T$, $\underline{b} = [0, 1, 0, 1]^T$ en $\underline{c} = [1, 3, 2, 6]^T$ uit \mathbb{R}^4 zijn onafhankelijk, want uit $r_1 \underline{a} + r_2 \underline{b} + r_3 \underline{c} = \underline{0}$ volgt:

$$\begin{cases} r_1 + r_3 = 0 \\ r_1 + r_2 + 3r_3 = 0 \\ 2r_1 + 2r_3 = 0 \\ 3r_1 + r_2 + 6r_3 = 0 . \end{cases}$$

Met $\underline{a}_1 = [1,0,1]^T$, $\underline{a}_2 = [1,1,3]^T$, $\underline{a}_3 = [2,0,2]^T$, $\underline{a}_4 = [3,1,6]^T$,
 $\underline{r} = [r_1, r_2, r_3]^T$ kunnen we dit stelsel schrijven als:

$$\begin{cases} (\underline{a}_1, \underline{r}) = 0 \\ (\underline{a}_2, \underline{r}) = 0 \\ (\underline{a}_3, \underline{r}) = 0 \\ (\underline{a}_4, \underline{r}) = 0 . \end{cases}$$

Hieruit volgt: $(\underline{a}_4 - \underline{a}_3 - \underline{a}_2, \underline{r}) = 0$, dus $r_3 = 0$. Daarna volgt uit $(\underline{a}_1, \underline{r}) = 0$ dat ook $r_1 = 0$ en tenslotte uit $(\underline{a}_2, \underline{r}) = 0$ dat $r_2 = 0$.

Opmerking. De meetkundige betekenis van afhankelijkheid.

In \mathbb{R}^2 en \mathbb{R}^3 geldt:

- 1) Twee rechten zijn evenwijdig (of samenvallend) als hun richtingsvectoren afhankelijk zijn.
- 2) Een vlak door 0 bestaat uit alle vectoren die een lineaire combinatie zijn van twee onafhankelijke vectoren.
- 3) De rechte $\underline{x} = \underline{a} + \lambda \underline{u}$ is evenwijdig met (of ligt in) het vlak $\underline{x} = \underline{b} + \mu \underline{v} + \nu \underline{w}$ als \underline{u} , \underline{v} en \underline{w} afhankelijk zijn. Omdat $\underline{u} \neq 0$ (anders is $\underline{x} = \underline{a} + \lambda \underline{u}$ geen rechte) en omdat \underline{v} en \underline{w} onafhankelijk zijn (vlak!) volgt in dit geval uit de afhankelijkheid van \underline{u} , \underline{v} en \underline{w} , dat \underline{u} een lineaire combinatie is van \underline{v} en \underline{w} (ga na).
- 4) De punten \underline{a} , \underline{b} en \underline{c} liggen op een rechte als $\underline{b} - \underline{a}$ en $\underline{c} - \underline{a}$ afhankelijk zijn.
- 5) De punten \underline{a} , \underline{b} , \underline{c} en \underline{d} liggen in een vlak als $\underline{b} - \underline{a}$, $\underline{c} - \underline{a}$, $\underline{d} - \underline{a}$ afhankelijk zijn.

Eigenschappen. Door nauwgezette toepassing van de definities zijn de volgende eigenschappen te bewijzen:

- 1) $\underline{0}$ is een afhankelijk stelsel.
- 2) Als $\underline{v}_1, \dots, \underline{v}_k$ een afhankelijk stelsel is, dan is ook $\underline{v}_1, \dots, \underline{v}_k, \underline{w}_1, \dots, \underline{w}_\ell$ afhankelijk voor elk stelsel $\underline{w}_1, \dots, \underline{w}_\ell$.
- 3) Is een van de vectoren $\underline{v}_1, \dots, \underline{v}_k$ de nulvector, dan is het stelsel $\underline{v}_1, \dots, \underline{v}_k$ afhankelijk.
- 4) Het stelsel $\underline{v}_1, \underline{v}_2, \dots, \underline{v}_k$ is dan en slechts dan onafhankelijk als $\underline{v}_1 + \underline{v}_2, \underline{v}_2, \dots, \underline{v}_k$ onafhankelijk is.
- 5) Zij $\alpha \neq 0$; het stelsel $\underline{v}_1, \underline{v}_2, \dots, \underline{v}_k$ is dan en slechts dan onafhankelijk als $\alpha \underline{v}_1, \underline{v}_2, \dots, \underline{v}_k$ onafhankelijk is.
- 6) Als $\underline{v}_1, \dots, \underline{v}_k$ een afhankelijk stelsel is, dan is tenminste één der \underline{v}_i een lineaire combinatie van de overige vectoren uit het stelsel.
Is nl. $r_1 \underline{v}_1 + \dots + r_k \underline{v}_k = \underline{0}$ en $r_i \neq 0$, dan is

$$\underline{v}_i = -\frac{r_1}{r_i} \underline{v}_1 - \dots - \frac{r_{i-1}}{r_i} \underline{v}_{i-1} - \frac{r_{i+1}}{r_i} \underline{v}_{i+1} - \dots - \frac{r_k}{r_i} \underline{v}_k .$$

Opmerking. Het onderzoek naar de afhankelijkheid van een stelsel.

Door herhaalde toepassing van de eigenschappen 4 en 5 tracht men het onderzoek naar de afhankelijkheid van een stelsel te vervangen door het onderzoek van een ander, eenvoudiger stelsel. Het systematisch uitvoeren van dit procédé noemen we "vegen". We zullen het spoedig als een nauwkeurig omschreven rekenrecept (algoritme) beschrijven. Vooruitlopend op dit recept twee voorbeelden:

Voorbeeld 3. Is $\underline{a} = [-1, 1, 1]^T$, $\underline{b} = [1, 2, 3]^T$, $\underline{c} = [5, 1, 3]^T$ afhankelijk? Schrijf de vectoren onder elkaar en veeg de tweede kolom schoon:

$$\begin{array}{l} \underline{a}^T = [-1, 1, 1] \\ \underline{b}^T = [1, 2, 3] \\ \underline{c}^T = [5, 1, 3] \end{array} \qquad \begin{array}{l} \underline{a}^T = [-1, 1, 1] \\ \underline{b}^T - 2\underline{a}^T = [3, 0, 1] \\ \underline{c}^T - \underline{a}^T = [6, 0, 2] . \end{array}$$

Omdat blijkbaar $\underline{c} - \underline{a} = 2(\underline{b} - 2\underline{a})$, dus $3\underline{a} - 2\underline{b} + \underline{c} = \underline{0}$, is het stelsel afhankelijk.

Voorbeeld 4. Is $\underline{a} = [1, 2, 3]^T$, $\underline{b} = [2, 7, 0]^T$, $\underline{c} = [1, 2, -1]^T$ afhankelijk?

We gaan analoog te werk als in voorbeeld 3:

$$\begin{array}{lll} \underline{a}^T = [1, 2, 3] & \underline{a}^T + 3\underline{c}^T = [4, 8, 0] & \underline{a}^T + 3\underline{c}^T - 2\underline{b}^T = [0, -6, 0] \\ \underline{b}^T = [2, 7, 0] & \underline{b}^T = [2, 7, 0] & \underline{b}^T = [2, 7, 0] \\ \underline{c}^T = [1, 2, -1] & \underline{c}^T = [1, 2, -1] & \underline{c}^T = [1, 2, -1] \end{array}$$

Het laatste stelsel is onafhankelijk en \underline{a} , \underline{b} , \underline{c} dus ook.

Immers uit $r_1[0, -6, 0]^T + r_2[2, 7, 0]^T + r_3[1, 2, -1]^T = \underline{0}$ volgt: $2r_2 + r_3 = 0$, $-6r_1 + 7r_2 + 2r_3 = 0$, $-r_3 = 0$ en dus $r_1 = r_2 = r_3 = 0$.

7.6.3. Bases

$$\forall b \langle x^1, x^2, \dots, x^k \rangle$$

7.6.3.1. Stelling. Zij L een vectorruimte, $\underline{v}_1, \dots, \underline{v}_k \in L$. Indien $\langle \underline{v}_1, \dots, \underline{v}_k \rangle \neq \{0\}$ dan heeft $\underline{v}_1, \dots, \underline{v}_k$ een onafhankelijk deelstelsel dat ook $\langle \underline{v}_1, \dots, \underline{v}_k \rangle$ voortbrengt.

Bewijs. Als $\underline{v}_1, \dots, \underline{v}_k$ onafhankelijk is, dan is er niets meer te bewijzen. Als $\underline{v}_1, \dots, \underline{v}_k$ afhankelijk is, dan is tenminste één der vectoren, zeg \underline{v}_i , een lineaire combinatie van de overige. In dit laatste geval is:

$$L := \langle \underline{v}_1, \dots, \underline{v}_k \rangle = \langle \underline{v}_1, \dots, \underline{v}_{i-1}, \underline{v}_{i+1}, \dots, \underline{v}_k \rangle$$

Indien $\underline{v}_1, \dots, \underline{v}_{i-1}, \underline{v}_{i+1}, \dots, \underline{v}_k$ nog steeds een afhankelijk stelsel is kunnen we wederom een der vectoren weglaten. Dit herhalen we tot we een onafhankelijk stelsel voortbrengenden overhouden. Daar $L \neq \{0\}$ wordt dit inderdaad bereikt. □

7.6.3.2. Definitie. Zij L een vectorruimte en zij $\underline{w}_1, \dots, \underline{w}_k$ een onafhankelijk stelsel in L zodanig dat $\langle \underline{w}_1, \dots, \underline{w}_k \rangle = L$, dan heet $\underline{w}_1, \dots, \underline{w}_k$ een basis van L .

Voorbeeld 1. Het stelsel $\underline{e}_1, \underline{e}_2, \dots, \underline{e}_n$ (7.5.4, voorbeeld 2) is een basis van \mathbb{R}^n . Deze noemen we de standaardbasis.

Opmerkingen. 1) Het is mogelijk dat een vectorruimte geen basis (met eindig veel elementen) heeft; voorbeeld: 7.5.3, voorbeeld 2. We zullen ons in het vervolg beperken tot vectorruimten die wel een basis hebben.

- 2) In 7.6.3.1 is bewezen dat $\langle \underline{v}_1, \dots, \underline{v}_k \rangle$ mits $\neq \{0\}$, een basis heeft.
 3) Een vectorruimte kan verschillende bases hebben: is $\underline{a} = [1, 1, 0, 0]^T$,
 $\underline{b} = [1, 2, 0, 0]^T$ dan heeft $L := \langle \underline{a}, \underline{b} \rangle = \langle \underline{e}_1, \underline{e}_2 \rangle \in \mathbb{R}^4$ de bases $\underline{a}, \underline{b}$ en $\underline{e}_1, \underline{e}_2$.

7.6.3.3. Stelling. Is $\underline{w}_1, \dots, \underline{w}_m$ een basis van de vectorruimte L , dan is elk element van L op éénduidige wijze te schrijven als lineaire combinatie van $\underline{w}_1, \dots, \underline{w}_m$.

Bewijs. Omdat $L = \langle \underline{w}_1, \dots, \underline{w}_m \rangle$ is elke $\underline{x} \in L$ te schrijven als lineaire combinatie van $\underline{w}_1, \dots, \underline{w}_m$. Stel nu dat

$$\underline{x} = r_1 \underline{w}_1 + \dots + r_m \underline{w}_m$$

en

$$\underline{x} = s_1 \underline{w}_1 + \dots + s_m \underline{w}_m,$$

dan is

$$(r_1 - s_1) \underline{w}_1 + \dots + (r_m - s_m) \underline{w}_m = \underline{0}.$$

Omdat $\underline{w}_1, \dots, \underline{w}_m$ onafhankelijk zijn is nu

$$r_1 - s_1 = \dots = r_m - s_m = 0.$$

□

7.6.3.4. Stelling (uitwisselingsstelling). Is $\underline{a}_1, \dots, \underline{a}_k$ een basis van de vectorruimte L en is $\underline{b}_1, \dots, \underline{b}_m$ een onafhankelijk stelsel in L , dan is $m \leq k$.

Bewijs. Daar $\underline{a}_1, \dots, \underline{a}_k$ een basis van L vormt, is $\underline{b}_1 = r_1 \underline{a}_1 + \dots + r_k \underline{a}_k$. Omdat \underline{b}_1 voorkomt in een onafhankelijk stelsel, is $\underline{b}_1 \neq \underline{0}$, dus $[r_1, \dots, r_k] \neq [0, \dots, 0]$. Uitsluitend voor het gemak van de notatie nemen we aan dat $r_1 \neq 0$ is (dit kunnen we altijd bereiken door passende verandering van de volgorde van $\underline{a}_1, \dots, \underline{a}_k$). Dan is

$$\underline{a}_1 = \frac{1}{r_1} \underline{b}_1 - \frac{r_2}{r_1} \underline{a}_2 - \dots - \frac{r_k}{r_1} \underline{a}_k.$$

Als we in iedere lineaire combinatie van $\underline{a}_1, \dots, \underline{a}_k$, de vector \underline{a}_1 vervangen door de zojuist gevonden uitdrukking zien we dat

$$L = \langle \underline{a}_1, \dots, \underline{a}_k \rangle = \langle \underline{b}_1, \underline{a}_2, \dots, \underline{a}_k \rangle.$$

Hieruit volgt:

$$\underline{b}_2 = s_1 \underline{b}_1 + s_2 \underline{a}_2 + \dots + s_k \underline{a}_k,$$

waarin $[s_2, \dots, s_k] \neq [0, \dots, 0]$ (anders was $\underline{b}_1, \underline{b}_2$ afhankelijk en $\underline{b}_1, \dots, \underline{b}_m$ dus ook). Neem gemakshalve aan dat $s_2 \neq 0$ is. Dan volgt

$$\underline{a}_2 = \frac{1}{s_2} \underline{b}_2 - \frac{s_1}{s_2} \underline{b}_1 - \frac{s_3}{s_2} \underline{a}_3 - \dots - \frac{s_k}{s_2} \underline{a}_k,$$

en dus

$$L = \langle \underline{b}_1, \underline{b}_2, \underline{a}_3, \dots, \underline{a}_k \rangle.$$

We herhalen nu dit uitwisselingsproces. Als $m \geq k$ is, hebben we na k uitwisselingen verkregen $L = \langle \underline{b}_1, \dots, \underline{b}_k \rangle$. Als nu $m > k$ zou zijn dan was

$\underline{b}_m = t_1 \underline{b}_1 + \dots + t_k \underline{b}_k$ in tegenspraak met de onafhankelijkheid van $\underline{b}_1, \dots, \underline{b}_m$. Dus $m \leq k$. □

Gevolg. Elk $(n + 1)$ -tal vectoren in \mathbb{R}^n is afhankelijk.

Gevolg. Alle bases van een vectorruimte bestaan uit evenveel vectoren.

Bewijs. Als $\underline{a}_1, \dots, \underline{a}_k$ en $\underline{b}_1, \dots, \underline{b}_m$ bases zijn van L , dan is $m \leq k$ en $k \leq m$. □

7.6.3.5. Definitie. Het aantal vectoren in een basis van een vectorruimte L heet de dimensie van L , notatie: $\dim L$, $\dim\{0\} := 0$. De dimensie van een nevenruimte $\underline{a} + L$ wordt gedefinieerd door $\dim(\underline{a} + L) := \dim L$.

Voorbeeld. $\dim \mathbb{R}^n = n$.

Stelling. Is V een deelruimte van een n -dimensionale vectorruimte L , dan heeft V een basis en $0 \leq \dim V \leq n$.

Bewijs. Als $V = \{0\}$, dan is $\dim V = 0$; als $V \neq \{0\}$, kies $\underline{v}_1 \in V$ met $\underline{v}_1 \neq 0$. Als $V = \langle \underline{v}_1 \rangle$, dan is $\dim V = 1$; als $V \neq \langle \underline{v}_1 \rangle$, kies $\underline{v}_2 \in V$ zó dat $\underline{v}_1, \underline{v}_2$ een onafhankelijk stelsel is.

Als $V = \langle \underline{v}_1, \underline{v}_2 \rangle$, dan is $\dim V = 2$; als $V \neq \langle \underline{v}_1, \underline{v}_2 \rangle$, kies $\underline{v}_3 \in V$ zó dat $\underline{v}_1, \underline{v}_2, \underline{v}_3$ een onafhankelijk stelsel is, enz.

Omdat elk $(n+1)$ -tal vectoren in L afhankelijk is, kan men zo niet onbeperkt doorgaan. We vinden $V = \langle \underline{v}_1, \dots, \underline{v}_k \rangle$ met $\underline{v}_1, \dots, \underline{v}_k$ onafhankelijk en $0 \leq k \leq n$.

Dus $\underline{v}_1, \dots, \underline{v}_k$ is een basis van V en $0 \leq \dim V = k \leq n$. \square

7.6.3.6. Stelling (uitbreidingsstelling). Zij V een k -dimensionale deelruimte van de n -dimensionale vectorruimte L met $0 < k < n$. Zij $\underline{v}_1, \dots, \underline{v}_k$ een basis van V , dan heeft L een basis $\underline{v}_1, \dots, \underline{v}_k, \underline{w}_1, \dots, \underline{w}_{n-k}$. (Deze basis van L ontstaat dus door uitbreiding van de basis van V met $n-k$ vectoren.)

Bewijs. Neem een basis $\underline{u}_1, \underline{u}_2, \dots, \underline{u}_n$ van L , dan is $L = \langle \underline{u}_1, \underline{u}_2, \dots, \underline{u}_n \rangle$. Volgens het uitwisselingsproces (7.6.3.4) zijn k van de \underline{u} 's uit de basis van L uit te wisselen tegen $\underline{v}_1, \dots, \underline{v}_k$. \square

7.6.3.7. Opmerking. We onderzoeken het verband tussen een k -dimensionale vectorruimte L en \mathbb{R}^k . Laat $\underline{v}_1, \dots, \underline{v}_k$ een basis van L zijn; elke vector $\underline{x} \in L$ is dan eenduidig bepaald door een kolomvector $\underline{r} = [r_1, \dots, r_k]^T \in \mathbb{R}^k$, nl. door middel van: $\underline{x} = r_1 \underline{v}_1 + \dots + r_k \underline{v}_k$. De getallen r_1, \dots, r_k heten de coördinaten (ook wel componenten) van \underline{x} t.o.v. de basis $\underline{v}_1, \dots, \underline{v}_k$. De vector $\underline{r} \in \mathbb{R}^k$ heet de kolom van \underline{x} t.o.v. deze basis. Let erop dat de kolom van \underline{x} afhangt van de gekozen basis. Het één-éénduidige verband tussen L en \mathbb{R}^k draagt ook de optelling en scalaire vermenigvuldiging over:

Als nl. $\underline{a} = a_1 \underline{v}_1 + \dots + a_k \underline{v}_k$, $\underline{b} = b_1 \underline{v}_1 + \dots + b_k \underline{v}_k$, met andere woorden als $\underline{a} \leftrightarrow [a_1, \dots, a_k]^T$, $\underline{b} \leftrightarrow [b_1, \dots, b_k]^T$, dan is $\underline{a} + \underline{b} \leftrightarrow [a_1 + b_1, \dots, a_k + b_k]^T$, $\alpha \underline{a} \leftrightarrow [\alpha a_1, \dots, \alpha a_k]^T$.

Iedere k -dimensionale lineaire ruimte ziet er precies zo uit als (we zeggen: is isomorf met) \mathbb{R}^k .

Dit resultaat betekent dat we ons in de praktijk kunnen beperken tot \mathbb{R}^k ($k \in \mathbb{N}$), hetgeen we van nu af meestal zullen doen.

7.6.4. Het bepalen van een basis, de normaalvorm van een matrix

In deze paragraaf zullen we een rekenprocedure bespreken die voor vele doeleinden bruikbaar zal zijn. Op deze plaats gaat het er allereerst om een basis te vinden voor $V := \langle \underline{v}_1, \dots, \underline{v}_k \rangle$ ($\underline{v}_1, \dots, \underline{v}_k \in \mathbb{R}^n$) en liefst zodanig, dat we voor een willekeurige $\underline{x} \in \mathbb{R}^n$ meteen kunnen vaststellen of $\underline{x} \in V$.

We gaan hierbij te werk volgens de veegtechniek uit 7.6.2, d.w.z. we maken gebruik van de eigenschap dat V niet verandert bij een van de volgende manipulaties met het stelsel voortbrengenden: verandering van de volgorde van de \underline{v} 's; vermenigvuldiging van één der \underline{v} 's met een getal $\alpha \neq 0$; vervanging van \underline{v}_i door $\underline{v}_i + \beta \underline{v}_j$ ($\beta \in \mathbb{R}$); het weglaten van $\underline{0}$ en het weglaten van een vector die een lineaire combinatie is van de andere vectoren. We bespreken eerst een voorbeeld.

Voorbeeld. Zij $\underline{u} = [0,0,1,1,1]^T$, $\underline{v} = [0,2,3,5,7]^T$, $\underline{w} = [0,3,6,2,-2]^T$, $\underline{x} = [0,1,2,3,4]^T$, $\underline{y} = [0,0,0,0,1]^T$. Bepaal een basis van $V := \langle \underline{u}, \underline{v}, \underline{w}, \underline{x}, \underline{y} \rangle \subset \mathbb{R}^5$.

Schrijf de vectoren onder elkaar en veeg de tweede kolom schoon:

$$\begin{array}{ll} \underline{x}^T = [0,1,2,3,4] & \underline{x}^T = [0,1,2,3,4] \\ \underline{v}^T = [0,2,3,5,7] & \underline{v}^T - 2\underline{x}^T = [0,0,-1,-1,-1] \\ \underline{w}^T = [0,3,6,2,-2] & \underline{w}^T - 3\underline{x}^T = [0,0,0,-7,-14] \\ \underline{u}^T = [0,0,1,1,1] & \underline{u}^T = [0,0,1,1,1] \\ \underline{y}^T = [0,0,0,0,1] & \underline{y}^T = [0,0,0,0,1] \end{array}$$

We zien nu: $V = \langle \underline{x}, \underline{v} - 2\underline{x}, \underline{w} - 3\underline{x}, \underline{u}, \underline{y} \rangle = \langle \underline{x}, \underline{u}, -\frac{1}{7}(\underline{w} - 3\underline{x}), \underline{y} \rangle$.

Van dit laatste viertal voortbrengenden is onmiddellijk te zien dat het lineair onafhankelijk en dus een basis is. We krijgen een fraaiere basis als we nog even verder werken.

Zij $\underline{t} := -\frac{1}{7}(\underline{w} - 3\underline{x})$, dan is:

$$\begin{array}{ll} \underline{x}^T = [0,1,2,3,4] & \underline{x}^T - 2\underline{u}^T = [0,1,0,1,2] \\ \underline{u}^T = [0,0,1,1,1] & \underline{u}^T = [0,0,1,1,1] \\ \underline{t}^T = [0,0,0,1,2] & \underline{t}^T = [0,0,0,1,2] \\ \underline{y}^T = [0,0,0,0,1] & \underline{y}^T = [0,0,0,0,1] \end{array}$$

$$\begin{array}{l}
 \underline{x}^T - 2\underline{u}^T - \underline{t}^T = [0, 1, 0, 0, 0] \\
 \underline{u}^T - \underline{t}^T = [0, 0, 1, 0, -1] \\
 \underline{t}^T = [0, 0, 0, 1, 2] \\
 \underline{y}^T = [0, 0, 0, 0, 1]
 \end{array}
 \qquad
 \begin{array}{l}
 \underline{x}^T - 2\underline{u}^T - \underline{t}^T = [0, 1, 0, 0, 0] \\
 \underline{u}^T - \underline{t}^T + \underline{y}^T = [0, 0, 1, 0, 0] \\
 \underline{t}^T - 2\underline{y}^T = [0, 0, 0, 1, 0] \\
 \underline{y}^T = [0, 0, 0, 0, 1]
 \end{array}
 .$$

Conclusie: $V = \langle \underline{e}_2, \underline{e}_3, \underline{e}_4, \underline{e}_5 \rangle$.

Het met de hand uitvoeren van het veegproces kan op veel manieren; de ene manier is handiger dan de andere. We willen nu het proces beschrijven op een meer systematische manier, die in de praktijk tot een computerprogramma zou kunnen worden uitgebouwd. We wijzen er met klem op dat de rekentechniek van het vegen het fundament is waar de hele verdere ontwikkeling op rust.

Terwille van later gebruik zullen we het geheel formuleren als een matrix-algoritme.

Laat $\underline{v}_1, \dots, \underline{v}_m$ vectoren uit \mathbb{R}^n zijn. Vorm dan een matrix $A^{(m \times n)}$ waarvan de rijen de vectoren $\underline{v}_1^T, \dots, \underline{v}_m^T$ zijn. We kunnen nu het veegprogramma uitvoeren voor de rijen van A . Voor later gebruik zullen we nu de rijen $\underline{0}^T$ niet weglaten. De overige veegbewerkingen zijn, geformuleerd voor rijen van A :

- (α) Verwisseling van de k -de en de ℓ -de rij.
- (β) Vermenigvuldiging van de k -de rij met $\alpha \in \mathbb{R}$, $\alpha \neq 0$.
- (γ) Vervanging van de k -de rij door de som van de k -de rij en β maal de ℓ -de rij ($\beta \in \mathbb{R}$).

Het standaardveegprogramma is nu het volgende:

- I. Begin bij de eerste kolom van A ; is deze $\underline{0}$ ($\in \mathbb{R}^m$), neem dan de tweede kolom; laat k_1 het rangnummer zijn van de eerste kolom van links af die $\neq \underline{0}$ is (als k_1 niet bestaat is $A^{(m \times n)} = \emptyset^{(m \times n)}$); verwissel, zo nodig, rijen zodat het eerste element van de k_1 -de kolom $\neq 0$ wordt; deel daarna de eerste rij door dit element.

Als bijv. $k_1 = 2$, $m = 4$, $n = 7$, dan heeft de matrix de vorm

$$\begin{bmatrix}
 0 & 1 & \cdot & \cdot & \cdot & \cdot & \cdot \\
 0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
 0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\
 0 & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot
 \end{bmatrix},$$

waarbij \cdot een element aanduidt waarvan we nog niets weten.

Trek veelvouden van de eerste rij van de tweede t/m m-de rij af zō dat de k_1 -de kolom de \underline{e}_1 uit \mathbb{R}^m wordt ("veeg de k_1 -de kolom schoon met de eerste rij"). A wordt dan

$$\begin{bmatrix} 0 & 1 & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & \cdot & \cdot & \cdot & \cdot & \cdot \end{bmatrix} .$$

- II. Als in de nu verkregen matrix de tweede t/m m-de rij alle gelijk $\underline{0}^T$ zijn, dan zijn we klaar. Laat anders k_2 het rangnummer zijn van de eerste kolom van links af, die een element $\neq 0$ heeft in de tweede t/m m-de rij (gevolg $k_2 > k_1$); verwissel, zo nodig, rijen zodat het tweede element van de k_2 -de kolom $\neq 0$ wordt; deel de tweede rij door dit element; veeg de k_2 -de kolom schoon met de tweede rij. De matrix wordt dan (gesteld $k_2 = 4$)

$$\begin{bmatrix} 0 & 1 & \cdot & 0 & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 1 & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & 0 & \cdot & \cdot & \cdot \end{bmatrix} .$$

- III. Herhaal de procedure nu; k_3 is het rangnummer van de nu meest linkse kolom die ergens in de derde t/m m-de rij een element $\neq 0$ heeft. Verwissel, zo nodig, rijen zodat het element in de derde rij, k_3 -de kolom $\neq 0$ wordt; deel de derde rij door dit element; veeg de k_3 -de kolom schoon. Daarna met k_4 , enz. Op zeker moment hebben we het volgende bereikt: de k_r -de kolom is in \underline{e}_r ($\in \mathbb{R}^m$) veranderd; verdergaan is niet mogelijk, hetzij doordat $r = m$, hetzij doordat $r < m$ en de rijen met nummers $r+1, \dots, m$ gelijk $\underline{0}^T$ zijn. We zijn nu klaar; A heeft zijn zgn.

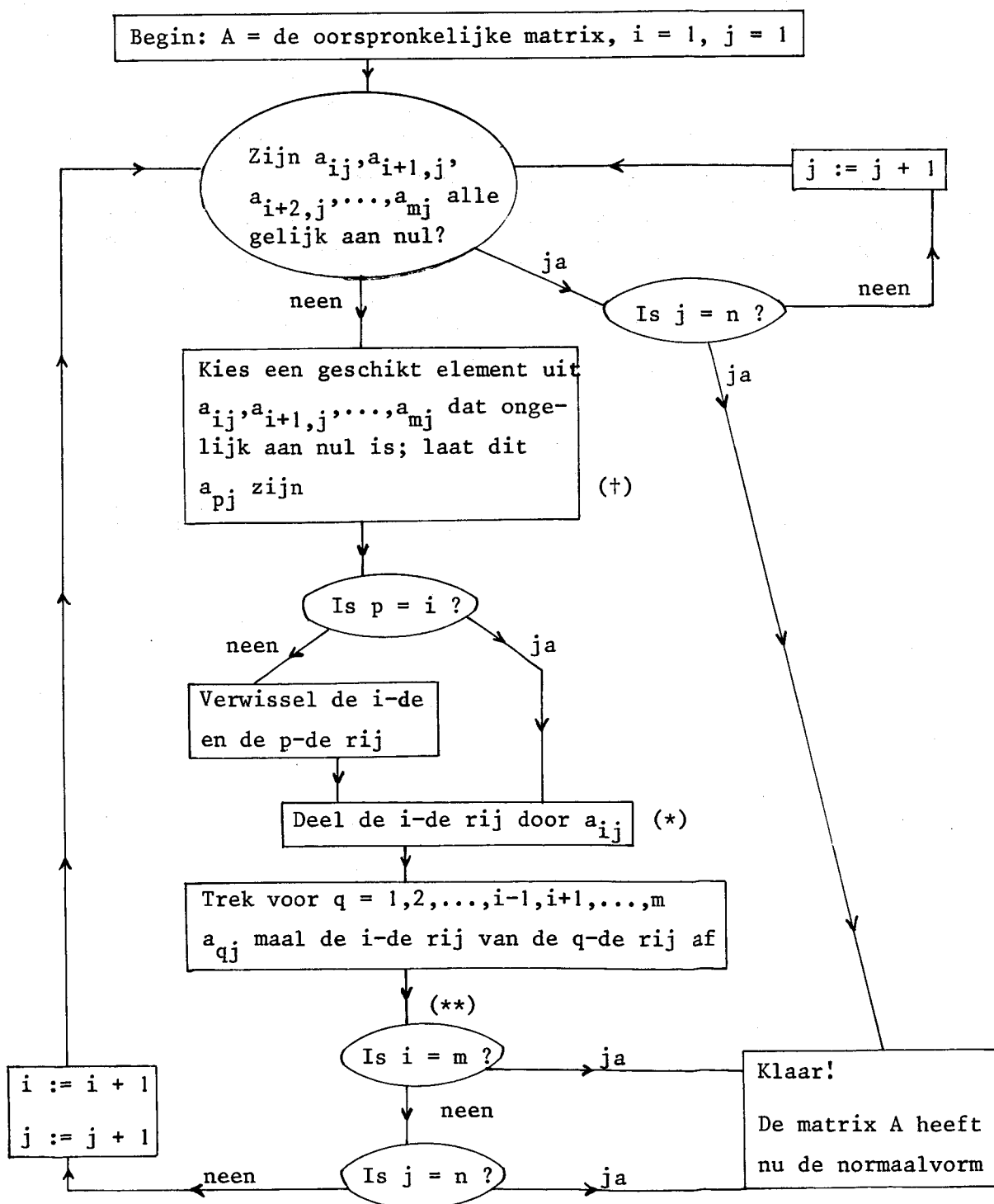
normaalvorm gekregen, bijv.

$$\begin{bmatrix} 0 & 1 & \cdot & 0 & 0 & \cdot & \cdot \\ 0 & 0 & 0 & 1 & 0 & \cdot & \cdot \\ 0 & 0 & 0 & 0 & 1 & \cdot & \cdot \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} , \quad k_1 = 2, k_2 = 4, k_3 = 5 ;$$

of

$$\begin{bmatrix} 0 & 1 & \cdot & 0 & 0 & 0 & \cdot \\ 0 & 0 & 0 & 1 & 0 & 0 & \cdot \\ 0 & 0 & 0 & 0 & 1 & 0 & \cdot \\ 0 & 0 & 0 & 0 & 0 & 1 & \cdot \end{bmatrix} , \quad k_1 = 2, k_2 = 4, k_3 = 5, k_4 = 6 .$$

We geven dit programma weer in een zgn. blokdiagram. Dit blokdiagram dient volgens de pijlrichting te worden doorlopen tot men bij klaar uitkomt. Op elk moment heeft men op papier of in het geheugen van de computer een matrix $A = A^{(m \times n)}$ en twee natuurlijke getallen i en j . Een rond hokje duidt op een vraag, een rechthoek op een opdracht. Een opdracht is steeds het veranderen van A , i of j . De oude waarden van deze grootheden worden dan vervangen door de nieuwe. Met $j := j + 1$ bedoelen we: maak de nieuwe waarde van j gelijk aan de oude plus 1; $j := j + 1$ is dus de opdracht "verhoog j met 1". We beginnen met $A =$ de oorspronkelijke matrix, $i = 1$ en $j = 1$.



Bij het vegen meten stelsels vgl opgelost.

- 82 -

(+) Als we dit programma uitvoeren op kladpapier, dan zal "geschikt" bij voorkeur +1 of -1 zijn of in ieder geval zo dat we weinig breuken krijgen. Als het voor een computer gebruikt wordt moeten we de keuzeprocedure voorschrijven; meestal neemt men dan het element met grootste absolute waarde.

Als we bovendien de kolomnummers k_1, \dots, k_r van de vectoren $\underline{e}_1, \dots, \underline{e}_r$ (uit \mathbb{R}^m) willen vastleggen, kunnen we dit doen door op de plaats aangeduid met (**) de opdracht: zet $k_i = j$ in te lassen.

De verkregen normaalvorm kan nu beschreven worden met:

- 1) Er zijn getallen $1 \leq k_1 < k_2 < \dots < k_r \leq n$ en de k_i -de kolom is \underline{e}_i uit \mathbb{R}^m .
- 2) Iedere kolom met rangnummer $< k_1$ is $\underline{0}$; iedere kolom met rangnummer k met $k_i < k < k_{i+1}$ heeft nullen op de laatste $(m - i)$ plaatsen; iedere kolom met rangnummer groter dan k_r heeft nullen op de laatste $(m - r)$ plaatsen.

Gevolg. De eerste r rijen zijn $\neq \underline{0}^T$; de laatste $m - r$ rijen zijn gelijk $\underline{0}^T$. De getransponeerden van deze eerste r rijen vormen een basis van $\langle \underline{v}_1, \dots, \underline{v}_m \rangle$.

- Opmerkingen. 1) Men kan bewijzen dat door de beschreven algoritme (onafhankelijk van de keuze bij (+)) de normaalvorm van A eënduidig bepaald is.
- 2) In de praktijk van het vraagstukken maken, zal men er soms de voorkeur aan geven te vegen met elementen $\neq 1$, d.w.z. men zal (*) nalaten en de volgende opdracht aanpassen.

7.6.5. Orthonormale bases in \mathbb{R}^n

Definitie. Een stelsel vectoren $\underline{v}_1, \dots, \underline{v}_k$ in \mathbb{R}^n heet orthonormaal, als voor $i, j = 1, \dots, k$ geldt: $(\underline{v}_i, \underline{v}_j) = \delta_{ij}$ (7.2.8; in woorden: De vectoren \underline{v}_i hebben lengte 1 en zijn onderling loodrecht).

Stelling. Een orthonormaal stelsel is onafhankelijk.

Bewijs. Zij $r_1 \underline{v}_1 + \dots + r_k \underline{v}_k = \underline{0}$, dan is

$$(r_1 \underline{v}_1 + \dots + r_k \underline{v}_k, \underline{v}_i) = (\underline{0}, \underline{v}_i) = 0 \quad (i = 1, \dots, k) .$$

Als $\underline{v}_1, \dots, \underline{v}_k$ orthonormaal is, is echter $(r_1 \underline{v}_1 + \dots + r_k \underline{v}_k, \underline{v}_i) = r_i$, dus $r_i = 0$ ($i = 1, \dots, k$). □

De basis $\underline{e}_1, \dots, \underline{e}_n$ van \mathbb{R}^n is tevens een orthonormaal stelsel; we noemen een basis die tevens een orthonormaal stelsel is een orthonormale basis.

Stelling. Elke deelruimte L van \mathbb{R}^n heeft een orthonormale basis.

Bewijs. Laat $\underline{u}_1, \dots, \underline{u}_k$ een basis van L zijn. We construeren uit deze basis een orthonormale basis door middel van het zgn. orthonormalisatieproces van Gram en Schmidt: We definiëren

$$\underline{v}_1 := \frac{1}{|\underline{u}_1|} \underline{u}_1 ; \quad \underline{w}_2 := \underline{u}_2 - (\underline{u}_2, \underline{v}_1) \underline{v}_1 ; \quad \underline{v}_2 := \frac{1}{|\underline{w}_2|} \underline{w}_2 .$$

Nu is $|\underline{v}_1| = |\underline{v}_2| = 1$; $(\underline{v}_1, \underline{v}_2) = 0$. Merk op dat $(\underline{u}_2, \underline{v}_1) \underline{v}_1$ de projectie van \underline{u}_2 is op de rechte door \underline{u}_1 (zie ook 7.7.2, voorbeeld 4).

Vervolgens voeren we in

$$\underline{w}_3 := \underline{u}_3 - (\underline{u}_3, \underline{v}_1) \underline{v}_1 - (\underline{u}_3, \underline{v}_2) \underline{v}_2 ; \quad \underline{v}_3 := \frac{1}{|\underline{w}_3|} \underline{w}_3$$

(merk op dat $|\underline{w}_3| \neq 0$ is). Ga na dat $(\underline{v}_3, \underline{v}_1) = (\underline{v}_3, \underline{v}_2) = 0$ en dat $(\underline{u}_3, \underline{v}_1) \underline{v}_1 + (\underline{u}_3, \underline{v}_2) \underline{v}_2$ de projectie is van \underline{u}_3 op het vlak door \underline{u}_1 en \underline{u}_2 . We zetten dit proces voort

$$\underline{w}_4 := \underline{u}_4 - (\underline{u}_4, \underline{v}_1) \underline{v}_1 - (\underline{u}_4, \underline{v}_2) \underline{v}_2 - (\underline{u}_4, \underline{v}_3) \underline{v}_3 ; \quad \underline{v}_4 := \frac{1}{|\underline{w}_4|} \underline{w}_4 ,$$

enz. Op deze wijze verkrijgt men een orthonormale basis $\underline{v}_1, \dots, \underline{v}_k$ van L . Met bovenstaande constructie is tevens $\underline{v}_1, \dots, \underline{v}_m$ een basis van $\langle \underline{u}_1, \dots, \underline{u}_m \rangle$ voor $m = 1, \dots, k$. □

Stelling. Zij L een k -dimensionale deelruimte van \mathbb{R}^n , dan heeft \mathbb{R}^n een orthonormale basis $\underline{v}_1, \dots, \underline{v}_k, \underline{v}_{k+1}, \dots, \underline{v}_n$ zodanig dat $\underline{v}_1, \dots, \underline{v}_k$ een orthonormale basis van L is en $\underline{v}_{k+1}, \dots, \underline{v}_n$ een orthonormale basis van L^\perp is.

Bewijs. Pas de uitbreidingsstelling (7.6.3.6) toe en daarna het orthonormalisatieproces van Gram en Schmidt. □

Gevolg 1. Is L een deelruimte van \mathbb{R}^n , dan is elke $\underline{x} \in \mathbb{R}^n$ op eenduidige wijze te schrijven als $\underline{x} = \underline{y} + \underline{z}$ met $\underline{y} \in L$, $\underline{z} \in L^\perp$.

Gevolg 2. Is L een deelruimte van \mathbb{R}^n , dan is

$$\dim L + \dim L^\perp = n ; \quad (L^\perp)^\perp = L .$$

Stelling. Zij L een k -dimensionale deelruimte van \mathbb{R}^n en $\underline{v}_1, \dots, \underline{v}_k$ een orthonormale basis van L , dan geldt voor alle $\underline{x} \in L$, $\underline{y} \in L$:

1)
$$\underline{x} = \sum_{i=1}^k (\underline{x}, \underline{v}_i) \underline{v}_i ,$$

2)
$$(\underline{x}, \underline{y}) = \sum_{i=1}^k (\underline{x}, \underline{v}_i) (\underline{y}, \underline{v}_i) \quad (\text{Parseval}),$$

*orthonormale basis, draagt
zieke eenheidsmatrix*

3)
$$|\underline{x}|^2 = (\underline{x}, \underline{x}) = \sum_{i=1}^k (\underline{x}, \underline{v}_i)^2 \quad (\text{Pythagoras}).$$

Bewijs. 1) Omdat $\underline{v}_1, \dots, \underline{v}_k$ basis is kan \underline{x} geschreven worden als $\underline{x} = \sum_{i=1}^k x_i \underline{v}_i$.
Nu is

$$(\underline{x}, \underline{v}_j) = \left(\sum_{i=1}^k x_i \underline{v}_i, \underline{v}_j \right) = \sum_{i=1}^k x_i (\underline{v}_i, \underline{v}_j) = x_j \quad (j = 1, \dots, k) .$$

2)
$$(\underline{x}, \underline{y}) = \left(\underline{x}, \sum_{i=1}^k (\underline{y}, \underline{v}_i) \underline{v}_i \right) = \sum_{i=1}^k (\underline{x}, \underline{v}_i) (\underline{y}, \underline{v}_i) .$$

3) Neem $\underline{x} = \underline{y}$ in het voorgaande resultaat. □

De stelling van Parseval drukt uit dat, als $[x_1, \dots, x_k]^T$ de kolom van \underline{x} is t.o.v. een orthonormale basis en $[y_1, \dots, y_k]^T$ de kolom van \underline{y} t.o.v. dezelfde orthonormale basis, voor het inwendig produkt geldt:

$$(\underline{x}, \underline{y}) = \sum_{i=1}^k x_i y_i .$$

"Ook wat betreft het inwendig produkt is een k -dimensionale deelruimte van \mathbb{R}^n isomorf met \mathbb{R}^k ."

ma 10-2-78

7.7. Lineaire afbeeldingen

7.7.1. Inleiding.

In deze paragraaf zullen we een type afbeeldingen van vectorruimten in vectorruimten (genaamd lineaire afbeeldingen) bestuderen die de eigenschappen 1) uit 7.2.9 delen met de afbeeldingen $\underline{x} \mapsto A\underline{x}$ van \mathbb{R}^n in \mathbb{R}^m (hierbij is A een $(m \times n)$ -matrix). In 7.7.2 geven we de definitie, enkele eigenschappen en voorbeelden voor het geval van willekeurige vectorruimten. In 7.7.3 zullen we aantonen dat elke lineaire afbeelding van \mathbb{R}^n in \mathbb{R}^m verkregen kan worden als vermenigvuldiging met een matrix; voor de bepaling van die matrix zullen we de rekenprocedure uit 7.6.4 gebruiken.

Nummer 7.7.4 geeft een aantal uiterst belangrijke stellingen o.a. over de dimensies van de rijenruimte en kolommenruimte van een matrix; deze stellingen vinden toepassing bij stelsels vergelijkingen.

7.7.2. Lineaire afbeeldingen, definities, eigenschappen, voorbeelden

Definitie. Laat L en M vectorruimten zijn. Een afbeelding $T: L \rightarrow M$ heet lineair indien voor alle $\underline{x}, \underline{y} \in L$ en alle $r \in \mathbb{R}$ geldt:

$$1) \quad T(\underline{x} + \underline{y}) = T\underline{x} + T\underline{y} ,$$

$$2) \quad T(r\underline{x}) = rT\underline{x} .$$

(Merk op dat we het beeld van \underline{x} meestal zonder haken noteren: $T\underline{x}$.)

Voorbeeld 1. Zij $A^{(m \times n)}$ een matrix, dan is $\underline{x} \mapsto A\underline{x}$ een lineaire afbeelding van \mathbb{R}^n in \mathbb{R}^m . (Zie 7.2.9.)

Definitie. De verzameling $TL := \{T\underline{x} \mid \underline{x} \in L\} \subset M$ heet de beeldruimte van T . De verzameling $N(T) := \{\underline{x} \in L \mid T\underline{x} = \underline{0}\} \subset L$ heet de nulruimte van T (of ook wel de kern van T).

Een lineaire afbeelding van L in \mathbb{R} heet ook wel een lineaire functionaal.

Eigenschappen. Als $T : L \rightarrow M$ een lineaire afbeelding is, dan geldt:

- 1) $T(r_1 \underline{a}_1 + r_2 \underline{a}_2 + \dots + r_k \underline{a}_k) = r_1 T \underline{a}_1 + r_2 T \underline{a}_2 + \dots + r_k T \underline{a}_k$ ($\underline{a}_i \in L, r_i \in \mathbb{R}$);
in het bijzonder is $T \underline{0} = \underline{0}$ en $T(-\underline{a}) = -T \underline{a}$.
- 2) TL is een deelruimte van M .
- 3) $N(T)$ is een deelruimte van L .
- 4) T is dan en slechts dan injectief als $N(T) = \{\underline{0}\}$.
- 5) Is $\underline{a} \in L, \underline{b} \in M, T \underline{a} = \underline{b}$ dan is $\{\underline{x} \in L \mid T \underline{x} = \underline{b}\}$ de nevenruimte $\underline{a} + N(T)$.
- 6) Als T bijectief is dan is de inverse afbeelding $T^{-1} : M \rightarrow L$ lineair.
- 7) Als $S : M \rightarrow N$ een lineaire afbeelding van M in de vectorruimte N is, dan is de samengestelde afbeelding $S \circ T : L \rightarrow N$ lineair.

Alle bewijzen zijn eenvoudige verificaties analoog aan 7.2.9.

Voorbeelden. 2) Hoewel we ons spoedig tot \mathbb{R}^n zullen beperken geven we eerst een paar voorbeelden voor vectorruimten waarvan de elementen functies zijn. Dit onderstreept de algemeenheid van het begrip lineaire afbeelding.

$$\text{Int}(f) := \int_0^1 f(t) dt$$

definieert een lineaire functionaal op de vectorruimte van de functies die integreerbaar zijn over $[0,1]$.

Als $T_a f$ gedefinieerd wordt door $(T_a f)(x) := f(x - a)$ is T_a een lineaire afbeelding van de vectorruimte van alle functies op \mathbb{R} in dezelfde vectorruimte.

Als $T_L f$ gedefinieerd wordt door

$$(T_L f)(x) = \int_0^{\infty} f(t) e^{-xt} dt$$

dan is T_L een lineaire afbeelding van een deelruimte van de vectorruimte van de functies op $[0, \infty)$ in de vectorruimte van de functies op $[0, \infty)$.

Als Df gedefinieerd wordt door $(Df)(x) = f'(x)$ dan is D een lineaire afbeelding van de vectorruimte der differentieerbare functies in de vectorruimte van alle functies.

De volgende voorbeelden zijn ontleend aan de meetkunde in \mathbb{R}^2 en \mathbb{R}^3 .

- 3) Zij V in \mathbb{R}^3 het vlak met vergelijking $(\underline{a}, \underline{x}) = 0$; gemakshalve nemen we aan dat de normaalvector \underline{a} de lengte 1 heeft dus $(\underline{a}, \underline{a}) = \underline{a}^T \underline{a} = 1$. De projectie van \underline{x} op V is $\underline{x} - (\underline{a}, \underline{x})\underline{a}$. De afbeelding $P_{\underline{a}} : \underline{x} \mapsto \underline{x} - (\underline{a}, \underline{x})\underline{a}$ is een lineaire afbeelding van \mathbb{R}^3 in \mathbb{R}^3 met V als beeldruimte en $\{\underline{ra} \mid r \in \mathbb{R}\}$ als nulruimte. We noemen $P_{\underline{a}}$ ook projectie. Omdat $(\underline{a}, \underline{x})\underline{a} = \underline{aa}^T \underline{x}$ kunnen we $P_{\underline{a}}$ voorstellen als matrixvermenigvuldiging met $(I - \underline{aa}^T)$. Voor $P_{\underline{a}} \circ P_{\underline{a}}$ geldt dan op grond van de associativiteit van de matrixvermenigvuldiging:

$$\begin{aligned} (P_{\underline{a}} \circ P_{\underline{a}})\underline{x} &= (I - \underline{aa}^T)(I - \underline{aa}^T)\underline{x} = (I - 2\underline{aa}^T + \underline{aa}^T \underline{aa}^T)\underline{x} = \\ &= (I - \underline{aa}^T)\underline{x} = P_{\underline{a}}\underline{x}. \end{aligned}$$

Dus $P_{\underline{a}} \circ P_{\underline{a}} = P_{\underline{a}}$; we noemen afbeeldingen met deze eigenschap idempotent.

- 4) Zij ℓ in \mathbb{R}^3 de rechte door $\underline{0}$ met parametervoorstelling $\underline{x} = \lambda \underline{b}$, waarbij gemakshalve $(\underline{b}, \underline{b}) = \underline{b}^T \underline{b} = 1$. Laat $Q_{\underline{b}}$ de projectie op ℓ zijn; deze is een lineaire afbeelding van \mathbb{R}^3 in \mathbb{R}^3 met ℓ als beeldruimte en het vlak V met vergelijking $(\underline{b}, \underline{x}) = 0$ als nulruimte.

Ook $Q_{\underline{b}}$ is als matrixvermenigvuldiging op te vatten want

$$Q_{\underline{b}}\underline{x} = (\underline{b}, \underline{x})\underline{b} = \underline{bb}^T \underline{x}.$$

Er geldt: $Q_{\underline{b}} \circ Q_{\underline{b}} = Q_{\underline{b}}$, hetgeen meetkundig duidelijk is, maar ook volgt

$$\text{uit } (\underline{bb}^T)(\underline{bb}^T) = \underline{b}(\underline{b}^T \underline{b})\underline{b}^T = \underline{bb}^T.$$

- 5) Zij $s_{\underline{a}}$ de spiegeling van \mathbb{R}^3 aan het vlak V met vergelijking $(\underline{a}, \underline{x}) = 0$; $(\underline{a}, \underline{a}) = 1$. $S_{\underline{a}}$ is een bijectieve lineaire afbeelding en $S_{\underline{a}}^{\star} = S_{\underline{a}}$ (afbeeldingen met deze eigenschap heten involuties). Dit laatste kunnen we ook met matrices bewijzen want

$$S_{\underline{a}}\underline{x} = \underline{x} - 2(\underline{a}, \underline{x})\underline{a} = (I - 2\underline{aa}^T)\underline{x}$$

en

$$(I - 2\underline{aa}^T)^2 = I - 4\underline{aa}^T + 4\underline{aa}^T \underline{aa}^T = I.$$

6) $A_a: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ zij gedefinieerd door $\underline{x} = [x, y]^T \rightarrow [x + ay, y]^T$.

A_a is een lineaire afbeelding geheten afschuiving. $A_a \underline{x}$ is te schrijven als matrixprodukt:

$$A_a \underline{x} = \begin{bmatrix} 1 & a \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}.$$

7) De draaiing D_φ van \mathbb{R}^2 over een hoek φ om de oorsprong is een lineaire afbeelding. $D_\varphi^{-1} = D_{-\varphi}$; $D_\varphi \circ D_\psi = D_{\varphi+\psi}$.

Ook D_φ is als matrixvermenigvuldiging voor te stellen, nl.

$$D_\varphi \underline{x} = \begin{bmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}.$$

7.7.3. De matrix van een lineaire afbeelding $\mathbb{R}^n \rightarrow \mathbb{R}^m$

Elk van de voorbeelden 3 t/m 7 uit 7.7.2 bleek als een speciaal geval van voorbeeld 1 opgevat te kunnen worden. Dit geldt voor alle lineaire afbeeldingen van \mathbb{R}^n in \mathbb{R}^m .

Stelling. Laat $A: \mathbb{R}^n \rightarrow \mathbb{R}^m$ lineair zijn. Als voor de vectoren $\underline{e}_1, \underline{e}_2, \dots, \underline{e}_n \in \mathbb{R}^n$ van de standaardbasis geldt

$$A\underline{e}_1 = \begin{bmatrix} a_{11} \\ \vdots \\ a_{m1} \end{bmatrix}, \quad A\underline{e}_2 = \begin{bmatrix} a_{12} \\ \vdots \\ a_{m2} \end{bmatrix}, \quad \dots, \quad A\underline{e}_n = \begin{bmatrix} a_{1n} \\ \vdots \\ a_{mn} \end{bmatrix},$$

dan is voor elke $\underline{x} \in \mathbb{R}^n$:

$$A\underline{x} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \vdots & \vdots & \dots & \vdots \\ \vdots & \vdots & \dots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

Bewijs. Voor elke $\underline{x} \in \mathbb{R}^n$ geldt: $\underline{x} = x_1 \underline{e}_1 + x_2 \underline{e}_2 + \dots + x_n \underline{e}_n$, zodat

$A\underline{x} = x_1 A\underline{e}_1 + x_2 A\underline{e}_2 + \dots + x_n A\underline{e}_n$. Op grond van de definitie van matrixvermenigvuldiging is dit juist hetgeen te bewijzen is. □

Gevolg. Voor elke lineaire functionaal $\mathbb{R}^n \rightarrow \mathbb{R}$ is er een $\underline{a} = [a_1, \dots, a_n]^T \in \mathbb{R}^n$ zodat de functionaal is $\underline{x} \mapsto (\underline{a}, \underline{x})$.

De stelling geeft aan hoe de matrix van een lineaire afbeelding opgebouwd wordt uit de beelden van de standaardbasis. Als we de beelden van een andere basis geven, dan is de matrix te bepalen op de wijze als hierna besproken aan de hand van een $\mathbb{R}^3 \rightarrow \mathbb{R}^4$ afbeelding.

Voorbeeld 1. Laat van A gegeven zijn

$$A \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ 10 \\ 5 \end{bmatrix}, \quad A \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \\ 8 \\ 0 \end{bmatrix}, \quad A \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} = \begin{bmatrix} 3 \\ 1 \\ 6 \\ 1 \end{bmatrix}.$$

We zullen door vegen de beelden van de standaardbasis van \mathbb{R}^3 bepalen. Gemakshalve schrijven we alles als rijen. Gegeven is

$$\begin{aligned} [0 \quad 1 \quad 1]^T &\mapsto [1 \quad 3 \quad 10 \quad 5]^T \\ [1 \quad 1 \quad 1]^T &\mapsto [2 \quad 2 \quad 8 \quad 0]^T \\ [1 \quad 0 \quad 2]^T &\mapsto [3 \quad 1 \quad 6 \quad 1]^T. \end{aligned}$$

Hieruit volgt

$$\begin{aligned} [1 \quad 1 \quad 1]^T &\mapsto [2 \quad 2 \quad 8 \quad 0]^T \\ [0 \quad 1 \quad 1]^T &\mapsto [1 \quad 3 \quad 10 \quad 5]^T \\ [1 \quad 0 \quad 2]^T &\mapsto [3 \quad 1 \quad 6 \quad 1]^T \end{aligned}$$

en dus

$$\begin{aligned} [1 \quad 1 \quad 1]^T &\mapsto [2 \quad 2 \quad 8 \quad 0]^T \\ [0 \quad 1 \quad 1]^T &\mapsto [1 \quad 3 \quad 10 \quad 5]^T \\ [0 \quad -1 \quad 1]^T &\mapsto [1 \quad -1 \quad -2 \quad 1]^T. \end{aligned}$$

Verder vegend vinden we

$$\begin{aligned} [1 \quad 0 \quad 0]^T &\mapsto [1 \quad -1 \quad -2 \quad -5]^T \\ [0 \quad 1 \quad 1]^T &\mapsto [1 \quad 3 \quad 10 \quad 5]^T \\ [0 \quad 0 \quad 2]^T &\mapsto [2 \quad 2 \quad 8 \quad 6]^T, \end{aligned}$$

vervolgens

$$\begin{aligned} [1 \ 0 \ 0]^T &\mapsto [1 \ -1 \ -2 \ -5]^T \\ [0 \ 1 \ 1]^T &\mapsto [1 \ 3 \ 10 \ 5]^T \\ [0 \ 0 \ 1]^T &\mapsto [1 \ 1 \ 4 \ 3]^T \end{aligned}$$

en tenslotte

$$\begin{aligned} [1 \ 0 \ 0]^T &\mapsto [1 \ -1 \ -2 \ -5]^T \\ [0 \ 1 \ 0]^T &\mapsto [0 \ 2 \ 6 \ 2]^T \\ [0 \ 0 \ 1]^T &\mapsto [1 \ 1 \ 4 \ 3]^T \end{aligned}$$

De matrix van A is dus

$$\begin{bmatrix} 1 & 0 & 1 \\ -1 & 2 & 1 \\ -2 & 6 & 4 \\ -5 & 2 & 3 \end{bmatrix}$$

(controleer het antwoord).

Het is duidelijk dat we precies het rekenproces van 7.6.4 uitvoeren. Dit leidt tot het volgende programma voor de bepaling van de matrix van een lineaire afbeelding $A : \mathbb{R}^n \rightarrow \mathbb{R}^m$ als de beelden van een basis $\underline{v}_1, \dots, \underline{v}_n$ van \mathbb{R}^n gegeven zijn.

Programma. Neem de gegeven basisvectoren $\underline{v}_1, \dots, \underline{v}_n$ van \mathbb{R}^n als de rijvectoren van een matrix M, hun beelden als de overeenkomstige rijen van een matrix N. Bepaal de normaalvorm van $[M, N]$ (bloknotatie); deze is $[I, \tilde{N}]$. (Het eerste blok wordt I omdat we uitgingen van een basis van \mathbb{R}^n !) Nu is \tilde{N}^T de matrix van de gegeven lineaire afbeelding.

In ons voorbeeld:

$$\begin{aligned} [M, N] &= \begin{bmatrix} 0 & 1 & 1 & 1 & 3 & 10 & 5 \\ 1 & 1 & 1 & 2 & 2 & 8 & 0 \\ 1 & 0 & 2 & 3 & 1 & 6 & 1 \end{bmatrix}, \\ [I, \tilde{N}] &= \begin{bmatrix} 1 & 0 & 0 & 1 & -1 & -2 & -5 \\ 0 & 1 & 0 & 0 & 2 & 6 & 2 \\ 0 & 0 & 1 & 1 & 1 & 4 & 3 \end{bmatrix}. \end{aligned}$$

Voorbeeld 2. Bepaal de matrix van de projectie P van \mathbb{R}^3 op het vlak $V : x + 2y + z = 0$. Merk op dat voor vectoren in V geldt $P\underline{x} = \underline{x}$, terwijl voor vectoren loodrecht op het vlak $P\underline{x} = \underline{0}$ is. Dus

$$P \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix} = \begin{bmatrix} 2 \\ -1 \\ 0 \end{bmatrix}, \quad P \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ -1 \end{bmatrix}, \quad P \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}.$$

Volgens het zoëven beschreven programma vinden we voor de matrix van P :

$$\frac{1}{6} \begin{bmatrix} 5 & -2 & -1 \\ -2 & 2 & -2 \\ -1 & -2 & 5 \end{bmatrix}.$$

De matrix van P is ook $I - \underline{a}\underline{a}^T$, waarbij \underline{a} een vector loodrecht op V is met $|\underline{a}| = 1$ (zie 7.7.2, voorbeeld 3). We vinden $\underline{a} = \lambda[1, 2, 1]^T$ met $\lambda^2 + 4\lambda^2 + \lambda^2 = 1$, $\lambda = 1/\sqrt{6}$. Dan is

$$\underline{a}\underline{a}^T = \frac{1}{6} \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix} [1, 2, 1] = \frac{1}{6} \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix},$$

hetgeen hetzelfde resultaat oplevert voor de matrix van P .

Opmerkingen. 1) Een lineaire afbeelding van een willekeurige n -dimensionale vectorruimte L in een m -dimensionale vectorruimte M is ook met een matrix te beschrijven. Kies daartoe bases in L en M en vervang iedere vector door zijn kolom t.o.v. de gekozen basis. De gegeven lineaire afbeelding bepaalt nu een $\mathbb{R}^n \rightarrow \mathbb{R}^m$ afbeelding van deze kolommen die te beschrijven is door een $(m \times n)$ -matrix. De aldus gevonden matrix is afhankelijk van de gekozen bases.

- 2) We beperken ons in dit college verder tot afbeeldingen van \mathbb{R}^n in \mathbb{R}^m die uitsluitend beschreven worden door hun matrices t.o.v. de standaardbases. Bij een bestudering van de lineaire algebra die verder gaat dan onze eerste kennismaking zal veel aandacht besteed worden aan het bepalen van bases t.o.v. waarvan een lineaire afbeelding een "mooie" matrix heeft.
- 3) In dit college zullen we geen notatieneel onderscheid maken tussen een lineaire afbeelding en haar matrix. Door de afspraak dat we uitsluitend standaardbases gebruiken ligt de matrix éénduidig vast. Merk op dat de term beeldruimte in 7.5.5 terecht gebruikt is.

7.7.4. De rang van een matrix, dimensiestellingen

Definitie. Zij A een $(m \times n)$ -matrix. De rang van A is de dimensie van $\mathbb{A}\mathbb{R}^n$. De rang is dus de dimensie van de kolommenruimte.

7.7.4.1. Stelling (dimensiestelling). Zij $A : \mathbb{R}^n \rightarrow \mathbb{R}^m$ lineair, dan is

$$\dim N(A) + \dim \mathbb{A}\mathbb{R}^n = n .$$

Bewijs. Daar $N(A) \subset \mathbb{R}^n$ geldt $k := \dim N(A) \leq n$.

Als $k = n$ is $N(A) = \mathbb{R}^n$, dus $A = \mathcal{O}$ en $\dim \mathbb{A}\mathbb{R}^n = 0$.

Zij nu $k < n$. Laat $\underline{u}_1, \dots, \underline{u}_k$ een basis van $N(A)$ zijn. Vul deze aan (7.6.3.6) met $\underline{v}_1, \dots, \underline{v}_{n-k}$ tot een basis van \mathbb{R}^n . Daar $A\underline{u}_1 = \dots = A\underline{u}_k = \underline{0}$, is

$$\mathbb{A}\mathbb{R}^n = \langle A\underline{v}_1, \dots, A\underline{v}_{n-k} \rangle .$$

Het bewijs is voltooid als we hebben aangetoond dat $A\underline{v}_1, \dots, A\underline{v}_{n-k}$ een onafhankelijk stelsel is.

Stel

$$r_1 A\underline{v}_1 + \dots + r_{n-k} A\underline{v}_{n-k} = \underline{0} ,$$

dan is

$$A(r_1 \underline{v}_1 + \dots + r_{n-k} \underline{v}_{n-k}) = \underline{0}$$

en dus

$$r_1 \underline{v}_1 + \dots + r_{n-k} \underline{v}_{n-k} \in N(A) .$$

Daar $\underline{u}_1, \dots, \underline{u}_k$ een basis van $N(A)$ is, is

$$r_1 \underline{v}_1 + \dots + r_{n-k} \underline{v}_{n-k} = s_1 \underline{u}_1 + \dots + s_k \underline{u}_k .$$

Hieruit volgt $0 = r_1 = \dots = r_{n-k}$ ($= s_1 = \dots = s_k$) omdat $\underline{u}_1, \dots, \underline{u}_k, \underline{v}_1, \dots, \underline{v}_{n-k}$ een onafhankelijk stelsel is. □

Stelling. Zij A een $(m \times n)$ -matrix, dan is de dimensie van de rijenruimte gelijk aan de dimensie van de kolommenruimte.

Bewijs. Met 7.5.5 en 7.5.6 zien we dat de rijenruimte gelijk is aan $(N(A))^\perp$. Uit 7.6.5, gevolg 2 volgt dat $\dim(N(A))^\perp = n - \dim(N(A))$. Op grond van de dimensiestelling is dit gelijk aan $\dim A\mathbb{R}^n$. \square

Gevolg. De rang van een matrix is gelijk aan de rang van haar normaalvorm.

Gevolg. Het getal r uit 7.6.4 is de rang van de matrix.

Gevolg. Voor de oplossingsruimte $N(A)$ van een homogeen stelsel vergelijkingen in n onbekenden $A\underline{x} = \underline{0}$ geldt:

$$\dim N(A) = n - \text{rang } A .$$

7.8. Lineaire afbeeldingen $\mathbb{R}^n \rightarrow \mathbb{R}^n$

7.8.1. Inleiding

De matrix van een lineaire afbeelding $A : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is vierkant. Voor zulke afbeeldingen en matrices kan men een aantal bijzondere begrippen en eigenschappen formuleren. In 7.8.2 onderzoeken we het verband tussen de inverse afbeelding A^{-1} van A , zo deze bestaat (zie 1.1), en de inverse matrix A^{-1} van A (zie 7.2.8.2). In 7.8.3 worden de reeds in 7.3.5 en 7.3.6 genoemde eigenwaarden nader besproken. In 7.8.4 wordt een summiere behandeling gegeven van het begrip determinant. Dit begrip heeft vroeger in de lineaire algebra een belangrijke rol gespeeld. Tegenwoordig acht men determinanten van minder belang. In 7.8.5 tenslotte, zullen we spreken over orthogonale afbeeldingen, dat zijn afbeeldingen die de lengtes van vectoren niet veranderen. We zullen aantonen dat in \mathbb{R}^2 en \mathbb{R}^3 orthogonale afbeeldingen draaiingen en spiegelingen zijn.

Draaiing in plat vlak steeds bekend

7.8.2. Inverteerbaarheid

Uit 7.7.2, eigenschap 4, en de dimensiestelling (7.7.4.1) volgt dat een lineaire afbeelding van \mathbb{R}^n in \mathbb{R}^n dan en slechts dan injectief is als ze surjectief is.

Zij A een bijectieve lineaire afbeelding van \mathbb{R}^n in \mathbb{R}^n . Deze afbeelding A heeft een inverse die weer lineair is. De matrix van deze inverse is de matrix A^{-1} waarvoor $A^{-1}A = I$ (dit volgt uit het feit dat de matrix van een samengestelde afbeelding het matrixproduct van de matrices van elk der afbeeldingen is. Omdat voor afbeeldingen geldt: $(A^{\leftarrow})^{\leftarrow} = A$ is dan ook $AA^{-1} = I$, wat in 7.2.8.2 al aangekondigd is.

We releveren een aantal eigenschappen van $(n \times n)$ -matrices:

- 1) A is inverteerbaar dan en slechts dan als $N(A) = \{0\}$.
- 2) A is inverteerbaar dan en slechts dan als $A\mathbb{R}^n = \mathbb{R}^n$.
- 3) A is inverteerbaar dan en slechts dan als $\text{rang } A = n$.
- 4) AB is inverteerbaar dan en slechts dan als A en B beide inverteerbaar zijn. In dat geval is $(AB)^{-1} = B^{-1}A^{-1}$.

Voor het bepalen van de inverse matrix is de procedure van 7.7.3 (en dus van 7.6.4) heel geschikt. We bespreken één voorbeeld.

Zij

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 2 & 1 & 0 \\ 0 & 1 & 3 \end{bmatrix} .$$

Als A inverteerbaar is vormen de kolommen een basis van \mathbb{R}^3 (dit is in ons voorbeeld inderdaad het geval). Voor de inverse matrix geldt nu:

$$A^{-1} \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix} , \quad A^{-1} \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} , \quad A^{-1} \begin{bmatrix} 1 \\ 0 \\ 3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} ,$$

waaruit we berekenen

$$A^{-1} = \begin{bmatrix} -3 & 2 & 1 \\ 6 & -3 & -2 \\ -2 & 1 & 1 \end{bmatrix} .$$

college 13-2-25

7.8.3. Eigenwaarden en eigenvectoren

Definitie. Zij $A: \mathbb{R}^n \rightarrow \mathbb{R}^n$ lineair. Een getal $\lambda \in \mathbb{R}$ heet een eigenwaarde van A indien $A - \lambda I$ niet inverteerbaar is. Dit is het geval dan en slechts dan als er een $\underline{v} \in \mathbb{R}^n$, $\underline{v} \neq \underline{0}$, bestaat zó dat $A\underline{v} = \lambda\underline{v}$. Zo'n vector \underline{v} heet een eigenvector bij de eigenwaarde λ .

We spreken ook wel van eigenwaarden en eigenvectoren van de matrix A .

We hebben al verschillende malen gezien hoe eigenwaarden en eigenvectoren in de toepassingen een rol spelen. Meetkundig zijn eigenvectoren vectoren $\underline{v} \neq \underline{0}$ waarvan het beeld op de rechte door $\underline{0}$ en \underline{v} ligt.

De volgende stellingen zijn eenvoudig te bewijzen.

Stelling. Als λ een eigenwaarde van $A^{(n \times n)}$ is, dan is $\{\underline{v} \in \mathbb{R}^n \mid A\underline{v} = \lambda\underline{v}\}$ een deelruimte van \mathbb{R}^n met positieve dimensie.

De deelruimte uit deze stelling heet de eigenruimte van A bij de eigenwaarde λ .

Stelling. Een getal λ is een eigenwaarde van A dan en slechts dan als $\dim N(A - \lambda I) > 0$ is. Alle vectoren $\neq \underline{0}$ uit $N(A - \lambda I)$ zijn eigenvectoren van A bij de eigenwaarde λ .

Stelling. Zij $\underline{v}_1, \dots, \underline{v}_k$ een lineair onafhankelijk stelsel van eigenvectoren van de matrix $A^{(n \times n)}$ met $A\underline{v}_i = \lambda_i \underline{v}_i$ ($i = 1, \dots, k$). Zij \underline{w} eveneens een eigenvector van A bij de eigenwaarde μ met $\mu \neq \lambda_i$ ($i = 1, \dots, k$). Dan is $\underline{v}_1, \dots, \underline{v}_k, \underline{w}$ een lineair onafhankelijk stelsel.

Bewijs. Als $\underline{v}_1, \dots, \underline{v}_k, \underline{w}$ afhankelijk zou zijn dan zou \underline{w} een lineaire combinatie zijn van $\underline{v}_1, \dots, \underline{v}_k$ (ga na). Veronderstel dat $\underline{w} = \alpha_1 \underline{v}_1 + \dots + \alpha_k \underline{v}_k$; omdat $\underline{w} \neq \underline{0}$ (eigenvector!) is $[\alpha_1, \dots, \alpha_k] \neq [0, \dots, 0]$. Nu is

$A\underline{w} = \alpha_1 A\underline{v}_1 + \dots + \alpha_k A\underline{v}_k$, waaruit volgt $\mu \underline{w} = \alpha_1 \lambda_1 \underline{v}_1 + \dots + \alpha_k \lambda_k \underline{v}_k$. Omdat ook geldt $\mu \underline{w} = \alpha_1 \mu \underline{v}_1 + \dots + \alpha_k \mu \underline{v}_k$, is $\alpha_1 (\lambda_1 - \mu) \underline{v}_1 + \dots + \alpha_k (\lambda_k - \mu) \underline{v}_k = \underline{0}$ en in deze lineaire combinatie zijn niet alle coëfficiënten 0. Dit is in tegenspraak met de onafhankelijkheid van $\underline{v}_1, \dots, \underline{v}_k$. □

Gevolg. Een $(n \times n)$ -matrix heeft ten hoogste n verschillende eigenwaarden.

Voorbeelden. 1) Voor de afbeelding $P_{\underline{a}}$ (7.7.2, voorbeeld 3) zijn alle vectoren $\neq \underline{0}$ in het vlak $(\underline{a}, \underline{x}) = 0$ eigenvectoren bij de eigenwaarde $+1$; alle vectoren $\lambda \underline{a}$ met $\lambda \neq 0$ zijn eigenvectoren bij de eigenwaarde 0 . Uit de voorafgaande stelling volgt dat er geen andere eigenwaarden en eigenvectoren zijn.

- 2) $Q_{\underline{b}}$ uit 7.7.2, voorbeeld 4, heeft de eigenwaarden 0 en 1 . De eigenruimte bij 0 is het vlak $(\underline{b}, \underline{x}) = 0$; de eigenruimte bij 1 is $\langle \underline{b} \rangle$.
- 3) De spiegeling $S_{\underline{a}}$ uit 7.7.2, voorbeeld 5, heeft eigenwaarden 1 (met als eigenruimte het spiegelvlak $(\underline{a}, \underline{x}) = 0$) en -1 (met als eigenruimte $\langle \underline{a} \rangle$).
- 4) Voor de afschuiving $A_{\underline{a}}$ uit 7.7.2, voorbeeld 6, is 1 de enige eigenwaarde met $\langle \underline{e}_1 \rangle$ als bijbehorende eigenruimte, mits $a \neq 0$ is (ga na).
- 5) De draaiing D_{φ} (7.7.2, voorbeeld 7) heeft geen eigenwaarden en eigenvectoren tenzij $\varphi = k\pi$ (k geheel).
- 6) Als $N(A) \neq \{0\}$ is $N(A)$ eigenruimte bij de eigenwaarde 0 .

Het bepalen van eigenwaarden en eigenvectoren van $A^{(n \times n)} = [a_{ij}]$ komt neer op het vinden van een λ en een $\underline{v} \neq \underline{0}$ die voldoen aan:

$$\begin{cases} (a_{11} - \lambda)v_1 + a_{12}v_2 + \dots + a_{1n}v_n = 0 \\ a_{21}v_1 + (a_{22} - \lambda)v_2 + \dots + a_{2n}v_n = 0 \\ \vdots \\ a_{n1}v_1 + a_{n2}v_2 + \dots + (a_{nn} - \lambda)v_n = 0. \end{cases}$$

ma 17-2-75

Een rekentechniek voor de uitvoering hiervan wordt besproken in 7.8.4.

7.8.4. Determinanten

De determinant van een vierkante matrix A is een getal, $\det A$, dat we aangeven door de elementen van de matrix tussen verticale strepen gezet.

De definitie van de determinant geven we recursief naar het aantal rijen en kolommen:

- i) Als $A = [a]$ is $\det A = a$.
- ii) Stel dat voor elke $(n \times n)$ -matrix de determinant bekend is, dan definiëren we:

$$\begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} & a_{1,n+1} \\ \vdots & \vdots & & \vdots & \vdots \\ a_{n+1,1} & \cdot & \dots & \cdot & a_{n+1,n+1} \end{vmatrix} :=$$

$$= a_{11}D_{11} - a_{12}D_{12} + a_{13}D_{13} - \dots + (-1)^n a_{1,n+1}D_{1,n+1},$$

waarin D_{ij} de determinant is van de $(n \times n)$ -matrix die uit A ontstaat door de i -de rij en de j -de kolom weg te laten.

Enkele gevolgen:

a) $\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}.$

b) $\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} =$

$$= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33}.$$

Meetkundig stelt $\det[\underline{a}, \underline{b}]$, $\underline{a}, \underline{b} \in \mathbb{R}^2$, de oppervlakte of het tegenovergestelde daarvan voor van het parallellogram waarvan de meetkundige vectoren \underline{a} en \underline{b} aanliggende zijden zijn. Merk op dat

$$\det[\underline{e}_1, \underline{e}_2] = \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = 1.$$

Evenzo stelt $\det[\underline{a}, \underline{b}, \underline{c}]$, $\underline{a}, \underline{b}, \underline{c} \in \mathbb{R}^3$, weer afgezien van een eventuele factor -1 , de inhoud voor van het parallellepipedum met de meetkundige vectoren \underline{a} , \underline{b} en \underline{c} als zijden. Merk op dat

$$\det[\underline{e}_1, \underline{e}_2, \underline{e}_3] = \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{vmatrix} = 1.$$

Het is voor determinanten met veel rijen en kolommen erg omslachtig de berekening uit te voeren volgens de definitie. Immers het aantal termen dat men zo krijgt bij de uitwerking van een $(n \times n)$ -determinant is $n!$.

Stelling. 1) $\det A = \det A^T$.

2) $\det AB = \det A \cdot \det B$.

Het bewijs van deze stelling laten we achterwege. Alle belangrijke eigenschappen van determinanten volgen nu zonder moeite:

Stelling. 1) $\det A = -\det B$ als B uit A ontstaat door verwisseling van twee rijen of van twee kolommen.

2) $\det A = \det B$ als de i-de rij (of kolom) van B gelijk is aan de som van de i-de rij (of kolom) van A en p maal de j-de rij (of kolom) van A ($p \in \mathbb{R}$, $i \neq j$), terwijl de overige rijen (kolommen) van B gelijk zijn aan de overeenkomstige rijen (kolommen) van A.

3) $\det A = p \det B$ als één rij (of kolom) van A gelijk is aan p maal de overeenkomstige rij (of kolom) van B, terwijl de overige rijen (kolommen) van B gelijk zijn aan de overeenkomstige rijen (kolommen) van A.

4) $\det A^{(n \times n)} \neq 0$ dan en slechts dan als $\text{rang } A = n$, dus als de kolommen van A een basis van \mathbb{R}^n vormen.

5) $\det(pA^{(n \times n)}) = p^n \det A$.

"Bewijs". 1) Verandering van de volgorde van de rijen in een matrix kan beschouwd worden als linksvermenigvuldiging met een permutatiematrix; verwisseling van de kolommen ontstaat door rechtsvermenigvuldiging met een permutatiematrix (zie 7.2.8.4). Ook kan men kolommen in rijen overvoeren door transponeren waarna $\det A^T = \det A$. We volstaan met één voorbeeld.

$$\begin{array}{cccc} \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} & \begin{bmatrix} a_1 & a_2 & a_3 & a_4 \\ b_1 & b_2 & b_3 & b_4 \\ c_1 & c_2 & c_3 & c_4 \\ d_1 & d_2 & d_3 & d_4 \end{bmatrix} & = & \begin{bmatrix} c_1 & c_2 & c_3 & c_4 \\ b_1 & b_2 & b_3 & b_4 \\ a_1 & a_2 & a_3 & a_4 \\ d_1 & d_2 & d_3 & d_4 \end{bmatrix} \\ P & A & = & B \end{array}$$

$\det P \cdot \det A = \det B$, en $\det P = -1$.

2) We volstaan weer met een voorbeeld.

$$\begin{matrix}
 \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & p & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} & \begin{bmatrix} a_1 & a_2 & a_3 & a_4 \\ b_1 & b_2 & b_3 & b_4 \\ c_1 & c_2 & c_3 & c_4 \\ d_1 & d_2 & d_3 & d_4 \end{bmatrix} & = & \begin{bmatrix} a_1 & a_2 & a_3 & a_4 \\ b_1+pc_1 & b_2+pc_2 & b_3+pc_3 & b_4+pc_4 \\ c_1 & c_2 & c_3 & c_4 \\ d_1 & d_2 & d_3 & d_4 \end{bmatrix} \\
 T & A & = & B
 \end{matrix}$$

det T · det A = det B, en det T = 1.

Evenzo

$$\begin{bmatrix} a_1 & a_2 & a_3 & a_4 \\ b_1 & b_2 & b_3 & b_4 \\ c_1 & c_2 & c_3 & c_4 \\ d_1 & d_2 & d_3 & d_4 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & p & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} a_1 & a_2 & pa_2+a_3 & a_4 \\ b_1 & b_2 & pb_2+b_3 & b_4 \\ c_1 & c_2 & pc_2+c_3 & c_4 \\ d_1 & d_2 & pd_2+d_3 & d_4 \end{bmatrix}$$

$$3) \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & p & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a_1 & a_2 & a_3 & a_4 \\ b_1 & b_2 & b_3 & b_4 \\ c_1 & c_2 & c_3 & c_4 \\ d_1 & d_2 & d_3 & d_4 \end{bmatrix} = \begin{bmatrix} a_1 & a_2 & a_3 & a_4 \\ pb_1 & pb_2 & pb_3 & pb_4 \\ c_1 & c_2 & c_3 & c_4 \\ d_1 & d_2 & d_3 & d_4 \end{bmatrix}$$

$$\begin{bmatrix} a_1 & a_2 & a_3 & a_4 \\ b_1 & b_2 & b_3 & b_4 \\ c_1 & c_2 & c_3 & c_4 \\ d_1 & d_2 & d_3 & d_4 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & p & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} a_1 & pa_2 & a_3 & a_4 \\ b_1 & pb_2 & b_3 & b_4 \\ c_1 & pc_2 & c_3 & c_4 \\ d_1 & pd_2 & d_3 & d_4 \end{bmatrix}$$

en $\det \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & p & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} = p$.

4) Als de kolommen (rijen) afhankelijk zijn, is een kolom (rij) lineaire combinatie van de overige. Herhaalde toepassing van 2) geeft een determinant waarvan een kolom of rij $\underline{0}$ resp. $\underline{0}^T$ is. Uit 3) volgt dat de determinant 0 is.

Als de kolommen (rijen) van A onafhankelijk zijn, kan A door bewerkingen als in 1) en 2) overgevoerd worden in een diagonaalmatrix met diagonaalelementen $\neq 0$. Van zo'n matrix is de determinant gelijk aan het produkt van de diagonaalelementen.

5) Herhaalde toepassing van 3).

Het berekenen van de determinanten gaat nu het handigst door gebruikmakend van de regels 1), 2) en 3) en transponeren, zoveel mogelijk nullen op de eerste rij te krijgen.

Voorbeeld 1.

$$\begin{array}{cccc|cccc}
 \underline{a} & \underline{b} & \underline{c} & \underline{d} & \underline{a} & \underline{b} - 2\underline{a} & \underline{c} - 3\underline{a} & \underline{d} - 4\underline{a} \\
 \left| \begin{array}{cccc} 1 & 2 & 3 & 4 \\ 1 & 2 & 3 & 5 \\ 1 & 2 & 4 & 5 \\ 1 & 3 & 4 & 5 \end{array} \right| & = & \left| \begin{array}{cccc} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{array} \right| & = & & & & \\
 & & & & \underline{a} & \underline{p} & \underline{q} & \underline{r} \\
 & & & & \underline{a} & \underline{p} & \underline{q} - \underline{p} & \underline{r} - \underline{p} \\
 = \left| \begin{array}{cccc} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \end{array} \right| & = & 1 \cdot 1 \cdot \left| \begin{array}{cc} 0 & 1 \\ 1 & 0 \end{array} \right| & = & -1 .
 \end{array}$$

Stelling. Als A een $(n \times n)$ -matrix is, dan heeft $A\underline{x} = \underline{0}$ dan en slechts dan een van $\underline{0}$ verschillende oplossing als $\det A = 0$.

Het belang van determinanten is gelegen in de volgende stelling.

Stelling. Een getal λ is een eigenwaarde van A dan en slechts dan als

$$\det(A - \lambda I) = 0 .$$

Deze algebraïsche vergelijking in λ (van de n -de graad) heet de karakteristische vergelijking van A .

Voorbeelden. 2) Voor de afschuiving $A_{\underline{a}}$ uit 7.7.2, voorbeeld 6 geldt:

$$\det(A_{\underline{a}} - \lambda I) = \begin{vmatrix} 1 - \lambda & a \\ 0 & 1 - \lambda \end{vmatrix} = (1 - \lambda)^2 .$$

3) Voor de spiegeling $S_{\underline{a}}$ uit 7.7.2, voorbeeld 5 geldt:

$$S_{\underline{a}} = \begin{bmatrix} 1 - 2a_1^2 & -2a_1a_2 & -2a_1a_3 \\ -2a_1a_2 & 1 - 2a_2^2 & -2a_2a_3 \\ -2a_1a_3 & -2a_2a_3 & 1 - 2a_3^2 \end{bmatrix}$$

en

$$\begin{aligned} \det(S_{\underline{a}} - \lambda I) &= \begin{vmatrix} 1 - 2a_1^2 - \lambda & -2a_1a_2 & -2a_1a_3 \\ -2a_1a_2 & 1 - 2a_2^2 - \lambda & -2a_2a_3 \\ -2a_1a_3 & -2a_2a_3 & 1 - 2a_3^2 - \lambda \end{vmatrix} = \\ &= (1 - 2a_1^2 - \lambda)(1 - 2a_2^2 - \lambda)(1 - 2a_3^2 - \lambda) - 8a_1^2a_2^2a_3^2 - 8a_1^2a_2^2a_3^2 + \\ &- 4a_1^2a_3^2(1 - 2a_2^2 - \lambda) - 4a_2^2a_3^2(1 - 2a_1^2 - \lambda) - 4a_1^2a_2^2(1 - 2a_3^2 - \lambda) = \\ &= -\lambda^3 + \lambda^2(3 - 2a_1^2 - 2a_2^2 - 2a_3^2) - \lambda(3 - 4a_1^2 - 4a_2^2 - 4a_3^2) + \\ &+ 1 - 2(a_1^2 + a_2^2 + a_3^2) . \end{aligned}$$

In verband met $a_1^2 + a_2^2 + a_3^2 = |\underline{a}|^2 = 1$, is

$$\det(S_{\underline{a}} - \lambda I) = -\lambda^3 + \lambda^2 + \lambda - 1 = -(\lambda - 1)^2(\lambda + 1) .$$

De eigenwaarden zijn dus 1 en -1.

4) Het werken met determinanten wordt nog geïllustreerd aan de volgende berekening. Uit $S_{\underline{a}} = I - 2Q_{\underline{a}}$ (7.7.2, voorbeeld 4) volgt:

$$\det(S_{\underline{a}} - \lambda I) = \det((1 - \lambda)I - 2Q_{\underline{a}}) = -8 \det(Q_{\underline{a}} - \frac{1}{2}(1 - \lambda)I) .$$

Dus λ is een eigenwaarde van $S_{\underline{a}}$ dan en slechts dan als $\frac{1}{2}(1 - \lambda)$ eigenwaarde is van $Q_{\underline{a}}$. Dit is in overeenstemming met het feit dat 0 en 1 de eigenwaarden zijn van $Q_{\underline{a}}$ en 1 en -1 die van $S_{\underline{a}}$.

De beide volgende toepassingen zijn voornamelijk van theoretisch belang.

Stelling (regel van Cramer). Als voor het stelsel vergelijkingen $A^{(n \times n)} \underline{x} = \underline{b}$ met $A = [\underline{a}_1, \underline{a}_2, \dots, \underline{a}_n]$ geldt $\text{rang } A = n$, dan is het stelsel eënduidig oplosbaar en de oplossing wordt gegeven door

$$x_i = \frac{\det[\underline{a}_1, \dots, \underline{a}_{i-1}, \underline{b}, \underline{a}_{i+1}, \dots, \underline{a}_n]}{\det A}, \quad i = 1, \dots, n .$$

Bewijs. Uit $\text{rang } A = n$ volgt dat A inverteerbaar is (zie 7.8.2), zodat het stelsel $A\underline{x} = \underline{b}$ precies één oplossing heeft. Voor de oplossing $[x_1, \dots, x_n]^T$ geldt

$$x_1 \underline{a}_1 + \dots + x_n \underline{a}_n = \underline{b} .$$

Dan is

$$\begin{aligned} \det[\underline{a}_1, \dots, \underline{a}_{i-1}, \underline{b}, \underline{a}_{i+1}, \dots, \underline{a}_n] &= \\ &= \sum_{k=1}^n x_k \det[\underline{a}_1, \dots, \underline{a}_{i-1}, \underline{a}_k, \underline{a}_{i+1}, \dots, \underline{a}_n] = \\ &= x_i \det[\underline{a}_1, \dots, \underline{a}_{i-1}, \underline{a}_i, \underline{a}_{i+1}, \dots, \underline{a}_n] = \\ &= x_i \det A, \quad \text{met } \det A \neq 0 . \end{aligned} \quad \square$$

Gevolg. Zij A een $(n \times n)$ -matrix en zij $\det A \neq 0$, zodat $\text{rang } A = n$. De vraag naar de inverse van A , oftewel de vraag naar een matrix

$$\underline{X} = [x_1, x_2, \dots, x_n]$$

zodat $A\underline{X} = I$, kunnen we dan ook aldus beantwoorden.

Uit $AX = I$ volgt:

$$[Ax_1, \dots, Ax_n] = [e_1, \dots, e_n],$$

zodat

$$\forall_{j=1, \dots, n} Ax_j = e_j.$$

Met

$$A = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{bmatrix} \quad \text{en} \quad X = \begin{bmatrix} x_{11} & \dots & x_{1n} \\ \vdots & & \vdots \\ x_{n1} & \dots & x_{nn} \end{bmatrix}$$

geldt volgens de regel van Cramer:

$$x_{ij} = \frac{\det[a_1, \dots, a_{i-1}, e_j, a_{i+1}, \dots, a_n]}{\det A}.$$

(a_1, \dots, a_n stellen weer de kolommen van A voor), zodat

$$x_{ij} = \frac{(-1)^{i+j} D_{ji}}{\det A},$$

waarbij D_{ji} de determinant is van de matrix die uit A ontstaat door weglating van de j-de rij en de i-de kolom. Daarmee is de inverse van A bepaald.

.8.5. Orthogonale afbeeldingen

Definitie. Een lineaire afbeelding $A : \mathbb{R}^n \rightarrow \mathbb{R}^n$ (een $(n \times n)$ -matrix A) heet orthogonaal indien voor alle $\underline{x} \in \mathbb{R}^n$ geldt

$$|Ax| = |\underline{x}|.$$

Orthogonale afbeeldingen zijn afbeeldingen die "lengtebehoudend" zijn. Spiegelingen en draaiingen zijn voorbeelden van orthogonale afbeeldingen. We zien onmiddellijk dat een orthogonale afbeelding geen andere eigenwaarden kan hebben dan +1 en -1.

Stelling (karakterisering van orthogonaliteit). Zij $A : \mathbb{R}^n \rightarrow \mathbb{R}^n$ lineair; de volgende beweringen zijn gelijkwaardig:

- a) A is orthogonaal;
- b) voor alle $\underline{x}, \underline{y} \in \mathbb{R}^n$ geldt $(Ax, Ay) = (\underline{x}, \underline{y})$;

- c) A is inverteerbaar en $A^{-1} = A^T$;
- d) $AA^T = I$;
- e) $A^T A = I$;
- f) de rijen van A vormen een orthonormaal stelsel;
- g) de kolommen van A vormen een orthonormaal stelsel.

Bewijs. Uit a) volgt b) want $(\underline{x}, \underline{x}) + 2(\underline{x}, \underline{y}) + (\underline{y}, \underline{y}) = (\underline{x} + \underline{y}, \underline{x} + \underline{y}) =$
 $= (A(\underline{x} + \underline{y}), A(\underline{x} + \underline{y})) = (A\underline{x}, A\underline{x}) + 2(A\underline{x}, A\underline{y}) + (A\underline{y}, A\underline{y}) =$
 $= (\underline{x}, \underline{x}) + 2(A\underline{x}, A\underline{y}) + (\underline{y}, \underline{y})$.

Uit b) volgt c) want $(\underline{x}, \underline{y}) = (A\underline{x}, A\underline{y}) = (\underline{x}, A^T A \underline{y})$ (zie 7.2.7); dus voor alle \underline{x} en \underline{y} geldt $(\underline{x}, (I - A^T A)\underline{y}) = 0$, derhalve is $I - A^T A = \mathcal{O}$.

Alle overige onderdelen van het bewijs zijn triviaal. □

Stelling. Is A orthogonaal, dan is $|\det A| = 1$.

Bewijs. $A^T A = I$, $(\det A)^2 = 1$. □

Voorbeeld 1. Iedere orthogonale afbeelding van \mathbb{R}^2 in \mathbb{R}^2 is of een draaiing om $\underline{0}$ of een spiegeling aan een rechte.

Bewijs. Zij

$$A = \begin{bmatrix} a & c \\ b & d \end{bmatrix}$$

een orthogonale (2×2) -matrix, dan is $a^2 + b^2 = 1$, $ac + bd = 0$, $c^2 + d^2 = 1$ en $\Delta := \det A = ad - bc = \pm 1$. Uit $a^2 + b^2 = 1$ volgt dat er precies één $\alpha \in [0, 2\pi)$ voldoet aan

$$\cos \alpha = a, \quad \sin \alpha = b.$$

Uit

$$\begin{cases} c \cos \alpha + d \sin \alpha = 0 \\ -c \sin \alpha + d \cos \alpha = \Delta \end{cases}$$

volgt

$$c = -\Delta \sin \alpha, \quad d = \Delta \cos \alpha,$$

zodat

$$A = \begin{bmatrix} \cos \alpha & -\Delta \sin \alpha \\ \sin \alpha & \Delta \cos \alpha \end{bmatrix}.$$

Als $\Delta = 1$ dan is $A = D_\alpha$, d.i. de draaiing over een hoek α om $\underline{0}$; zie 7.7.2, voorbeeld 7. Als $\Delta = -1$ dan is

$$A = \begin{bmatrix} \cos \alpha & \sin \alpha \\ \sin \alpha & -\cos \alpha \end{bmatrix}$$

en A is de spiegeling aan de rechte $\underline{x} = \lambda \underline{a}$ met $\underline{a} = [\cos \frac{1}{2}\alpha, \sin \frac{1}{2}\alpha]^T$ (ga na).

Voorbeeld 2. In \mathbb{R}^3 zijn er meer typen orthogonale afbeeldingen. We sommen nu de mogelijkheden op. Omdat de karakteristieke vergelijking van de derde graad is, is er steeds tenminste één eigenwaarde.

- 1) $\lambda = 1$ is een eigenwaarde.
 - a) De eigenruimte bij $\lambda = 1$ is driedimensionaal; $A = I$.
 - b) De eigenruimte bij $\lambda = 1$ (we noemen deze E_1) is tweedimensionaal. Uit de orthogonaliteit volgt dan dat E_1^\perp eigenruimte bij $\lambda = -1$ is; A is een spiegeling aan het vlak E_1 .
 - c) De eigenruimte, E_1 , bij $\lambda = 1$ is één dimensionaal. Dan geldt dat $AE_1^\perp = E_1^\perp$; de restrictie van A tot E_1^\perp is een orthogonale afbeelding van een tweedimensionale ruimte die geen eigenwaarde 1 heeft. A is een draaiing om de rechte E_1 .
- 2) $\lambda = -1$ is de enige eigenwaarde.
 - a) De eigenruimte, E_{-1} , bij $\lambda = -1$ is driedimensionaal; $A = -I$, spiegeling t.o.v. het punt $\underline{0}$.
 - b) De eigenruimte, E_{-1} , bij $\lambda = -1$ is één dimensionaal ($\dim E_{-1} = 2$ is uitgesloten; vergelijk met 1b). A is een draaiing om de rechte E_{-1} gecombineerd met een spiegeling aan het vlak E_{-1}^\perp .

Voorbeeld.

$$\begin{cases} x_2 + 2x_3 + x_5 + x_7 = 0 \\ x_4 + x_6 + 2x_7 = 0 \\ x_6 + x_7 = 0 \\ x_2 + 2x_3 + x_4 + x_5 + 2x_7 = 0 \end{cases}$$

Met behulp van de veegalgoritme (7.6.4) is de matrix van dit stelsel te herleiden tot de normaalvorm:

$$\tilde{A} = \begin{bmatrix} 0 & 1 & 2 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

In dit geval is $k_1 = 2$, $k_2 = 4$, $k_3 = 6$, $\text{rang } \tilde{A} = 3$, dus $\dim N(\tilde{A}) = 7 - 3 = 4$.

De oplossing van $A\underline{x} = \underline{0}$ is nu: $x_1 = \lambda_1$, $x_3 = \lambda_2$, $x_5 = \lambda_3$, $x_7 = \lambda_4$, derhalve $x_2 = -2\lambda_2 - \lambda_3 - \lambda_4$, $x_4 = -\lambda_4$, $x_6 = -\lambda_4$.

Dit kunnen we ook schrijven als:

$$\underline{x} = \lambda_1[1,0,0,0,0,0,0]^T + \lambda_2[0,-2,1,0,0,0,0]^T + \\ + \lambda_3[0,-1,0,0,1,0,0]^T + \lambda_4[0,-1,0,-1,0,-1,1]^T.$$

Uit het feit dat $\dim N(A) = 4$ volgt ook dat we de algemene oplossing hebben.

We formuleren het nog eens anders:

Laat A een $(m \times n)$ -matrix zijn. Een basis voor de oplossingsruimte van $A\underline{x} = \underline{0}$ verkrijgt men aldus:

- 1) Breng A op normaalvorm. Laat k_1, \dots, k_r de rangnummers van de kolommen zijn waarin de vectoren $\underline{e}_1, \dots, \underline{e}_r$ uit \mathbb{R}^m staan.
- 2) Splits de onbekenden in twee groepen: x_{k_1}, \dots, x_{k_r} , en de overige, zeg: $x_{\ell_1}, \dots, x_{\ell_{n-r}}$.
- 3) Stel, voor elke $i \in \{1, 2, \dots, n-r\}$, $x_{\ell_i} = 1$, $x_{\ell_j} = 0$ ($j \neq i$) en bereken x_{k_1}, \dots, x_{k_r} . Zo verkrijgt men telkens een oplossing.
- 4) De aldus verkregen $n-r$ oplossingen vormen een basis van $N(A)$.

7.9.3. Inhomogene lineaire vergelijkingen

We beschouwen het stelsel vergelijkingen

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n = b_2 \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \dots + a_{mn}x_n = b_m \end{cases} \quad (2)$$

Met $A, \underline{a}_1, \dots, \underline{a}_m$ als in 7.9.2, $\underline{b} = [b_1, \dots, b_m]^T$ en $\underline{x} = [x_1, \dots, x_n]^T$ is het stelsel (2) te schrijven als $A\underline{x} = \underline{b}$. Men noemt A en $[A, \underline{b}]$ resp. de coëfficiëntenmatrix en de aangevulde matrix van (2).

Door herhaling van reeds gegeven argumenten kan men alle onderstaande beweringen bewijzen.

- a) Het stelsel (2) is dan en slechts dan niet strijdig (d.w.z. oplosbaar) indien $\underline{b} \in A\mathbb{R}^n$.
- b) Het stelsel (2) is dan en slechts dan oplosbaar indien $\text{rang } A = \text{rang}[A, \underline{b}]$.
- c) Als $A\underline{a} = \underline{b}$ dan is het stelsel (2) oplosbaar en de algemene oplossing is $\underline{a} + N(A)$ (m.a.w. $\underline{a} + N(A)$ is de verzameling van alle oplossingen).
- d) Als (2) oplosbaar is, is de dimensie van de nevenruimte der oplossingen (gelijk aan $\dim N(A)$) gelijk aan $n - \text{rang } A$.
- e) Als $[\tilde{A}, \tilde{\underline{b}}]$ de normaalvorm is van $[A, \underline{b}]$, dan hebben $\tilde{A}\underline{x} = \tilde{\underline{b}}$ en $A\underline{x} = \underline{b}$ dezelfde oplossingen.
- f) Als $n = m$ en A is een injectie (d.w.z. $N(A) = \{0\}$), dan is $A\underline{x} = \underline{b}$ éénduidig oplosbaar met oplossing $\underline{x} = A^{-1}\underline{b}$.

Uit bewering e) volgt tevens een methode om het stelsel (2) op te lossen.

- 1) Breng $[A, \underline{b}]$ op normaalvorm. Laat k_1, \dots, k_r de rangnummers van de kolommen zijn waarin de vectoren $\underline{e}_1, \dots, \underline{e}_r$ uit \mathbb{R}^m staan. Er zijn nu twee mogelijkheden:
 - i) $k_r = n+1$, dan is $\text{rang}[A, \underline{b}] > \text{rang } A$; het stelsel is strijdig.
 - ii) $k_r \leq n$, dan is $\text{rang}[A, \underline{b}] = \text{rang } A$; het stelsel is oplosbaar.
- 2) Laat het stelsel oplosbaar zijn en laat $[\tilde{A}, \tilde{\underline{b}}]$ de normaalvorm zijn van de aangevulde matrix, dan is $\tilde{\underline{b}} = [\tilde{b}_1, \dots, \tilde{b}_r, 0, \dots, 0]^T$.
Is nu $r = n$ dan is het stelsel eenduidig oplosbaar; is $r < n$ dan is er een $(n-r)$ -dimensionale nevenruimte van oplossingen. Op grond van c) is het voldoende één oplossing aan te geven en verder $N(A)$ volgens 7.9.2 te

bepalen (dit is, als \tilde{A} bekend is, niet veel werk meer). Eén oplossing is $x_{k_i} = \tilde{b}_i$ voor $i = 1, \dots, r$, en de overige x -en gelijk aan 0.

Voorbeelden.

$$1) \quad \begin{cases} x_2 + x_3 = 10 \\ x_1 + x_2 + x_3 = 8 \\ x_1 + 2x_3 = 6, \end{cases}$$

$$[A, \underline{b}] = \begin{bmatrix} 0 & 1 & 1 & 10 \\ 1 & 1 & 1 & 8 \\ 1 & 0 & 2 & 6 \end{bmatrix},$$

$$[\tilde{A}, \tilde{\underline{b}}] = \begin{bmatrix} 1 & 0 & 0 & -2 \\ 0 & 1 & 0 & 6 \\ 0 & 0 & 1 & 4 \end{bmatrix}.$$

Oplossing: $x_1 = -2$, $x_2 = 6$, $x_3 = 4$ of $\underline{x} = [-2, 6, 4]^T$.

$$2) \quad \begin{cases} x_1 + 2x_2 + 3x_3 - x_4 = 1 \\ 2x_1 + 3x_2 - 2x_3 + 3x_4 = 1 \\ 4x_1 + 7x_2 + 4x_3 + x_4 = 2, \end{cases}$$

$$[A, \underline{b}] = \begin{bmatrix} 1 & 2 & 3 & -1 & 1 \\ 2 & 3 & -2 & 3 & 1 \\ 4 & 7 & 4 & 1 & 2 \end{bmatrix},$$

$$[\tilde{A}, \tilde{\underline{b}}] = \begin{bmatrix} 1 & 0 & -13 & 9 & 0 \\ 0 & 1 & 8 & -5 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

In dit voorbeeld is $\text{rang}[\tilde{A}, \tilde{\underline{b}}] = 3 > 2 = \text{rang } \tilde{A}$. Het stelsel is strijdig, er zijn geen oplossingen.

$$3) \quad \begin{cases} x_1 + 2x_2 + 3x_3 - x_4 = 1 \\ 2x_1 + 3x_2 - 2x_3 + 3x_4 = 1 \\ 4x_1 + 7x_2 + 4x_3 + x_4 = 3, \end{cases}$$

$$[A, \underline{b}] = \begin{bmatrix} 1 & 2 & 3 & -1 & 1 \\ 2 & 3 & -2 & 3 & 1 \\ 4 & 7 & 4 & 1 & 3 \end{bmatrix},$$

$$[\tilde{A}, \tilde{\underline{b}}] = \begin{bmatrix} 1 & 0 & -13 & 9 & -1 \\ 0 & 1 & 8 & -5 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}.$$

In dit voorbeeld is $\text{rang}[\tilde{A}, \tilde{b}] = \text{rang } \tilde{A} = 2$. De dimensie van de nevenruimte der oplossingen is $4 - 2 = 2$.

Eén oplossing is $x_1 = -1, x_2 = 1, x_3 = x_4 = 0$.

De algemene oplossing luidt

$$\underline{x} = [-1, 1, 0, 0]^T + \lambda [13, -8, 1, 0]^T + \mu [-9, 5, 0, 1]^T .$$

ma 24.2.78

7.9.4. Positieve oplossingen van stelsels lineaire vergelijkingen

In veel van de toepassingen van stelsels $A\underline{x} = \underline{b}$ gaat het er niet om de algemene oplossing te vinden, maar is men slechts geïnteresseerd in positieve oplossingen, d.w.z. oplossingen die voldoen aan $\underline{x} \geq \underline{0}$ (dus $x_i \geq 0$ voor $i = 1, \dots, n$).

N.B. Het gebruik van het woord positief wijkt hier af van het gewone gebruik; een vector \underline{x} heet positief als alle elementen van \underline{x} niet negatief zijn.

Als voorbereiding op het bestuderen van positieve oplossingen dienen de volgende beschouwingen over zgn. convexe verzamelingen.

7.9.4.1. Definitie. Zij $\underline{y}, \underline{z} \in \mathbb{R}^n$; een vector \underline{x} heet een convexe combinatie van \underline{y} en \underline{z} indien er een getal $\lambda \in [0, 1]$ bestaat zó dat

$$\underline{x} = \lambda \underline{y} + (1 - \lambda) \underline{z} .$$

De verzameling van alle convexe combinaties van \underline{y} en \underline{z} noemen we het segment tussen \underline{y} en \underline{z} .

7.9.4.2. Definitie. Een deelverzameling U van \mathbb{R}^n heet convex indien voor elke $\underline{y}, \underline{z} \in U$ en elke $\lambda \in [0, 1]$ geldt:

$$\lambda \underline{y} + (1 - \lambda) \underline{z} \in U .$$

Voorbeelden van convexe verzamelingen

- 1) In \mathbb{R}^2 is de cirkelschijf $x_1^2 + x_2^2 \leq 1$ convex.
- 2) Het binnengebied van een driehoek in \mathbb{R}^2 en van een viervlak in \mathbb{R}^3 is convex.
- 3) In \mathbb{R}^n is $\{\underline{x} \mid \underline{x} \geq \underline{0}\}$ convex.
- 4) Zij $\underline{a} \in \mathbb{R}^n, r \in \mathbb{R}$; de verzamelingen $\{\underline{x} \mid (\underline{a}, \underline{x}) > r\}$ en $\{\underline{x} \mid (\underline{a}, \underline{x}) \geq r\}$ zijn convexe verzamelingen (geheten open en gesloten halfruimte).

Opmerking. Het verband tussen de convexe functies uit 2.3.3, 2.6.9 en convexe verzamelingen wordt gegeven door de volgende eigenschap:

Zij I een interval in \mathbb{R} , dan is $f: I \rightarrow \mathbb{R}$ dan en slechts dan een convexe functie als $\{(x,y) \in \mathbb{R}^2 \mid x \in I, y \geq f(x)\}$ een convexe verzameling is.

7.9.4.3. Definitie. Zij $\underline{x}_1, \underline{x}_2, \dots, \underline{x}_m \in \mathbb{R}^n$; een vector \underline{z} heet een convexe combinatie van $\underline{x}_1, \underline{x}_2, \dots, \underline{x}_m$ indien er getallen $\lambda_1, \lambda_2, \dots, \lambda_m$ bestaan met $\lambda_1 \geq 0, \lambda_2 \geq 0, \dots, \lambda_m \geq 0$ en $\lambda_1 + \lambda_2 + \dots + \lambda_m = 1$, zó dat $\underline{z} = \lambda_1 \underline{x}_1 + \lambda_2 \underline{x}_2 + \dots + \lambda_m \underline{x}_m$.

Eigenschappen. 1) Als U_1 en U_2 convex zijn, dan is $U_1 \cap U_2$ convex.

De doorsnede van een willekeurige collectie convexe verzamelingen is convex.

2) Iedere nevenruimte (en in het bijzonder iedere deelruimte) is convex.

3) Als U convex is en als $\underline{x}_1, \dots, \underline{x}_k \in U$, dan ligt ook elke convexe combinatie van $\underline{x}_1, \dots, \underline{x}_k$ in U .

Bewijs. We bewijzen alleen eigenschap 3) door volledige inductie naar k .

Voor $k = 2$ volgt de eigenschap uit de definitie van convexe verzameling.

Onderstel nu dat iedere convexe combinatie van een $(k-1)$ -tal elementen uit U , in U ligt. Zij $\underline{x} = \lambda_1 \underline{x}_1 + \dots + \lambda_k \underline{x}_k$, $\lambda_i \geq 0$, $\sum \lambda_i = 1$. Als $\lambda_1 = 1$ zijn de overige $\lambda_2, \dots, \lambda_k$ gelijk 0 en is $\underline{x} \in U$; anders is

$$\underline{x} = \lambda_1 \underline{x}_1 + (1 - \lambda_1) \left[\frac{\lambda_2}{1 - \lambda_1} \underline{x}_2 + \dots + \frac{\lambda_k}{1 - \lambda_1} \underline{x}_k \right] =: \lambda_1 \underline{x}_1 + (1 - \lambda_1) \underline{y}.$$

Daar $\underline{y} \in U$ krachtens aanname, is $\underline{x} \in U$. □

4) Bij iedere verzameling V is er een kleinste convexe verzameling die V bevat; deze is de doorsnede van alle convexe verzamelingen die V bevatten.

Notatie: $C(V)$; naam: convexe omhulsel van V .

5) Als $V = \{\underline{x}_1, \dots, \underline{x}_k\}$, dan is

$$C(V) = \{ \lambda_1 \underline{x}_1 + \dots + \lambda_k \underline{x}_k \mid \lambda_1 \geq 0, \dots, \lambda_k \geq 0, \lambda_1 + \dots + \lambda_k = 1 \}.$$

Bewijs. Merk op dat alle convexe combinaties van eindig veel vectoren een convexe verzameling vormen. □

Let op de analogie tussen $C(V)$ en $\langle V \rangle$. In plaats van $C(\{\underline{a}_1, \dots, \underline{a}_k\})$ schrijven we $C(\underline{a}_1, \dots, \underline{a}_k)$. De verzameling $C(\underline{a}_1, \dots, \underline{a}_k)$ heet het convexe polytoop opgespannen door $\underline{a}_1, \dots, \underline{a}_k$.

- 6) Een convex polytoop is begrensd. (Een verzameling $V \subset \mathbb{R}^n$ heet begrensd als er een $M > 0$ bestaat zodat $|\underline{x}| \leq M$ voor alle $\underline{x} \in V$; N.B. \emptyset is begrensd.)
- 7) Als $\underline{x} \in C(\underline{a}_1, \dots, \underline{a}_k)$ dan is $C(\underline{x}, \underline{a}_1, \dots, \underline{a}_k) = C(\underline{a}_1, \dots, \underline{a}_k)$.
- 8) Er bestaat een e nduidig bepaalde verzameling $\{\underline{b}_1, \dots, \underline{b}_\ell\} \subset \{\underline{a}_1, \dots, \underline{a}_k\}$ z  dat:
 - i) $C(\underline{b}_1, \dots, \underline{b}_\ell) = C(\underline{a}_1, \dots, \underline{a}_k)$.
 - ii) Geen der \underline{b}_i is gelegen in het convexe polytoop opgespannen door de overige \underline{b} 's.

Bewijs. Probeer $\underline{a}_1, \dots, \underline{a}_k$ achtereenvolgens weg te laten. □

De \underline{b}_i heten de hoekpunten van het polytoop.

- 9) $\{\underline{x} \mid \underline{x} \geq 0, x_1 + \dots + x_n = 1\}$ is een convex polytoop met hoekpunten $\underline{e}_1, \dots, \underline{e}_n$.
- 10) Een convexe combinatie van convexe combinaties van $\underline{x}_1, \dots, \underline{x}_k$ is zelf een convexe combinatie van $\underline{x}_1, \dots, \underline{x}_k$.

7.9.4.4. Definitie. Zij U een convexe verzameling. Een vector $\underline{x} \in U$ heet een extreem element van U als uit $\underline{x} = \lambda \underline{y} + (1 - \lambda) \underline{z}$ met $\underline{y}, \underline{z} \in U, \underline{y} \neq \underline{z}$, volgt $\lambda = 1$ en $\underline{x} = \underline{y}$ of $\lambda = 0$ en $\underline{x} = \underline{z}$.

Gelijkwaardige formulering: \underline{x} is extreem element als uit $\underline{x} = \alpha \underline{y} + (1 - \alpha) \underline{z}$ met $\underline{y}, \underline{z} \in U, 0 < \alpha < 1$, volgt: $\underline{x} = \underline{y} = \underline{z}$; of \underline{x} is extreem element als \underline{x} niet tussenpunt is van een segment in U ; of \underline{x} is extreem element als \underline{x} geen convexe combinatie is van twee van \underline{x} verschillende elementen van U ; of \underline{x} is een extreem element van U als $U \setminus \{\underline{x}\}$ convex is.

Voorbeelden. 5) Open en gesloten halfruimten hebben geen extreme elementen.

6) De hoekpunten van een convex polytoop zijn de extreme elementen.

7.9.4.5. Stelling. De positieve oplossingen van $A\underline{x} = \underline{b}$ vormen een convexe verzameling.

Bewijs. De verzameling van positieve oplossingen kan leeg zijn. De lege verzameling is convex. Zijn \underline{y} en \underline{z} positieve oplossingen en is $0 \leq \lambda \leq 1$, dan is ook $\lambda\underline{y} + (1 - \lambda)\underline{z}$ een positieve oplossing. \square

Voorbeeld. 7) De verzameling, V , van de positieve oplossingen van $\frac{x}{a} + \frac{y}{b} = 1$ in \mathbb{R}^2 kan de volgende gedaanten hebben:

- i) Als $a > 0$, $b > 0$ dan is V begrensd met twee extreme elementen.
- ii) Als $a > 0$, $b < 0$ of $a < 0$, $b > 0$ dan is V niet begrensd en V heeft slechts één extreem element.
- iii) Als $a < 0$, $b < 0$ dan is $V = \emptyset$.

Laat A een $(m \times n)$ -matrix zijn, en laat $\underline{a}_1, \dots, \underline{a}_n$ de kolomvectoren van A zijn zodat het stelsel vergelijkingen $A\underline{x} = \underline{b}$ te schrijven is als

$$x_1\underline{a}_1 + x_2\underline{a}_2 + \dots + x_n\underline{a}_n = \underline{b}.$$

De extreme elementen van de verzameling der positieve oplossingen heten extreme oplossingen.

7.9.4.6. Stelling. Zij $\underline{x} = [x_1, x_2, \dots, x_p, 0, 0, \dots, 0]^T$ met $x_1 > 0, x_2 > 0, \dots, x_p > 0$ een positieve oplossing van $A\underline{x} = \underline{b}$. Nu is \underline{x} een extreme oplossing dan en slechts dan als $\underline{a}_1, \dots, \underline{a}_p$ een onafhankelijk stelsel is.

Bewijs. i) Veronderstel dat $\underline{a}_1, \dots, \underline{a}_p$ onafhankelijk is en dat

$\underline{x} = \lambda_0\underline{u} + (1 - \lambda_0)\underline{v}$ voor zekere λ_0 met $0 < \lambda_0 < 1$ en positieve oplossingen \underline{u} en \underline{v} .

Uit $\lambda_0 u_k + (1 - \lambda_0)v_k = 0$, $u_k \geq 0$, $v_k \geq 0$, volgt $u_k = 0$, $v_k = 0$.

We hebben dus: $u_{p+1} = \dots = u_n = 0$, $v_{p+1} = \dots = v_n = 0$. Nu is

$$\begin{aligned} x_1\underline{a}_1 + \dots + x_p\underline{a}_p &= \underline{b} \\ u_1\underline{a}_1 + \dots + u_p\underline{a}_p &= \underline{b} \\ v_1\underline{a}_1 + \dots + v_p\underline{a}_p &= \underline{b}. \end{aligned}$$

Omdat $\underline{a}_1, \dots, \underline{a}_p$ onafhankelijk is, volgt hieruit $\underline{x} = \underline{u} = \underline{v}$.

ii) Veronderstel dat $\underline{a}_1, \dots, \underline{a}_p$ niet onafhankelijk is, dan zijn er getallen y_1, \dots, y_p , niet alle gelijk aan nul, zó dat $y_1 \underline{a}_1 + \dots + y_p \underline{a}_p = \underline{0}$.

Beschouw $\underline{y} := [y_1, y_2, \dots, y_p, 0, 0, \dots, 0]^T \in \mathbb{R}^n$, dan is $\underline{y} \neq \underline{0}$ (N.B. \underline{y} kan negatieve elementen hebben!).

Zij $m := \min\{x_1, \dots, x_p\}$, $M := \max\{|y_1|, \dots, |y_p|\}$, dan is $m > 0$, $M > 0$.

Zij $\epsilon = \frac{1}{2} \frac{m}{M}$. Nu zijn $\underline{u} := \underline{x} + \epsilon \underline{y}$ en $\underline{v} := \underline{x} - \epsilon \underline{y}$ twee verschillende positieve oplossingen en $\underline{x} = \frac{1}{2} \underline{u} + \frac{1}{2} \underline{v}$, dus \underline{x} is geen extreme oplossing. \square

7.9.4.7. Stelling. Zij $\underline{x} = [x_1, \dots, x_n]^T$ een positieve oplossing van $A\underline{x} = \underline{b}$. Nu is \underline{x} een extreme oplossing dan en slechts dan indien het stelsel van de kolomvectoren van A met rangnummers i, waarvoor $x_i \neq 0$ is, een onafhankelijk stelsel is.

Bewijs. Pas 7.9.4.6 toe, eventueel na henummering van de onbekenden. \square

Stelling. $A\underline{x} = \underline{b}$ heeft ten hoogste eindig veel extreme oplossingen.

Bewijs. De verzameling $\{\underline{a}_1, \dots, \underline{a}_n\}$ der kolomvectoren van A heeft 2^n deelverzamelingen. Alleen die deelverzamelingen waarvan de elementen een onafhankelijk stelsel vormen, corresponderen elk met ten hoogste één extreme oplossing. \square

Stelling 7.9.4.7 geeft ons tevens een mogelijkheid de extreme oplossingen te bepalen. We gaan daarbij als volgt te werk:

Bepaal een onafhankelijk stelsel kolomvectoren van A, bijv. $\underline{a}_1, \underline{a}_3, \underline{a}_7, \underline{a}_8$.

Probeer op te lossen: $x_1 \underline{a}_1 + x_3 \underline{a}_3 + x_7 \underline{a}_7 + x_8 \underline{a}_8 = \underline{b}$; er is ten hoogste één oplossing. Laat deze zijn $x_1 = \alpha_1$, $x_3 = \alpha_3$, $x_7 = \alpha_7$, $x_8 = \alpha_8$. Beschouw nu $[\alpha_1, 0, \alpha_3, 0, 0, 0, \alpha_7, \alpha_8, 0, \dots, 0]^T$; als deze vector positief is, is het een extreme oplossing.

Belangrijk zijn gevallen waarin de verzameling van de positieve oplossingen van $A\underline{x} = \underline{b}$ begrensd is. Een voldoende (maar niet nodige!) voorwaarde hiervoor is:

$$\sum_{i=1}^m a_{ij} > 0 \quad (j = 1, \dots, n) \quad \text{en} \quad \sum_{i=1}^m b_i \geq 0 .$$

Algemener: de verzameling van de positieve oplossingen van $A\underline{x} = \underline{b}$ is begrensd als er reële getallen r_1, \dots, r_m bestaat zodat

$$\sum_{i=1}^m r_i a_{ij} > 0 \quad (j = 1, \dots, n) \quad \text{en} \quad \sum_{i=1}^m r_i b_i \geq 0 .$$

Als de verzameling der positieve oplossingen begrensd is gelden fraaie stellingen.

7.9.4.8. Stelling. Laat de verzameling der positieve oplossingen van $A\underline{x} = \underline{b}$ begrensd zijn, laat \underline{x}_0 een positieve oplossing zijn en laat $\underline{a} \in N(A)$ zijn. Dan vormen alle positieve oplossingen van de vorm $\underline{x}_0 + t\underline{a}$ met $t \in \mathbb{R}$ een segment in \mathbb{R}^n , eventueel bestaande uit het punt \underline{x}_0 alleen.

Bewijs. Triviaal. □

We komen nu tot de hoofdstelling van deze paragraaf.

7.9.4.9. Stelling. Laat de verzameling V van de positieve oplossingen van $A\underline{x} = \underline{b}$ niet-leeg en begrensd zijn. Dan bestaat V uit alle convexe combinaties van de extreme oplossingen, m.a.w. V is het convexe polytoop met de extreme oplossingen als hoekpunten.

Bewijs. We geven een bewijs uit het ongerijmde. Stel dat er een positieve oplossing \underline{z} is die geen convexe combinatie van extreme oplossingen is; \underline{z} is dan zelf zeker geen extreme oplossing. Er zijn dan twee verschillende positieve oplossingen \underline{u} en \underline{v} en een λ met $0 < \lambda < 1$, zó dat $\underline{z} = \lambda\underline{u} + (1 - \lambda)\underline{v}$. Alle vectoren van de vorm $t\underline{u} + (1 - t)\underline{v} = \underline{v} + t(\underline{u} - \underline{v})$ met $t \in \mathbb{R}$ zijn oplossingen van $A\underline{x} = \underline{b}$. Op grond van 7.9.4.8 vormen de positieve oplossingen van deze vorm een segment S in \mathbb{R}^n . Laat \underline{u}^* en \underline{v}^* de eindpunten van S zijn; dan zijn \underline{u} en \underline{v} , en dus ook \underline{z} , convexe combinaties van \underline{u}^* en \underline{v}^* , terwijl $\underline{u}^* \neq \underline{v}^*$, $\underline{u}^* \neq \underline{z}$, $\underline{v}^* \neq \underline{z}$. Als \underline{u}^* en \underline{v}^* nu beide convexe combinaties van extreme oplossingen zouden zijn, dan zou hetzelfde gelden voor \underline{z} . Tenminste één van beide \underline{u}^* of \underline{v}^* , zeg \underline{u}^* , is dus geen convexe combinatie van extreme oplossingen.

Laat van de vector \underline{z} nu p elementen gelijk aan 0 zijn, dan zijn ook de overeenkomstige elementen van \underline{u}^* en \underline{v}^* gelijk aan 0. We zullen aantonen dat van de vector \underline{u}^* tenminste $p+1$ elementen 0 zijn. Als dit nl. niet zo is, dan kunnen we als in het bewijs van 7.9.4.6 een $\epsilon > 0$ vinden, zo klein dat $\underline{u}^* + \epsilon(\underline{v}^* - \underline{u}^*)$ en $\underline{u}^* - \epsilon(\underline{v}^* - \underline{u}^*)$ beide positieve oplossingen zijn. Het laatste is in tegenspraak met het feit dat \underline{u}^* een eindpunt is van het segment s van positieve oplossingen van de vorm $\underline{u}^* + t(\underline{v}^* - \underline{u}^*)$. Het aantal elementen van \underline{u}^* dat gelijk aan 0 is, is dus $\geq p+1$.

Bedenk dat \underline{u}^* geen convexe combinatie van extreme oplossingen is. We laten nu \underline{u}^* de rol van \underline{z} spelen en herhalen het bovenstaande betoog. We vinden dan een positieve oplossing die geen convexe combinatie van extreme oplossingen is en met tenminste $p+2$ elementen gelijk aan 0. We herhalen het argument; zodoende vinden we positieve oplossingen die geen convexe combinatie van extreme oplossingen zijn, met steeds meer elementen 0. Op een gegeven moment (ga na) bevat de oplossing nog zo weinig positieve elementen, dat de overeenkomstige kolomvectoren van A onafhankelijk zijn en de vector zelf een extreme oplossing is, òf we eindigen in $\underline{0}$, d.i. ook een extreme oplossing. □

Voorbeelden. 8) Bepaal alle positieve oplossingen van:

$$\begin{cases} 2x_1 + 2x_2 - 2x_3 = -1 \\ x_1 - x_2 + 3x_3 = 2 \end{cases} .$$

De verzameling van de positieve oplossingen is begrensd, immers elke oplossing moet ook voldoen aan: $3x_1 + x_2 + x_3 = 1$. Voor we de extreme oplossingen zoeken gaan we over op de normaalvorm.

Zij

$$[A, \underline{b}] = \begin{bmatrix} 1 & -1 & 3 & 2 \\ 2 & 2 & -2 & -1 \end{bmatrix}$$

(verwisseling der vergelijkingen!), dan is

$$[\tilde{A}, \tilde{\underline{b}}] = \begin{bmatrix} 1 & 0 & 1 & \frac{3}{4} \\ 0 & 1 & -2 & -\frac{5}{4} \end{bmatrix} .$$

De oplossing van $A\underline{x} = \underline{b}$ is dan: $\underline{x} = [\frac{3}{4}, -\frac{5}{4}, 0]^T + \lambda[-1, 2, 1]^T$. Er is dus ten hoogste een segment van positieve oplossingen, derhalve twee extreme oplossingen. Om deze te bepalen kunnen we twee wegen bewandelen:

a) Oplossingen zijn positief als $\frac{3}{4} - \lambda \geq 0$, $-\frac{5}{4} + 2\lambda \geq 0$ en $\lambda \geq 0$, dus $\frac{5}{8} \leq \lambda \leq \frac{3}{4}$. De extreme oplossingen zijn

$$(\lambda = \frac{5}{8}) : [\frac{1}{8}, 0, \frac{5}{8}]^T \quad \text{en} \quad (\lambda = \frac{3}{4}) : [0, \frac{1}{4}, \frac{3}{4}]^T.$$

Uit 7.9.4.6 volgt dat deze oplossingen inderdaad extreem zijn.

Willekeurige positieve oplossingen worden gegeven door:

$$\underline{x} = \lambda[\frac{1}{8}, 0, \frac{5}{8}]^T + (1 - \lambda)[0, \frac{1}{4}, \frac{3}{4}]^T \quad (0 \leq \lambda \leq 1).$$

b) Beschouw alle lineair onafhankelijke deelstelsels van de verzameling der kolomvectoren van \tilde{A} :

$$\left\{ \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -2 \end{bmatrix} \right\}.$$

Voor elk van deze deelstelsels proberen we een oplossing van $\tilde{A}\underline{x} = \tilde{\underline{b}}$ te vinden waarvan alleen de elementen met rangnummers gelijk aan die van de kolomvectoren uit het deelstelsel ongelijk 0 zijn.

Onafhankelijk deelstelsel	op te lossen vergelijkingen	oplossing $[x_1, x_2, x_3]^T$
$\begin{bmatrix} 1 \\ 0 \end{bmatrix}$	$x_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} \frac{3}{4} \\ -\frac{5}{4} \end{bmatrix}, x_2 = 0, x_3 = 0$	strijdig
$\begin{bmatrix} 0 \\ 1 \end{bmatrix}$	$x_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{3}{4} \\ -\frac{5}{4} \end{bmatrix}, x_1 = 0, x_3 = 0$	strijdig
$\begin{bmatrix} 1 \\ -2 \end{bmatrix}$	$x_3 \begin{bmatrix} 1 \\ -2 \end{bmatrix} = \begin{bmatrix} \frac{3}{4} \\ -\frac{5}{4} \end{bmatrix}, x_1 = 0, x_2 = 0$	strijdig

Onafhankelijk
deelstelsel

op te lossen
vergelijkingen

oplossing
 $[x_1, x_2, x_3]^T$

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

$$x_1 \begin{bmatrix} 1 \\ 0 \end{bmatrix} + x_2 \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{3}{4} \\ -\frac{5}{4} \end{bmatrix}, x_3 = 0$$

$$\left[\frac{3}{4}, -\frac{5}{4}, 0\right]^T$$

niet positief

$$\begin{bmatrix} 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ -2 \end{bmatrix}$$

$$\begin{cases} x_1 + x_3 = \frac{3}{4} \\ -2x_3 = -\frac{5}{4} \\ x_2 = 0 \end{cases}$$

$$\left[\frac{1}{8}, 0, \frac{5}{8}\right]^T$$

extreme oplossing!

$$\begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -2 \end{bmatrix}$$

$$\begin{cases} x_3 = \frac{3}{4} \\ x_2 - 2x_3 = -\frac{5}{4} \\ x_1 = 0 \end{cases}$$

$$\left[0, \frac{1}{4}, \frac{3}{4}\right]^T$$

extreme oplossing!

$$9) \begin{cases} x_1 - x_3 = 1 \\ x_2 - x_3 = 1 \end{cases}$$

De algemene oplossing is $\underline{x} = [1, 1, 0]^T + \lambda[1, 1, 1]^T$.

We zien dat de verzameling der positieve oplossingen niet begrensd is.

Iedere oplossing waarvoor $\lambda \geq 0$, is positief. Er is één extreme oplossing, nl. $[1, 1, 0]^T$.

We zullen nog eens rechtstreeks afleiden dat $[1, 1, 0]^T$ een extreem element is van de verzameling $\{[1+\alpha, 1+\alpha, \alpha]^T \mid \alpha \geq 0\}$. Zij nl.

$$[1, 1, 0]^T = \lambda[1+\alpha, 1+\alpha, \alpha]^T + (1-\lambda)[1+\beta, 1+\beta, \beta]^T$$

met $0 < \lambda < 1$. Dan volgt

$$\begin{aligned} (1+\alpha)\lambda + (1+\beta)(1-\lambda) &= 1 \\ \alpha\lambda + \beta(1-\lambda) &= 0, \end{aligned}$$

dus $\alpha\lambda = 0$, $\beta(1-\lambda) = 0$, $\alpha = 0$, $\beta = 0$.

De enige extreme oplossing is ook te bepalen met behulp van 7.9.4.7 en wel als volgt:

$$[A, \underline{b}] = \begin{bmatrix} 1 & 0 & -1 & 1 \\ 0 & 1 & -1 & 1 \end{bmatrix}$$

$$\begin{matrix} \underline{a}_1 & \underline{a}_2 & \underline{a}_3 & \underline{b} \end{matrix}$$

Onafhankelijk deelstelsel	op te lossen vergelijkingen	oplossing $[x_1, x_2, x_3]^T$
\underline{a}_1	$x_1 \underline{a}_1 = \underline{b}, x_2 = 0, x_3 = 0$	strijdig
\underline{a}_2	$x_2 \underline{a}_2 = \underline{b}, x_1 = 0, x_3 = 0$	strijdig
\underline{a}_3	$x_3 \underline{a}_3 = \underline{b}, x_1 = 0, x_2 = 0$	$[0, 0, -1]^T$, niet positief
$\underline{a}_1, \underline{a}_2$	$x_1 = 1, x_2 = 1, x_3 = 0$	$[1, 1, 0]^T$, extreme oplossing
$\underline{a}_1, \underline{a}_3$	$\begin{cases} x_1 - x_3 = 1, x_2 = 0 \\ -x_3 = 1 \end{cases}$	$[0, 0, -1]^T$, niet positief
$\underline{a}_2, \underline{a}_3$	$\begin{cases} -x_3 = 1, x_1 = 0 \\ x_2 - x_3 = 1 \end{cases}$	$[0, 0, -1]^T$, niet positief

10) $\begin{cases} x_1 - x_2 + x_3 = 2 \\ x_1 + 2x_2 - x_3 = 2 \end{cases}$

Merk op dat niet geldt: $\sum_{i=1}^2 a_{ij} > 0$ ($j = 1, 2, 3$). Toch is de verzameling der positieve oplossingen begrensd, want een positieve oplossing moet ook voldoen aan $\frac{3}{2} \times$ de eerste vergelijking + $1 \times$ de tweede vergelijking:

$$\frac{5}{2} x_1 + \frac{1}{2} x_2 + \frac{1}{2} x_3 = 5 .$$

We bepalen nu de extreme oplossingen zonder eerst op normaalvorm over te gaan:

$$\underline{a}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \underline{a}_2 = \begin{bmatrix} -1 \\ 2 \end{bmatrix}, \quad \underline{a}_3 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \quad \underline{b} = \begin{bmatrix} 2 \\ 2 \end{bmatrix} .$$

Onafhankelijk deelstelsel	op te lossen vergelijkingen	oplossing $[x_1, x_2, x_3]^T$
\underline{a}_1	$\begin{cases} x_1 = 2, x_2 = 0, x_3 = 0 \\ x_1 = 2 \end{cases}$	$[2, 0, 0]^T$ extreme oplossing
\underline{a}_2	$\begin{cases} -x_2 = 2, x_1 = 0, x_3 = 0 \\ 2x_2 = 2 \end{cases}$	strijdig
\underline{a}_3	$\begin{cases} x_3 = 2, x_1 = 0, x_2 = 0 \\ -x_3 = 2 \end{cases}$	strijdig
$\underline{a}_1, \underline{a}_2$	$\begin{cases} x_1 - x_2 = 2, x_3 = 0 \\ x_1 + 2x_2 = 2 \end{cases}$	$[2, 0, 0]^T$ extreme oplossing, zie boven
$\underline{a}_1, \underline{a}_3$	$\begin{cases} x_1 + x_3 = 2, x_2 = 0 \\ x_1 - x_3 = 2 \end{cases}$	$[2, 0, 0]^T$ zie boven
$\underline{a}_2, \underline{a}_3$	$\begin{cases} -x_2 + x_3 = 2, x_1 = 0 \\ 2x_2 - x_3 = 2 \end{cases}$	$[0, 4, 6]^T$ extreme oplossing

Willekeurige positieve oplossingen worden gegeven door:

$$\underline{x} = \lambda [2, 0, 0]^T + (1 - \lambda) [0, 4, 6]^T = [2\lambda, 4 - 4\lambda, 6 - 6\lambda]^T$$

$(0 \leq \lambda \leq 1)$.

7.9.5. Lineaire ongelijkheden

We beschouwen stelsels ongelijkheden van de vorm

$$\begin{cases} a_1 x_1 + \dots + a_n x_n \leq \alpha \\ b_1 x_1 + \dots + b_n x_n \geq \beta \\ \vdots \qquad \qquad \qquad \vdots \qquad \qquad \qquad \vdots \\ f_1 x_1 + \dots + f_n x_n \leq \eta. \end{cases}$$

Na vermenigvuldiging van sommige ongelijkheden met -1 kan elk stelsel geschreven worden in de vorm

$$\underline{Ax} \leq \underline{b} . \quad (1)$$

Laat de rijvectoren van A zijn: $\underline{v}_1^T := [a_{11}, \dots, a_{1n}]$, \dots , $\underline{v}_m^T := [a_{m1}, \dots, a_{mn}]$; dan is het stelsel $\underline{Ax} \leq \underline{b}$ ook te schrijven als:

$$\begin{cases} (\underline{v}_1, \underline{x}) \leq b_1 \\ \vdots \\ (\underline{v}_m, \underline{x}) \leq b_m . \end{cases} \quad (2)$$

Stelling. De verzameling V van de vectoren $\underline{x} \in \mathbb{R}^n$ die voldoen aan $\underline{Ax} \leq \underline{b}$ is convex.

Bewijs. In de schrijfwijze (2) is duidelijk dat V de doorsnede is van convexe halfruimten. □

Opmerking. In een stelsel als (2) kunnen overtollige ongelijkheden voorkomen; deze kan men weglaten zonder dat de oplossingsverzameling verandert.

Voorbeeld.

$$\begin{cases} -2x_1 - x_2 \leq 4 \\ -x_1 + x_2 \leq 2 \\ -x_1 - 2x_2 \leq 0 \\ x_1 + x_2 \leq 2 . \end{cases}$$

Als we in \mathbb{R}^2 de oplossingsverzameling tekenen, dan blijkt dat de eerste ongelijkheid overtollig is. De oplossingsverzameling is het convexe polytoop met hoekpunten $[0, 2]^T$, $[4, -2]^T$ en $[-\frac{4}{3}, \frac{2}{3}]^T$.

7.9.6. Iets over lineaire programmering

Voorbeeld. Het dieetprobleem.

We bespreken nu een opgave waar het essentieel gaat om het vinden van het maximum of minimum van een lineaire vorm in \underline{x} , $f(\underline{x}) := (\underline{a}, \underline{x})$, als \underline{x} tevens voldoet aan een aantal vergelijkingen of ongelijkheden. Optimaliseringsproblemen als deze behoren tot het terrein der lineaire programmering. In het klassieke voorbeeld van het dieet gaat het over de samenstelling van een gezonde voeding tegen minimale kosten. Laat er n levensmiddelen L_1, \dots, L_n beschikbaar zijn, bijv. brood, aardappelen, bonen, enz. Een dieet

$$\underline{d} = \begin{bmatrix} d_1 \\ \vdots \\ d_n \end{bmatrix}$$

is een lijst van de dagelijks te eten hoeveelheden d_1, \dots, d_n van L_1, \dots, L_n . De benodigde voedingsstoffen zijn S_1, \dots, S_m , bijv. eiwit, vet, koolhydraat, calcium, vitamine A, enz. Laat in $A^{(m \times n)} = [a_{ij}]$ het element a_{ij} aangeven het aantal eenheden S_i in een eenheid L_j . Laat de dagelijkse behoefte aan S_i , b_i eenheden bedragen en zij

$$\underline{b} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix} .$$

Voor een goed dieet is $A\underline{d} \geq \underline{b}$. Als p_j de prijs is van een eenheid L_j en $\underline{p} = [p_1, \dots, p_n]^T$, dan is de dagelijkse prijs van het dieet: $pr(\underline{d}) := (\underline{p}, \underline{d})$. De taak van de diëtist is nu: Vind bij gegeven A , \underline{b} en \underline{p} onder alle vectoren \underline{d} die voldoen aan

$$\begin{cases} \underline{d} \geq \underline{0} \\ A\underline{d} \geq \underline{b} \end{cases} ,$$

een vector waarvoor $(\underline{p}, \underline{d})$ minimaal is.

Algemeen geformuleerd gaat het er in lineaire programmeringsproblemen om een \underline{x} te bepalen die voldoet aan

$$\begin{cases} \underline{x} \geq \underline{0} , \\ A\underline{x} \geq \underline{b} \text{ of } A'\underline{x} \leq \underline{b}' , \end{cases}$$

en waarvoor een lineaire functionaal $(\underline{p}, \underline{x})$ minimaal of maximaal is.

Korter geformuleerd luidt het probleem: Vind onder de positieve oplossingen van een stelsel lineaire ongelijkheden een oplossing waarvoor een lineaire functionaal maximaal of minimaal is.

In 7.9.5 is opgemerkt dat de oplossingen van (1) een (eventueel lege) convexe verzameling vormen. Het probleem komt dus neer op het bepalen van de extreme waarden van een lineaire functionaal op een convexe verzameling.

Loze veranderlijken

Door een eenvoudige kunstgreep is het mogelijk het stelsel lineaire ongelijkheden te vervangen door een stelsel gelijkheden zodat we in de situatie van 7.9.4 terechtkomen.

Laat het stelsel zijn:

$$\begin{cases} a_{11}x_1 + \dots + a_{1n}x_n \leq b_1 \\ \vdots \qquad \qquad \qquad \vdots \qquad \qquad \qquad \vdots \\ a_{m1}x_1 + \dots + a_{mn}x_n \leq b_m \\ x_1 \geq 0, \dots, x_n \geq 0. \end{cases} \quad (2)$$

Aan de x_1, \dots, x_n voegen we nu toe m nieuwe variabelen y_1, \dots, y_m , de zgn. loze veranderlijken. We proberen op te lossen:

$$\begin{cases} y_1 + a_{11}x_1 + \dots + a_{1n}x_n = b_1 \\ y_2 + a_{21}x_1 + \dots + a_{2n}x_n = b_2 \\ \vdots \qquad \qquad \qquad \vdots \qquad \qquad \qquad \vdots \qquad \qquad \qquad \vdots \\ y_m + a_{m1}x_1 + \dots + a_{mn}x_n = b_m \\ y_1 \geq 0, \dots, y_m \geq 0, x_1 \geq 0, \dots, x_n \geq 0. \end{cases} \quad (3)$$

Als $y_1 = \beta_1, \dots, y_m = \beta_m, x_1 = \alpha_1, \dots, x_n = \alpha_n$ een oplossing van (3) is, is $\underline{x} = [\alpha_1, \dots, \alpha_n]^T$ een oplossing van (2), de afgesneden oplossing. Omgekeerd, is $\underline{x} = \underline{\alpha}$ een oplossing van (2) dan vindt men een uitbreiding tot een oplossing van (3) door te stellen:

$$y_i = b_i - (a_{i1}\alpha_1 + \dots + a_{in}\alpha_n) \quad (i = 1, \dots, m).$$

Eigenschap. Als we een extreme oplossing van (2) uitbreiden tot een oplossing van (3), dan is deze een extreme oplossing. (Bewijs dit.) Als we dus alle extreme oplossingen van (3) kennen, dan komen alle extreme oplossingen van (2) voor onder de afgesneden oplossingen.

We formuleren nu de hoofdstelling van deze paragraaf.

Stelling. Zij C een convex polytoop in \mathbb{R}^n met hoekpunten $\underline{x}_1, \dots, \underline{x}_k$, dan heeft de lineaire vorm $(\underline{p}, \underline{x})$ op C een globaal minimum en een globaal maximum. Het minimum wordt aangenomen in een hoekpunt (maar mogelijk ook elders); hetzelfde geldt voor het maximum.

Bewijs. We beperken ons tot het geval van een minimum; voor het maximum is het bewijs analoog. Bepaal het kleinste van de k getallen $(\underline{p}, \underline{x}_1), \dots, (\underline{p}, \underline{x}_k)$; laat dit m zijn en zij $(\underline{p}, \underline{x}_\ell) = m$. Het is mogelijk dat $m = (\underline{p}, \underline{x}_i)$ voor meerdere waarden van i . Voor een willekeurige $\underline{x} = \lambda_1 \underline{x}_1 + \dots + \lambda_k \underline{x}_k$ ($\lambda_i \geq 0$, $\sum \lambda_i = 1$) uit C geldt nu

$$\begin{aligned}(\underline{p}, \underline{x}) &= \lambda_1 (\underline{p}, \underline{x}_1) + \lambda_2 (\underline{p}, \underline{x}_2) + \dots + \lambda_k (\underline{p}, \underline{x}_k) \geq \\ &\geq (\lambda_1 + \lambda_2 + \dots + \lambda_k) m = m = (\underline{p}, \underline{x}_\ell) .\end{aligned}$$

□

Gevolg. Als de verzameling V van de positieve oplossingen van $A\underline{x} = \underline{b}$ (of $A\underline{x} \leq \underline{b}$) begrensd is wordt het minimum (of maximum) van $(\underline{p}, \underline{x})$ op V aangenomen in een extreme oplossing.

Opmerkingen. a) Als de verzameling der positieve oplossingen niet begrensd is hoeft $(\underline{p}, \underline{x})$ geen maximum of minimum te bezitten. We beschouwen dit geval niet verder. In het geval dat de verzameling der positieve oplossingen van $A\underline{x} = \underline{b}$ begrensd is, biedt de methode uit 7.9.4 dus de mogelijkheid een \underline{x} te vinden waarvoor $(\underline{p}, \underline{x})$ minimaal (of maximaal) is.

2) De verzameling van alle \underline{x} waar het minimum (resp. maximum) wordt aangenomen is het convexe polytoop opgespannen door die extreme oplossingen waar de waarde van $(\underline{p}, \underline{x})$ gelijk aan het minimum (resp. maximum) is.

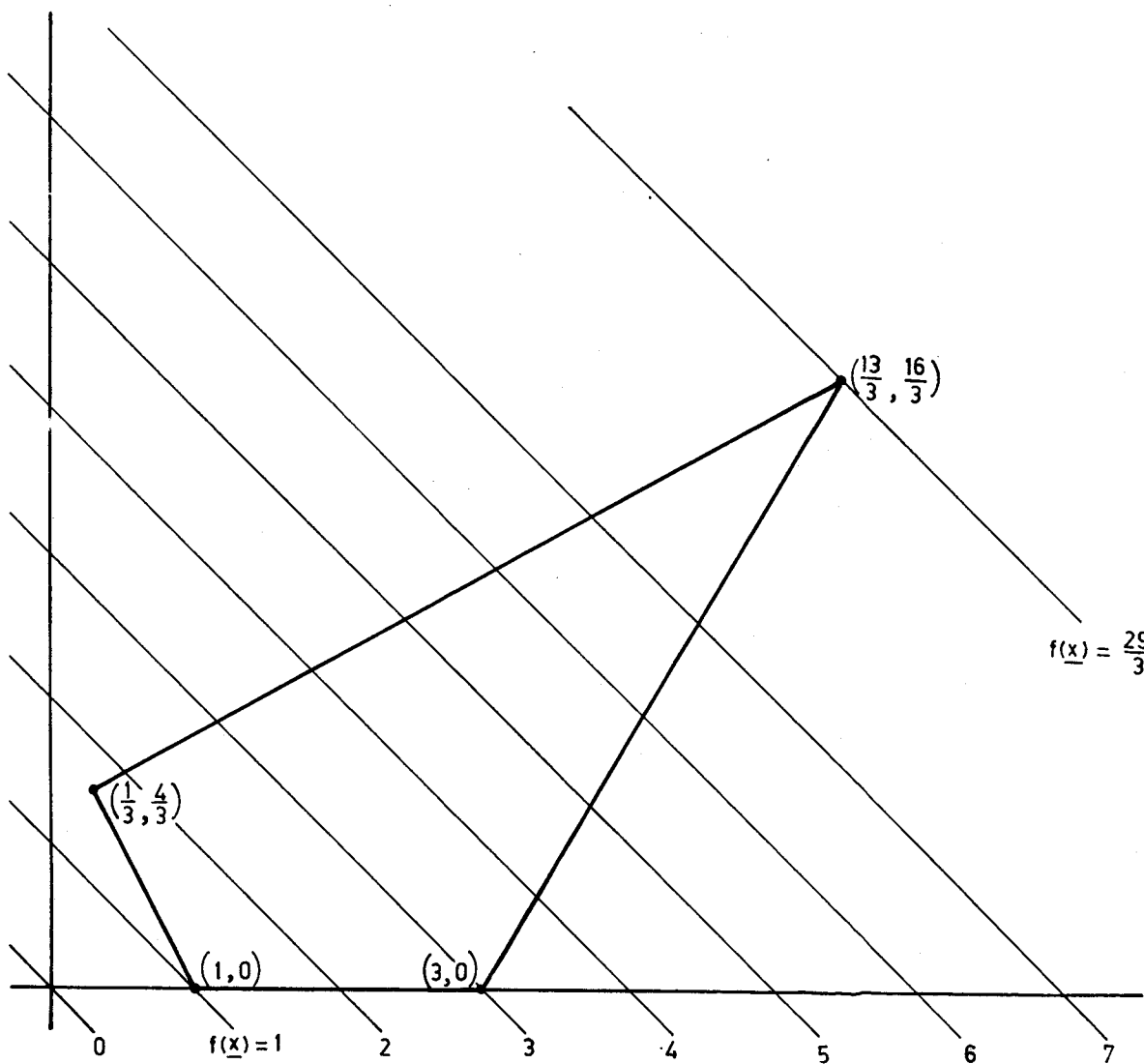
- 3) In het wiskundige vakgebied der operations research wordt uitvoeriger ingegaan op het ontwikkelen van goede algoritmen voor de oplossing van lineaire programmeringsproblemen. De op 7.9.4 berustende werkwijze is in de praktijk vaak zeer omslachtig.
- 4) Lineaire programmeringsproblemen in twee variabelen kunnen grafisch opgelost worden.

Voorbeeld: Bepaal het minimum van $x + y = ([1,1]^T, \underline{x}) = f(\underline{x})$ op de oplossingsverzameling van

$$\begin{cases} -x + y \leq 1 \\ 4x - y \leq 12, & x \geq 0, y \geq 0. \\ -2x - y \leq -2 \end{cases}$$

Deze oplossingsverzameling is in \mathbb{R}^2 voor te stellen door het convexe polytoop met vier extreme elementen: $[0,1]^T$; $[3,0]^T$; $[\frac{13}{3}, \frac{16}{3}]^T$; $[\frac{1}{3}, \frac{4}{3}]^T$.

De waarde van $f(\underline{x})$ in deze hoekpunten is: 1; 3; $\frac{29}{3}$; $\frac{5}{3}$. Het gezochte minimum is 1 en wordt aangenomen in $[1,0]^T$. Dit blijkt ook uit de hoogtekaart van $f(\underline{x})$.



Hoofdstuk 8. Meervoudige integralen

8.1. Integratiegebieden en infinitesimale elementen

We spreken in het volgende over het begrip k-dimensionaal gebied in \mathbb{R}^n , waarbij $1 \leq k \leq n$. Dit begrip zullen we niet algemeen definiëren. We volstaan met het geven van een aantal voorbeelden. Behalve in 8.5 zullen we steeds aannemen dat het gebied begrensd en gesloten is.

k = 1: een kromme in \mathbb{R}^n (vergelijk 6.1.1).

Voorbeelden: een lijnstuk, een gebroken lijn, een cirkel, de rand van een halve cirkelschijf.

k = 2, n = 2: een verzameling in \mathbb{R}^2 , begrensd door een gesloten kromme, bijv. een cirkelschijf, een veelhoek, een verzameling van de vorm $\{(x,y) \mid a \leq x \leq b, 0 \leq y \leq f(x)\}$, waarbij f een positieve functie is.

k = 2, n = 3: een deelverzameling van een gebogen oppervlak in \mathbb{R}^3 , al of niet begrensd door een gesloten kromme, bijv. een boloppervlak, een cylindermantel, een kubusoppervlak.

k = 3, n = 3: een verzameling in \mathbb{R}^3 begrensd door een twee-dimensionaal gebied, bijv. een bol, de cylinder $\{(x,y,z) \mid 0 \leq z \leq 1, x^2 + y^2 \leq 1\}$, een convex polytoop.

In het algemeen zullen we aannemen dat de krommen en oppervlakken die de rand van het gebied vormen, continu zijn. In sommige gevallen zullen we differentieerbaarheid eisen.

De integraal $\int_a^b f(x)dx$, waarin $f: [a,b] \rightarrow \mathbb{R}$ een continue functie is, is als volgt ingevoerd: Men verdeelt het interval $[a,b]$ in kleine deelintervallen $[x_{i-1}, x_i]$ (waarbij $a = x_0 < x_1 < \dots < x_n = b$), kiest getallen $\xi_i \in [x_{i-1}, x_i]$ en vormt de som $\sum_{i=1}^n f(\xi_i)\Delta x_i$, waarbij $\Delta x_i := x_i - x_{i-1}$. Men kan bewijzen dat de limiet van deze som, als $n \rightarrow \infty$ en $\max_i \Delta x_i \rightarrow 0$, bestaat. Deze limiet noemt men als $\int_a^b f(x)dx$.

Bij algemenere integralen gaan we op analoge wijze te werk. Zij G een gebied en zij de functie f continu op G . We verdelen G in kleine deelgebiedjes E_i , die we (infinitesimale) elementen zullen noemen. Van deze elementen bepalen we de maat (= lengte, oppervlakte of volume) $\mu(E_i)$. We kiezen punten $\xi_i \in E_i$ en berekenen de som

8.1.1.
$$\sum_i f(\xi_i) \mu(E_i) .$$

Laat nu het aantal elementen tot ∞ en de afmetingen van de elementen tot 0 naderen. Indien dan de limiet van de som 8.1.1 bestaat, noemt men deze limiet de integraal van f over G . We zullen eerst spreken over de aard van de elementen en over de manier waarop we daaraan een maat toekennen. Daartoe geven we een aantal voorbeelden.

A. Lijnelementen.

- A1) Als $x_1, x_2 \in \mathbb{R}$, $x_1 < x_2$ en $\Delta x := x_2 - x_1$ klein is, dan is het interval $E := [x_1, x_2]$ een element met $\mu(E) = \Delta x$.
- A2) Als $G = K$ een kromme is in \mathbb{R}^2 of \mathbb{R}^3 , gegeven door de parametervoorstelling $\underline{x}(t)$ ($a \leq t \leq b$), en als $a \leq t_1 < t_2 \leq b$ en $\Delta t := t_2 - t_1$ klein is, dan is $E := \{\underline{x}(t) \mid t_1 \leq t \leq t_2\}$ een element met

$$\mu(E) = \Delta s \approx |\underline{x}(t_2) - \underline{x}(t_1)| \approx |\dot{\underline{x}}(t_1)| \Delta t .$$

s heet de booglengte van E . Als de kromme K in \mathbb{R}^2 ligt en gegeven wordt door de vergelijking $y = y(x)$, kunnen we K ook beschrijven door de parametervoorstelling $\underline{x}(t) = (t, y(t))$. Dan geldt $\dot{\underline{x}}(t) = (1, y'(t))$ en dus $|\dot{\underline{x}}(t)| = \sqrt{1 + (y'(t))^2}$. In dat geval is de maat van E (met t vervangen door x):

$$\mu(E) = \Delta s \approx \sqrt{1 + (y'(x))^2} \Delta x .$$

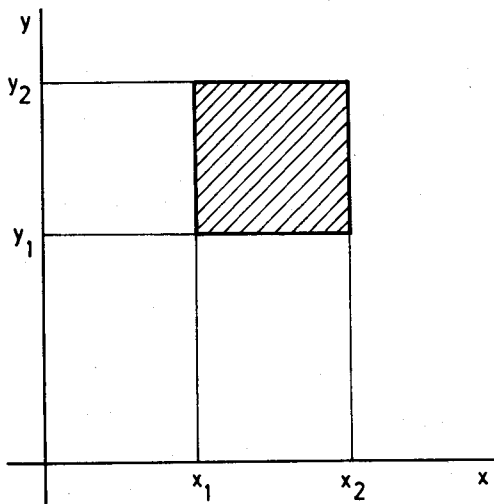
B. Oppervlakte-elementen

B1) In \mathbb{R}^2 is een rechthoek

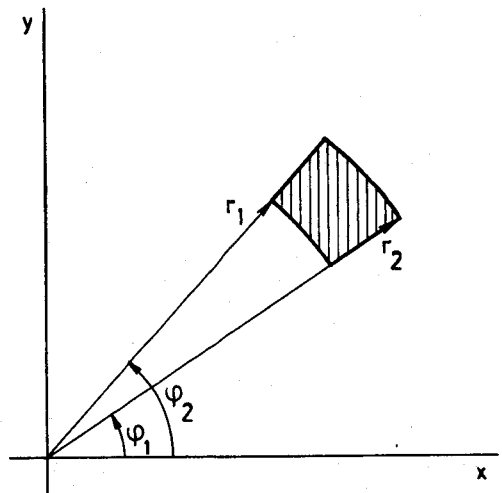
$$E := [x_1, x_2] \times [y_1, y_2] = \{(x, y) \mid x_1 \leq x \leq x_2, y_1 \leq y \leq y_2\}$$

een element als $\Delta x := x_2 - x_1$, $\Delta y := y_2 - y_1$ klein zijn. De maat is $\mu(E) = \Delta x \Delta y$.

B2) In \mathbb{R}^2 is de verzameling $E = \{(r \cos \varphi, r \sin \varphi) \mid \varphi_1 \leq \varphi \leq \varphi_2, r_1 \leq r \leq r_2\}$ een element als $\Delta\varphi := \varphi_2 - \varphi_1$, $\Delta r := r_2 - r_1$ klein zijn. Voor de maat geldt $\mu(E) \approx r_1 \Delta r \Delta\varphi$.



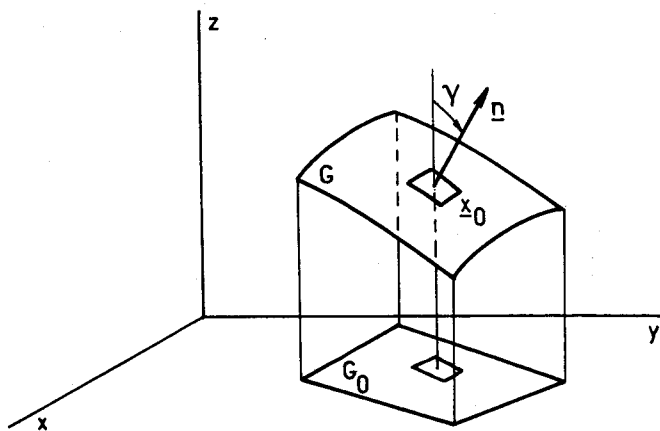
B1



B2

Oppervlakte-elementen.

B3) Zij G een twee-dimensionaal gebied op het oppervlak $z = f(x,y)$ in \mathbb{R}^3 . De projectie van G op het (x,y) -vlak noemen we G_0 .



Zij $\underline{x}_0 = (x_0, y_0, z_0) \in G$ en neem aan dat f differentieerbaar is. Dan wordt het raakvlak V aan G in \underline{x}_0 gegeven door

$$z - z_0 = f_x(x_0, y_0)(x - x_0) + f_y(x_0, y_0)(y - y_0)$$

oftewel $(\underline{n}, \underline{x} - \underline{x}_0) = 0$, waarbij $\underline{n} := (-f_x, -f_y, 1)$ (afgeleiden genomen in (x_0, y_0)). De vector \underline{n} staat dus loodrecht op G . De hoek γ die deze vector

met de positieve z-as maakt voldoet aan $\cos \gamma = 1/\sqrt{1+f_x^2+f_y^2}$. Laat nu E_0 een infinitesimaal element in G_0 zijn en zij $(x_0, y_0) \in E_0$. Met E_0 correspondeert een infinitesimaal element $E := \{(x, y, z) \in G \mid (x, y) \in E_0\}$ van G en een element $E_r := \{(x, y, z) \in V \mid (x, y) \in E_0\}$ van V . Men kan aantonen dat de oppervlakte van E_r een benadering is voor de oppervlakte van E . De oppervlakte van E_r is gelijk aan $\frac{\mu(E_0)}{\cos \gamma}$. Derhalve vinden we, als $\mu(E_0) = \Delta x \Delta y$ (dus als E_0 van het type B1 is):

$$\mu(E) = \Delta \sigma \approx \mu(E_r) = \sqrt{1+f_x^2+f_y^2} \Delta x \Delta y .$$

C. Volume-elementen

C1) Het blok $E := \{(x, y, z) \mid x_1 \leq x \leq x_2, y_1 \leq y \leq y_2, z_1 \leq z \leq z_2\} = [x_1, x_2] \times [y_1, y_2] \times [z_1, z_2]$ is een element in \mathbb{R}^3 als $\Delta x := x_2 - x_1$, $\Delta y := y_2 - y_1$, $\Delta z := z_2 - z_1$ klein zijn. Er geldt $\mu(E) = \Delta x \Delta y \Delta z$.

C2) De verzameling

$$E := \{(r \cos \varphi, r \sin \varphi, z) \mid r_1 \leq r \leq r_2, \varphi_1 \leq \varphi \leq \varphi_2, z_1 \leq z \leq z_2\}$$

met $\Delta r := r_2 - r_1$, $\Delta \varphi := \varphi_2 - \varphi_1$, $\Delta z := z_2 - z_1$ klein, is een element in \mathbb{R}^3 met maat $\mu(E) \approx r_1 \Delta r \Delta \varphi \Delta z$; (r, φ, z) zijn cilindercoördinaten (1.6.2).

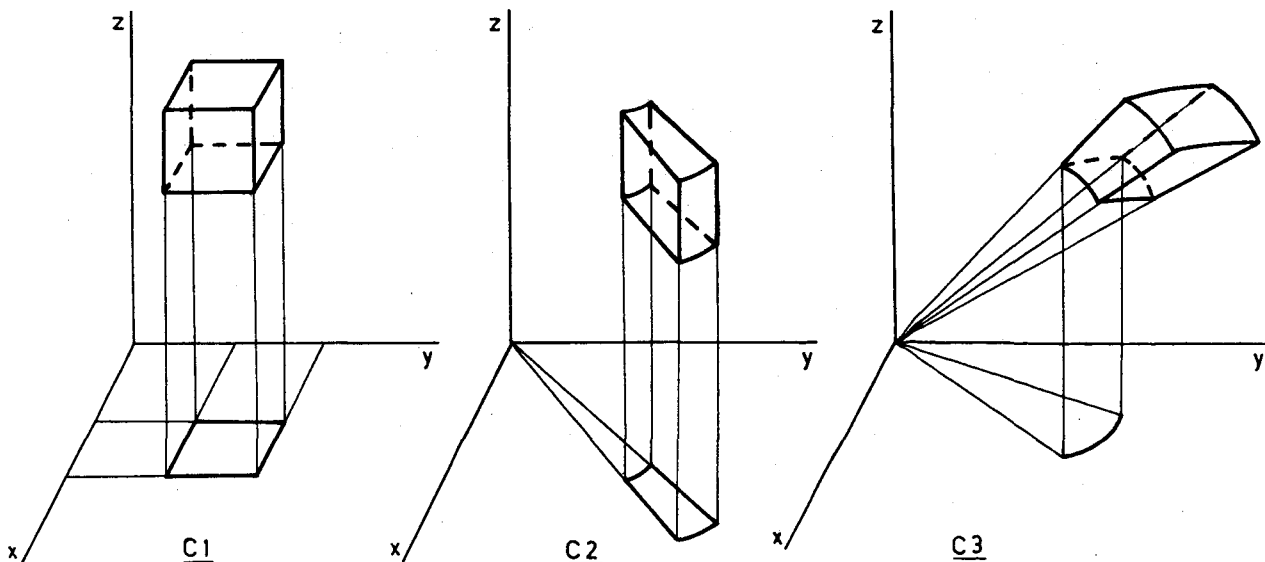
C3) De verzameling

$$E := \{(\rho \sin \theta \cos \varphi, \rho \sin \theta \sin \varphi, \rho \cos \theta) \mid \rho_1 \leq \rho \leq \rho_2, \theta_1 \leq \theta \leq \theta_2, \varphi_1 \leq \varphi \leq \varphi_2\}$$

met $\Delta \rho := \rho_2 - \rho_1$, $\Delta \theta := \theta_2 - \theta_1$, $\Delta \varphi := \varphi_2 - \varphi_1$ klein, is een element in \mathbb{R}^3 met maat

$$\mu(E) \approx \rho_1^2 \sin \theta_1 \Delta \rho \Delta \theta \Delta \varphi ;$$

(ρ, θ, φ) zijn bolcoördinaten (1.6.3).



Volume -elementen .

8.1.2. We maken een paar opmerkingen naar aanleiding van de benaderende formules voor de maat van sommige elementen. In de eerste plaats hebben we niet gedefinieerd wat de echte maat van een element is. We beschouwen dit begrip intuïtief bekend. Bij de benaderende formules is het van belang, dat de fout die bij de benadering wordt gemaakt van kleinere orde is dan de maat zelf. We kunnen deze uitspraak als volgt preciseren: De diameter $\delta(E)$ van een element E is de maximale afstand tussen twee punten van E . (Men kan aantonen dat de maximum betsaat omdat E begrensd en gesloten is (6.7.1).) Als we nu de maat $\mu(E)$ van een element benaderen door de grootheid $\tilde{\mu}(E)$, dan zeggen we dat de fout $\mu(E) - \tilde{\mu}(E)$ van de benadering van kleinere orde is dan $\mu(E)$ als

$$\lim_{\delta(E) \rightarrow 0} \frac{\mu(E) - \tilde{\mu}(E)}{\mu(E)} = 0 .$$

Als aan deze voorwaarde voldaan is, dan kan men aantonen dat in 8.1.1 de echte maat $\mu(E_i)$ vervangen mag worden door de benaderende waarde $\tilde{\mu}(E_i)$, omdat dan bij limietovergang de in de som gemaakte fout tot 0 nadert als $\delta(E_i) \rightarrow 0$. Men kan bewijzen dat in de voorgaande voorbeelden de fout inderdaad van kleinere orde is dan de maat. Zo is bijv. de oppervlakte van het oppervlakte-element van type B2 gelijk aan

$$\mu(E) = \frac{\Delta\varphi}{2\pi} (\pi r_2^2 - \pi r_1^2) = \frac{1}{2}(r_1 + r_2)\Delta r \Delta\varphi .$$

We hebben als benadering genomen $\tilde{\mu}(E) = r_1 \Delta r \Delta\varphi$. De gemaakte fout is dus $\frac{1}{2}(r_2 - r_1)\Delta r \Delta\varphi = \frac{1}{2}(\Delta r)^2 \Delta\varphi$, en er geldt

$$\frac{\mu(E) - \tilde{\mu}(E)}{\mu(E)} = \frac{\Delta r}{r_1 + r_2} \rightarrow 0 \quad \text{als } \delta(E) \rightarrow 0 .$$

Als we voor het element van type B3 nemen $\mu(E_0) = \Delta x \Delta y$ i.p.v.

$\mu(E_r) = \Delta x \Delta y / \cos \gamma$ als benadering voor $\mu(E)$, dan is de fout niet van kleinere orde dan $\mu(E)$.

8.1.3. Definitie. Indien de som 8.1.1 een limiet heeft als het aantal elementen E_i tot ∞ nadert en $\max_i \delta(E_i) \rightarrow 0$, dan noemt men deze limiet de integraal van f over G ; algemene notatie: $\int_G f d\mu$.

Gewoonlijk zullen we echter een andere notatie gebruiken die aangeeft met welk type integraal we te maken hebben. Zonder bewijs vermelden we het volgende resultaat.

8.1.4. Stelling. Zij G een begrensd en gesloten gebied met voldoende gladde rand en zij de functie f continu op G , dan bestaat $\int_G f d\mu$. De integraal is onafhankelijk van het type infinitesimale elementen dat wordt gebruikt in 8.1.1. Bovendien mag $\mu(E_i)$ bij de berekening van de som 8.1.1 worden vervangen door een benaderende uitdrukking, mits de gemaakte fout van kleinere orde is dan $\mu(E_i)$.

We zullen niet precies omschrijven wat we met "voldoend gladde rand" bedoelen. We merken alleen op dat in alle voorbeelden die we in dit hoofdstuk zullen tegenkomen, aan deze voorwaarde is voldaan.

Tenslotte merken we op dat de elementen E_i het gebied G niet volledig hoeven te overdekken, mits de maat van het niet overdekte stuk tot 0 nadert. Zo kan men bijv. een cirkelschijf in \mathbb{R}^2 niet verdelen in elementen van het type B1. Door de diameter klein genoeg te kiezen kan men echter wel de cirkelschijf "op een willekeurig klein gebied na" overdekken.

De volgende eigenschappen volgen onmiddellijk uit de definitie van integraal.

8.1.5. Eigenschappen.

1)
$$\int_G (\lambda_1 f + \lambda_2 g) d\mu = \lambda_1 \int_G f d\mu + \lambda_2 \int_G g d\mu \quad \text{voor alle } \lambda_1, \lambda_2 \in \mathbb{R}.$$

2)
$$\int_{G_1} f d\mu + \int_{G_2} f d\mu = \int_{G_1 \cup G_2} f d\mu,$$

mits G_1 en G_2 disjunct zijn, of ten hoogste de rand gemeenschappelijk hebben.

3) Zij $f \geq g$ op G , dan is
$$\int_G f d\mu \geq \int_G g d\mu.$$

4)
$$\int_G d\mu = \mu(G),$$
 d.i. de lengte, de oppervlakte of het volume van G , al naar gelang G een één-, twee- of drie-dimensionaal gebied is.

8.1.6. We geven een aantal voorbeelden van situaties waarbij meervoudige integralen optreden.

1) Zij $G = [a,b]$, $f(x) \geq 0$ voor $x \in [a,b]$, dan is $\int_G f d\mu$ de oppervlakte van het vlakdeel $\{(x,y) \mid a \leq x \leq b, 0 \leq y \leq f(x)\}$.

2) Zij G een twee-dimensionaal gebied in \mathbb{R}^2 en zij $f(x,y) \geq 0$ voor $(x,y) \in G$, dan is $\int_G f d\mu$ het volume van het deel van de ruimte $\{(x,y,z) \mid (x,y) \in G, 0 \leq z \leq f(x,y)\}$.

3) Een groot aantal voorbeelden zijn te ontleen aan de mechanica. Zij G een (één-, twee of drie-dimensionaal) lichaam met massadichtheid ρ .

(i) De totale massa van G is gelijk aan $\int_G \rho d\mu$.

(ii) Het massatraagheidsmoment van G t.o.v. de z -as is $\int_G (x^2 + y^2) \rho d\mu$.

(iii) Het zwaartepunt (ξ, η, ζ) van G wordt gegeven door

$$\xi = \frac{1}{M} \int_G x \rho d\mu, \quad \eta = \frac{1}{M} \int_G y \rho d\mu, \quad \zeta = \frac{1}{M} \int_G z \rho d\mu$$

waarbij $M := \int_G \rho d\mu$ de totale massa van G is.

4) Zij G een (één-, twee- of drie-dimensionaal) gebied belegd met elektrische lading met dichtheid ρ , dan is de totale lading in G gelijk aan

$$\int_G \rho d\mu.$$

8.2. Lijnintegralen

Zij $G = [a,b]$ een interval in \mathbb{R} . Neem als elementen in G lijnelementen van het type A_1 . De integraal van f over G is dan juist de gewone integraal. We gebruiken dan ook de oude notatie: $\int_a^b f(x) dx = \int_G f d\mu$.

Zij $G = K$ een kromme in \mathbb{R}^2 of \mathbb{R}^3 met parametervoorstelling $\underline{x}(t)$ ($a \leq t \leq b$).
Neem als elementen in G lijnelementen van het type A2. De integraal van f over G noteren we dan als

$$\int_K f(\underline{x}) ds := \int_G f d\mu .$$

Op grond van de benadering gegeven voor Δs geldt

$$\int_K f(\underline{x}) ds = \int_a^b f(\underline{x}(t)) |\dot{\underline{x}}(t)| dt .$$

De booglengte van K wordt gegeven door

8.2.1.
$$\int_K ds = \int_a^b |\dot{\underline{x}}(t)| dt .$$

Als K een kromme is in het (x,y) -vlak, gegeven door de vergelijking $y = y(x)$ ($a \leq x \leq b$) met $y(x)$ differentieerbaar, dan volgt

$$\int_K f(\underline{x}) ds = \int_a^b f(x, y(x)) \sqrt{1 + (y'(x))^2} dx .$$

De booglengte van K wordt dan gegeven door

8.2.2.
$$\int_K ds = \int_a^b \sqrt{1 + (y'(x))^2} dx .$$

Voorbeelden. 1) De booglengte van het parabolsegment $y = x^2$, $0 \leq x \leq 1$, is volgens 8.2.2:

$$\begin{aligned} s &= \int_0^1 \sqrt{1 + 4x^2} dx = x\sqrt{1 + 4x^2} \Big|_0^1 - \int_0^1 \frac{4x^2}{\sqrt{4x^2 + 1}} = \\ &= \sqrt{5} - s + \int_0^1 \frac{dx}{\sqrt{4x^2 + 1}} . \end{aligned}$$

Daar

$$\int_0^1 \frac{dx}{\sqrt{4x^2 + 1}} = \frac{1}{2} \int_0^1 \frac{dx}{\sqrt{x^2 + \frac{1}{4}}} = \frac{1}{2} \log(x + \sqrt{x^2 + \frac{1}{4}}) \Big|_0^1 = \frac{1}{2} \log(2 + \sqrt{5}) ,$$

vinden we

$$s = \frac{1}{2}\sqrt{5} + \frac{1}{4} \log(2 + \sqrt{5}) .$$

- 2) De lengte van een boog van de cycloïde $x = at - a \sin t$, $y = a - a \cos t$, $0 \leq t \leq 2\pi$, is volgens 8.2.1''

$$\begin{aligned} s &= \int_0^{2\pi} \sqrt{a^2(1 - \cos t)^2 + a^2 \sin^2 t} dt = a \int_0^{2\pi} \sqrt{2 - 2 \cos t} dt = \\ &= 2a \int_0^{2\pi} \sin \frac{1}{2}t dt = 8a . \end{aligned}$$

Laat K een kromme zijn in het (x,y) -vlak, gegeven door de vergelijking $r = r(\varphi)$ ($\alpha \leq \varphi \leq \beta$), waarbij r, φ poolcoördinaten zijn. Uit $x = r \cos \varphi$, $y = r \sin \varphi$ volgt $\dot{x}^2 + \dot{y}^2 = \dot{r}^2 + \dot{\varphi}^2 r^2$ ($\dot{}$ betekent afgeleide naar φ), waarna 8.2.1 overgaat in

$$s = \int_{\alpha}^{\beta} \sqrt{\dot{r}^2 + r^2} d\varphi .$$

Voorbeeld 3. De booglengte van de cardioïde $r = 2a(1 - \cos \varphi)$ (zie 1.6.1, voorbeeld b) is

$$s = \int_0^{2\pi} \sqrt{4a^2(1 - \cos \varphi)^2 + 4a^2 \sin^2 \varphi} d\varphi = 2a \int_0^{2\pi} \sqrt{2 - 2 \cos \varphi} d\varphi = 16a .$$

Een kromme K gegeven door $y = y(x)$, $a \leq x \leq b$, met $y(x) \geq 0$, wordt gewenteld om de x -as. Gevraagd wordt de oppervlakte te berekenen van het aldus ontstane omwentelingsoppervlak. Men noemt deze oppervlakte wel de ronde oppervlakte van het omwentelingslichaam dat door het omwentelingsoppervlak en de vlakken $x = a$ en $x = b$ begrensd wordt. Ter bepaling van de ronde oppervlakte

verdelen we het omwentelingsoppervlak in "repen" door middel van vlakken loodrecht op de x-as. Deze repen zijn bij benadering kegelmantels met straal y en apothema Δs , dus met oppervlakte $2\pi y \Delta s$. De ronde oppervlakte wordt nu gegeven door

$$\mathcal{O} = \int_K 2\pi y ds = \int_a^b 2\pi y \sqrt{1 + (y'(x))^2} dx .$$

Als de kromme K gegeven is in poolcoördinaten door $r = r(\varphi)$, $\alpha \leq \varphi \leq \beta$, dan is de ronde oppervlakte

$$\mathcal{O} = \int_{\alpha}^{\beta} 2\pi r \sin \varphi \sqrt{r^2 + \dot{r}^2} d\varphi .$$

Voorbeelden. 4) De lemniscaat $r^2 = \cos 2\varphi$ wordt gewenteld om de x-as. Met gebruik van $r\dot{r} = -\sin 2\varphi$ volgt voor de ronde oppervlakte

$$\begin{aligned} \mathcal{O} &= 2 \int_0^{\pi/4} 2\pi \sin \varphi \sqrt{r^4 + r^2 \dot{r}^2} d\varphi = \\ &= 4\pi \int_0^{\pi/4} \sin \varphi \sqrt{\cos^2 2\varphi + \sin^2 2\varphi} d\varphi = 4\pi \int_0^{\pi/4} \sin \varphi d\varphi = 4\pi(1 - \frac{1}{2}\sqrt{2}) . \end{aligned}$$

5) Gevraagd de oppervlakte \mathcal{O} van het gedeelte van de cylinder $x^2 + y^2 = 2x$, dat buiten de kegel $z^2 = x^2 + y^2$ ligt.

Merk op dat het betreffende gedeelte van de cylinder ook wordt afgesneden door het oppervlak $z^2 = 2x$. Verdeel het cylindergedeelte in repen evenwijdig aan de z-as, met oppervlakte $2z\Delta s$, waarin $z = \sqrt{2x}$ en Δs een lijnelement van de grondcirkel van de cylinder is.

De gevraagde oppervlakte is voor te stellen door een lijnintegraal over de cirkel C: $x^2 + y^2 = 2x$,

$$\mathcal{O} = 2 \int_C z ds = 4 \int_0^2 \sqrt{2x} \sqrt{1 + (y')^2} dx ,$$

waarin $y = \sqrt{2x - x^2}$, dus

$$\sigma = 4 \int_0^2 \sqrt{2x} \frac{1}{\sqrt{2x-x^2}} dx = 4\sqrt{2} \int_0^2 \frac{dx}{\sqrt{2-x}} = 16.$$

- 6) Een oneindig lange draad geplaatst langs de x-as, is elektrisch geladen met constante ladingsdichtheid. Als gevolg heerst in het (x,y)-vlak een elektrisch veld ter grootte c/y en gericht loodrecht op het (x,y)-vlak. De kracht door dit vlak uitgeoefend op een homogene ladingverdeling langs het parabolosegment K: $y = 1 + \frac{1}{4}x^2$, $-1 \leq x \leq 1$, is gelijk aan

$$\begin{aligned} F &= c_1 \int_K \frac{c}{y} ds = c_1 c \int_{-1}^1 \frac{\sqrt{1+(y')^2}}{y} dx = \\ &= 2c_1 c \int_0^1 \frac{dx}{\sqrt{1+\frac{1}{4}x^2}} = 4c_1 c \int_0^1 \frac{dx}{\sqrt{x^2+4}} = 4c_1 c \log(x + \sqrt{x^2+4}) \Big|_0^1 = \\ &= 4c_1 c [\log(1+\sqrt{5}) - \log 2]. \end{aligned}$$

8.3. Oppervlakte-integralen

Zij G een twee-dimensionaal gebied in \mathbb{R}^2 . Neem als elementen in G oppervlakte-elementen van het type B1. We noteren dan de integraal van f over G als

$$\iint_G f(x,y) dx dy := \int_G f d\mu.$$

Indien oppervlakte-elementen van het type B2 worden genomen, dan schrijven we

$$\iint_G f(r,\varphi) r dr d\varphi := \int_G f d\mu.$$

Hierbij is aangenomen dat f als functie van r en φ gegeven is. Als f gegeven is als functie van x en y, dan volgt door een eenvoudige substitutie:

$$f(x,y) = f(r \cos \varphi, r \sin \varphi) =: \hat{f}(r,\varphi).$$

Volgens 8.1.4 geldt

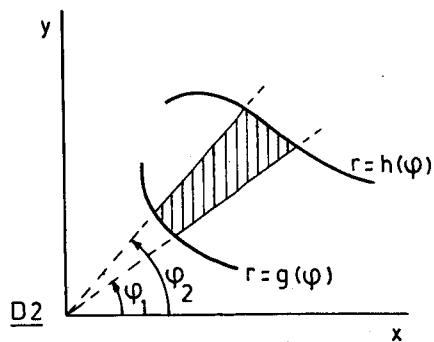
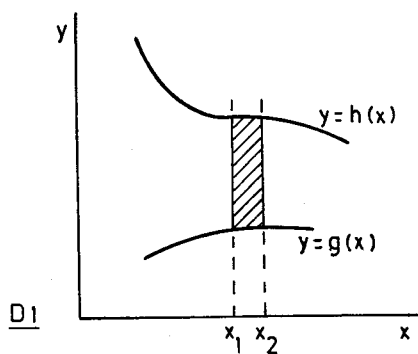
$$\iint_G f(r \cos \varphi, r \sin \varphi) r dr d\varphi = \iint_G f(x,y) dx dy.$$

De oppervlakte van G wordt gegeven door

$$\iint_G dx dy \quad \text{resp.} \quad \iint_G r dr d\varphi .$$

We beschrijven nu een methode om de berekening van oppervlakte-integralen te herleiden tot de berekening van enkelvoudige integralen. Daartoe voeren we in het begrip partieel infinitesimaal element (p.i. element). Een p.i. element is een gebied waarvan sommige (maar niet alle) afmetingen klein zijn. We geven enige voorbeelden:

- D1) Als $x_1 < x_2$ en $\Delta x := x_2 - x_1$ klein is, dan is het gebied $\{(x,y) \mid x_1 \leq x \leq x_2, g(x) \leq y \leq h(x)\}$ een p.i. element in \mathbb{R}^2 . Hierbij zijn g en h continue functies met $g(x) \leq h(x)$ voor $x_1 \leq x \leq x_2$.
- D2) Als $\varphi_1 < \varphi_2$ en $\Delta\varphi := \varphi_2 - \varphi_1$ klein is, dan is de verzameling (in poolcoördinaten) gegeven door $\varphi_1 \leq \varphi \leq \varphi_2, g(\varphi) \leq r \leq h(\varphi)$, een p.i. element in \mathbb{R}^2 . Hierbij zijn g en h continue functies met $0 < g(\varphi) \leq h(\varphi)$ voor $\varphi_1 \leq \varphi \leq \varphi_2$.
- D3) De cirkelring $r_1 \leq r \leq r_2$ is een p.i. element in \mathbb{R}^2 als $\Delta r := r_2 - r_1$ klein is.



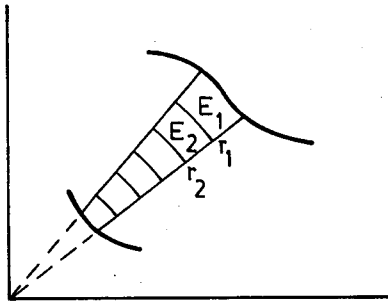
Zij E een p.i. element in \mathbb{R}^2 en zij de functie f continu op E , dan kan men $\int_E f d\mu$ benaderen door een enkelvoudige integraal:

$$(D1): \quad \int_E f d\mu = \iint_E f(x,y) dx dy \approx \left(\int_{g(x_1)}^{h(x_1)} f(x_1,y) dy \right) \Delta x ,$$

$$(D2): \quad \int_E f d\mu = \iint_E f(r,\varphi) r dr d\varphi \approx \left(\int_{g(\varphi_1)}^{h(\varphi_1)} f(r,\varphi_1) r dr \right) \Delta\varphi ,$$

$$(D3): \quad \int_E f \, d\mu = \iint_E f(r, \varphi) r \, dr \, d\varphi \approx \left(\int_0^{2\pi} f(r_1, \varphi) \, d\varphi \right) r_1 \Delta r .$$

Op deze manier worden inderdaad benaderingen verkregen zoals we nog toelichten aan de hand van het tweede voorbeeld.



Verdeel E in elementen E_1, E_2, \dots van het type B2. Op grond van de definitie van de integraal, is $\int_E f \, d\mu$ limiet van en dus te benaderen door $\sum_i f(\xi_i) \mu(E_i)$, waarbij $\xi_i \in E_i$.

Vervang $\mu(E_i)$ door de benadering $r_i \Delta r_i \Delta \varphi$ (bedenk dat $\Delta \varphi = \varphi_2 - \varphi_1$ voor alle elementen E_i hetzelfde is) en kies voor ξ_i het punt met poolcoördinaten (r_i, φ_1) . Dan vinden we

$$\iint_E f(r, \varphi) r \, dr \, d\varphi \approx \left(\sum_i f(r_i, \varphi_1) r_i \Delta r_i \right) \Delta \varphi .$$

Merk nu op dat de som $\sum_i f(r_i, \varphi_1) r_i \Delta r_i$ een benadering is voor de integraal

$$\int_{g(\varphi_1)}^{h(\varphi_1)} f(r, \varphi_1) r \, dr ,$$

immers als $\max_i \Delta r_i \rightarrow 0$ nadert de som tot de integraal. Derhalve vinden we

$$\iint_E f(r, \varphi) r \, dr \, d\varphi \approx \left(\int_{g(\varphi_1)}^{h(\varphi_1)} f(r, \varphi_1) r \, dr \right) \Delta \varphi .$$

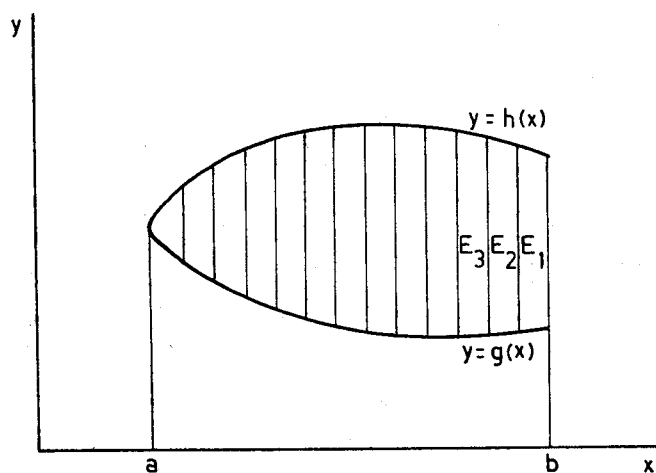
Men kan bewijzen dat de fout die bij deze benadering wordt gemaakt van kleinere orde is dan $\Delta \varphi$. Een dergelijke redenering kan men houden in elk van de genoemde voorbeelden.

We zullen nu laten zien hoe p.i. elementen worden gebruikt bij de berekening van oppervlakte-integralen.

Zij G het gebied

$$8.3.1. \quad G := \{(x,y) \mid a \leq x \leq b, g(x) \leq y \leq h(x)\},$$

waarbij g en h continue functies zijn met $g(x) \leq h(x)$ voor $a \leq x \leq b$.



We verdelen G in p.i. elementen van het type D1. Als f continu is op G dan geldt

$$\iint_G f(x,y) dx dy = \sum_i \iint_{E_i} f(x,y) dx dy \approx \sum_i \left(\int_{g(x_i)}^{h(x_i)} f(x_i, y) dy \right) \Delta x_i.$$

Definiëren we $u(x) := \int_{g(x)}^{h(x)} f(x,y) dy$, dan kan men bewijzen dat u continu is op $[a,b]$ (zie ook 8.7, voorbeeld 2). We vinden

$$8.3.2. \quad \iint_G f(x,y) dx dy \approx \sum_i u(x_i) \Delta x_i.$$

Als $\max_i \Delta x_i \rightarrow 0$, dan nadert de uitdrukking in het rechterlid tot $\int_a^b u(x) dx$.

Bovendien wordt de in 8.3.2 gegeven benadering steeds beter, omdat de fouten in de termen van de som van kleinere orde zijn dan Δx_i . Daarom vinden we bij limietovergang

$$8.3.3. \quad \iint_G f(x,y) dx dy = \int_a^b \left(\int_{g(x)}^{h(x)} f(x,y) dy \right) dx.$$

De integraal in het rechterlid noteren we gemakshalve als

$$\int_a^b dx \int_{g(x)}^{h(x)} f(x,y) dy .$$

We noemen zo'n uitdrukking een herhaalde integraal.

We kunnen in het voorgaande de rol van x en y verwisselen. Als we aannemen dat het gebied gegeven wordt door $G := \{(x,y) \mid \alpha \leq y \leq \beta, k(y) \leq x \leq \ell(y)\}$, dan vinden we

$$8.3.4. \quad \iint_G f(x,y) dx dy = \int_{\alpha}^{\beta} dy \int_{k(y)}^{\ell(y)} f(x,y) dx .$$

Nemen we p.i. elementen van het type D2, dan volgt

$$8.3.5. \quad \iint_G f(r,\varphi) r dr d\varphi = \int_{\alpha}^{\beta} d\varphi \int_{g(\varphi)}^{h(\varphi)} f(r,\varphi) r dr ,$$

aangenomen dat het gebied G in poolcoördinaten wordt beschreven door $\alpha \leq \varphi \leq \beta, g(\varphi) \leq r \leq h(\varphi)$.

Als het gebied G (in poolcoördinaten) gegeven wordt door $a \leq r \leq b, k(r) \leq \varphi \leq \ell(r)$, kunnen we de rol van φ en r verwisselen. We krijgen dan

$$8.3.6. \quad \iint_G f(r,\varphi) r dr d\varphi = \int_a^b r dr \int_{k(r)}^{\ell(r)} f(r,\varphi) d\varphi .$$

Voorbeelden. 1) Zij G de kwart-cirkelschijf gegeven door $x^2 + y^2 \leq 1, x \geq 0, y \geq 0$. Gevraagd wordt $\iint_G e^{x^2+y^2} xy dx dy$ te bepalen.

1°. Met gebruik van 8.3.3 vinden we

$$\begin{aligned} \iint_G e^{x^2+y^2} xy dx dy &= \int_0^1 xe^{x^2} dx \int_0^{\sqrt{1-x^2}} ye^{y^2} dy = \int_0^1 xe^{x^2} dx \cdot \frac{1}{2}(e^{1-x^2} - 1) = \\ &= \frac{1}{2}e \int_0^1 x dx - \frac{1}{2} \int_0^1 xe^{x^2} dx = \frac{1}{4}e - \left(\frac{1}{4}e - \frac{1}{4}\right) = \frac{1}{4} . \end{aligned}$$

2°. Door overgang op poolcoördinaten vinden we met behulp van 8.3.5:

$$\begin{aligned} \iint_G e^{x^2+y^2} xy dx dy &= \iint_G r^3 e^{r^2} \sin \varphi \cos \varphi dr d\varphi = \\ &= \int_0^1 r^3 e^{r^2} dr \int_0^{\pi/2} \sin \varphi \cos \varphi d\varphi = \frac{1}{4} \int_0^1 t e^t dt = \frac{1}{4}, \end{aligned}$$

waarbij de substitutie $r^2 = t$ is toegepast.

2) Zij G de driehoek in het (x,y)-vlak die wordt ingesloten door de rechten $y = 2x$, $2y = x$, $x + y = 3$.

Gevraagd wordt $\iint_G xy dx dy$ te berekenen. Daartoe verdelen we G door de rechte $x = 1$ in twee gebieden G_1 en G_2 , waarna de integralen over G_1 en G_2 worden bepaald met behulp van 8.3.3:

$$\begin{aligned} \iint_G &= \iint_{G_1} + \iint_{G_2} = \int_0^1 x dx \int_{\frac{1}{2}x}^{2x} y dy + \int_1^2 x dx \int_{\frac{1}{2}x}^{3-x} y dy = \\ &= \int_0^1 x dx \cdot \frac{1}{2}(4x^2 - \frac{1}{4}x^2) + \int_1^2 x dx \cdot \frac{1}{2}[(3-x)^2 - \frac{1}{4}x^2] = \\ &= \frac{15}{8} \int_0^1 x^3 dx + \frac{1}{8} \int_1^2 (3x^3 - 24x^2 + 36x) dx = \frac{13}{8}. \end{aligned}$$

Ook hier kunnen we overgaan op poolcoördinaten en 8.3.5 gebruiken:

$$\begin{aligned} \iint_G xy dx dy &= \iint_G r^3 \sin \varphi \cos \varphi dr d\varphi = \\ &= \int_{\arctan \frac{1}{2}}^{\arctan 2} \sin \varphi \cos \varphi d\varphi \int_0^{3(\sin \varphi + \cos \varphi)^{-1}} r^3 dr = \\ &= \frac{81}{4} \int_{\arctan \frac{1}{2}}^{\arctan 2} \frac{\sin \varphi \cos \varphi}{(\sin \varphi + \cos \varphi)^4} d\varphi = \\ &= \frac{81}{4} \int_{\frac{1}{2}}^2 \frac{t}{(1+t)^4} dt = \frac{13}{8}, \end{aligned}$$

waarbij de substitutie $\tan \varphi = t$ is toegepast.

3) Gevraagd $\int_0^{\pi} dx \int_0^{\pi/x} \sin(x+y) e^{\sin y} dy$.

De binnenste integraal is moeilijk. We schrijven de herhaalde integraal eerst als een oppervlakte-integraal over het gebied G gegeven door $0 \leq x \leq \pi$, $0 \leq y \leq \pi/2$. Daarna passen we 8.3.4 toe:

$$\begin{aligned} \int_0^{\pi} dx \int_0^{\pi/2} f(x,y) dy &= \iint_G f(x,y) dx dy = \int_0^{\pi/2} dy \int_0^{\pi} f(x,y) dx = \\ &= \int_0^{\pi/2} e^{\sin y} dy \int_0^{\pi} \sin(x+y) dx = \\ &= - \int_0^{\pi/2} e^{\sin y} dy \cdot [\cos(\pi+y) - \cos y] = \\ &= 2 \int_0^{\pi/2} e^{\sin y} \cos y dy = 2[e^{\sin \frac{\pi}{2}} - e^0] = 2(e-1) . \end{aligned}$$

4) Gevraagd het volume van het lichaam dat wordt ingesloten door de cilindrs $x^2 + y^2 = 1$ en $x^2 + z^2 = 1$. Het lichaam is te beschouwen als deel van het cilindrische gebied $x^2 + y^2 \leq 1$, dat wordt afgesneden door de oppervlakken $z = \sqrt{1-x^2}$ en $z = -\sqrt{1-x^2}$. Op grond van symmetrie is het gevraagde volume gelijk aan

$$8 \iint_G \sqrt{1-x^2} dx dy ,$$

waarbij G het gebied is gegeven door $x^2 + y^2 \leq 1$, $x \geq 0$, $y \geq 0$.

Met behulp van 8.3.3 volgt

$$\begin{aligned} 8 \iint_G \sqrt{1-x^2} dx dy &= 8 \int_0^1 \sqrt{1-x^2} dx \int_0^{\sqrt{1-x^2}} dy = 8 \int_0^1 (1-x^2) dx = \\ &= 8 - \frac{8}{3} = \frac{16}{3} . \end{aligned}$$

8.3.7. Voor de oppervlakte van een gebied G in het (x,y) -vlak geldt $\mathcal{O} = \iint_G dx dy$. Is G gegeven als in 8.3.1, dan vinden we

$$\mathcal{O} = \int_a^b dx \int_{g(x)}^{h(x)} dy = \int_a^b h(x) dx - \int_a^b g(x) dx ,$$

een bekend resultaat.

Is G gegeven in poolcoördinaten door bijv. $\alpha \leq \varphi \leq \beta$, $0 \leq r \leq h(\varphi)$, dan vinden we

8.3.8.
$$\mathcal{O} = \int_{\alpha}^{\beta} d\varphi \int_0^{h(\varphi)} r dr = \frac{1}{2} \int_{\alpha}^{\beta} (h(\varphi))^2 d\varphi .$$

Voorbeelden. 5) Gevraagd de oppervlakte \mathcal{O} van het gebied begrensd door de lemniscaat, in poolcoördinaten gegeven door de vergelijking $r = a\sqrt{\cos 2\varphi}$. Schets de lemniscaat en merk op dat er symmetrie t.o.v. de x -as en t.o.v. de y -as bestaat. De vergelijking $r = a\sqrt{\cos 2\varphi}$ is slechts gedefinieerd voor $0 \leq \varphi \leq \pi/4$, $3\pi/4 \leq \varphi \leq 5\pi/4$ en $7\pi/4 \leq \varphi \leq 2\pi$.

Met behulp van 8.3.8 vinden we voor de gevraagde oppervlakte:

$$\mathcal{O} = 4 \cdot \frac{1}{2} \int_0^{\pi/4} a^2 \cos 2\varphi d\varphi = a^2 .$$

6) Gevraagd wordt de oppervlakte \mathcal{O} van het gebied begrensd door de "lus" van het folium van Descartes (zie 6.5, voorbeeld 3). De vergelijking van het folium luidt in rechthoekige coördinaten, resp. poolcoördinaten:

$$x^3 + y^3 = 3xy \quad \text{resp.} \quad r = \frac{3 \cos \varphi \sin \varphi}{\cos^3 \varphi + \sin^3 \varphi} .$$

Overeenkomstig 8.3.8 is de gevraagde oppervlakte

$$\begin{aligned} \mathcal{O} &= \frac{1}{2} \int_0^{\pi/2} \frac{9 \sin^2 \varphi \cos^3 \varphi}{(\cos^3 \varphi + \sin^3 \varphi)^2} d\varphi = \\ &= \frac{9}{2} \int_0^{\infty} \frac{t^2 dt}{(t^3 + 1)^2} = \frac{3}{2} \int_0^{\infty} \frac{(t^3 + 1)'}{(t^3 + 1)^2} dt = \frac{3}{2} , \end{aligned}$$

waarbij de substitutie $\tan \varphi = t$ is toegepast.

8.3.9. Zij G een twee-dimensionaal gebied gelegen op een oppervlak $z = f(x,y)$ in \mathbb{R}^3 . De projectie van G op het (x,y) -vlak noemen we G_0 , dan is dus $G = \{(x,y,z) \mid (x,y) \in G_0, z = f(x,y)\}$. Zij de functie $h: G \rightarrow \mathbb{R}$ continu op G . De integraal van h over G noteren we als

$$\iint_G h \, d\sigma := \iint_G h \, d\mu .$$

Gebruikmakend van de benadering voor de maat van een infinitesimaal element van het type B3 (zie 8.1) kunnen we schrijven

8.3.10.
$$\iint_G h \, d\sigma = \iint_{G_0} h(x,y,f(x,y)) \sqrt{1 + f_x^2 + f_y^2} \, dx dy .$$

Voor de oppervlakte van G vinden we

8.3.11.
$$\iint_G d\sigma = \iint_{G_0} \sqrt{1 + f_x^2 + f_y^2} \, dx dy .$$

Door 8.3.10, 8.3.11 wordt de integraal over een gebogen oppervlak in \mathbb{R}^3 herleid tot een integraal over een vlak gebied in \mathbb{R}^2 . De laatste is te berekenen door herleiding tot een herhaalde integraal. In sommige situaties is het voordelig de integraal te schrijven in poolcoördinaten:

$$\iint_G h \, d\sigma = \iint_{G_0} h(r,\varphi,\hat{f}(r,\varphi)) \sqrt{1 + \hat{f}_r^2 + \frac{1}{r^2} \hat{f}^2} \, r dr d\varphi ,$$

waarin $\hat{f}(r,\varphi) := f(r \cos \varphi, r \sin \varphi)$.

Voorbeelden. 7) Gevraagd de oppervlakte O van het gedeelte van de kegel $z^2 = x^2 + y^2$, dat binnen de cylinder $x^2 + y^2 = 2x$ ligt.

Voor $z = \sqrt{x^2 + y^2}$ is $\frac{\partial z}{\partial x} = \frac{x}{\sqrt{x^2 + y^2}}$ en $\frac{\partial z}{\partial y} = \frac{y}{\sqrt{x^2 + y^2}}$, waarna met 8.3.11 volgt:

$$O = 2 \iint_{G_0} \sqrt{1 + \left(\frac{\partial z}{\partial x}\right)^2 + \left(\frac{\partial z}{\partial y}\right)^2} \, dx dy = 2 \iint_{G_0} \sqrt{2} \, dx dy = 2\sqrt{2} \pi ;$$

hierbij is G_0 het cirkelgebied $\{(x,y) \mid x^2 + y^2 - 2x \leq 0\}$.

8) Beschouw de bol $B := \{(x,y,z) \mid x^2 + y^2 + z^2 = 1\}$, die belegd is met massa met dichtheid $\rho(x,y,z) = z^2$. Gevraagd wordt de totale massa op B. Deze is op grond van symmetrie gelijk aan twee maal de massa op de halve bol $G := \{(x,y,z) \mid x^2 + y^2 + z^2 = 1, z \geq 0\}$, dus gelijk aan $2 \iint_G \rho \, d\sigma$.

We gebruiken poolcoördinaten. Er geldt: $\rho = 1 - r^2$, $\hat{f}(r,\varphi) = \sqrt{1 - r^2}$, dus $\hat{f}_r = -r/\sqrt{1 - r^2}$, zodat

$$\begin{aligned} 2 \iint_G \rho \, d\sigma &= 2 \iint_{G_0} (1 - r^2) \sqrt{1 + \frac{r^2}{1 - r^2}} \, r \, dr \, d\varphi = \\ &= 2 \iint_{G_0} r \sqrt{1 - r^2} \, dr \, d\varphi = 2 \int_0^{2\pi} d\varphi \int_0^1 r \sqrt{1 - r^2} \, dr = \frac{4}{3} \pi ; \end{aligned}$$

G_0 is de projectie van G op het (x,y)-vlak, i.e. $G_0 = \{(x,y) \mid x^2 + y^2 \leq 1\}$.

8.4. Volume-integralen

8.4.1. Zij G een drie-dimensionaal gebied in \mathbb{R}^3 . Neem als elementen in G volume-elementen van het type C1 (zie 8.1). We noteren dan de integraal van f over G als

$$\iiint_G f(x,y,z) \, dx \, dy \, dz := \int_G f \, d\mu .$$

Indien volume-elementen van het type C2 worden genomen, dan schrijven we

$$\iiint_G f(r,\varphi,z) \, r \, dr \, d\varphi \, dz := \int_G f \, d\mu ,$$

en in geval van volume-elementen van het type C3:

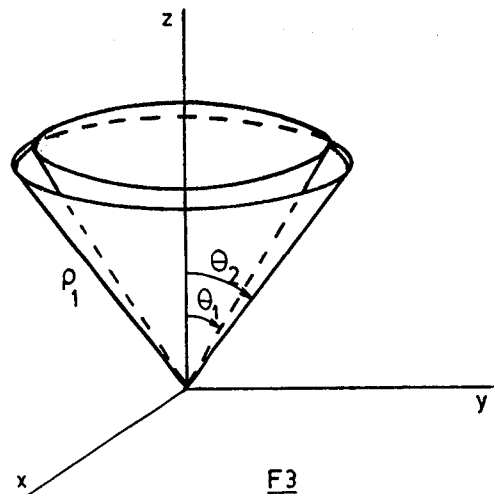
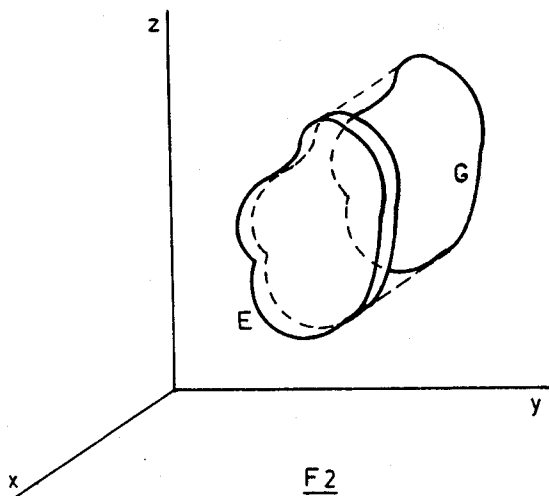
$$\iiint_G f(\rho,\theta,\varphi) \rho^2 \sin \theta \, d\rho \, d\theta \, d\varphi := \int_G f \, d\mu .$$

We merken op dat de integraal weer onafhankelijk is van de keuze van de volume-elementen. Ook hier voeren we p.i. elementen in; we geven enige voorbeelden:

F1) Als $x_1 < x_2$, $y_1 < y_2$ en $\Delta x := x_2 - x_1$, $\Delta y := y_2 - y_1$ klein zijn, dan is het gebied $\{(x,y,z) \mid x_1 \leq x \leq x_2, y_1 \leq y \leq y_2, g(x,y) \leq z \leq h(x,y)\}$ een p.i. element in \mathbb{R}^3 . Hierbij zijn g en h continue functies met $g(x,y) \leq h(x,y)$ voor $x_1 \leq x \leq x_2$, $y_1 \leq y \leq y_2$.

F2) Als $x_1 < x_2$ en $\Delta x := x_2 - x_1$ klein is, dan is de verzameling $\{(x,y,z) \mid x_1 \leq x \leq x_2, (y,z) \in G\}$ een p.i. element in \mathbb{R}^3 . Hierbij is G een gebied in \mathbb{R}^2 .

F3) Als $\theta_1 < \theta_2$ en $\Delta\theta := \theta_2 - \theta_1$ klein is, dan is de verzameling (in bolcoördinaten) gegeven door $0 \leq \rho \leq \rho_1$, $\theta_1 \leq \theta \leq \theta_2$, een p.i. element in \mathbb{R}^3 .



8.4.2. Evenals bij oppervlakte-integralen kan men hier de integraal over een p.i. element benaderen door een integraal van lagere dimensie. Zo geldt, als f continu is op E :

$$(F1) \quad \int_E f \, d\mu = \iiint_E f(x,y,z) \, dx \, dy \, dz \approx \left(\int_{g(x,y)}^{h(x,y)} f(x_1, y_1, z) \, dz \right) \Delta x \Delta y,$$

$$(F2) \quad \int_E f \, d\mu = \iiint_E f(x,y,z) \, dx \, dy \, dz \approx \left(\iint_G f(x_1, y, z) \, dy \, dz \right) \Delta x,$$

$$(F3) \quad \int_E f \, d\mu = \iiint_E f(\rho, \theta, \varphi) \rho^2 \sin \theta \, d\rho \, d\theta \, d\varphi \approx \left(\iint_G h(\rho, \theta_1, \varphi) \rho^2 \, d\rho \, d\varphi \right) \sin \theta_1 \Delta\theta$$

waarbij $G := \{(\rho, \varphi) \mid 0 \leq \rho \leq \rho_1, 0 \leq \varphi \leq 2\pi\}$.

De afleiding van deze formules is dezelfde als in 8.3.

Laat G een drie-dimensionaal gebied in \mathbb{R}^3 zijn en zij de functie f continu op G . Ter berekening van $\int_G f \, d\mu$ verdelen we G in p.i. elementen E_i . We benaderen de integraal van f over E_i door een één- of twee-dimensionale integraal en we sommeren over alle elementen E_i . Zo kunnen we de volume-integraal herleiden tot een herhaalde integraal.

Zij bijv. G het gebied gegeven door

$$8.4.3. \quad G := \{(x, y, z) \mid (x, y) \in G_0, g(x, y) \leq z \leq h(x, y)\},$$

waarbij G_0 een gebied in \mathbb{R}^2 is. Dan vinden we, gebruikmakend van p.i. elementen van het type F_1 ,

$$\iiint_G f(x, y, z) \, dx \, dy \, dz = \iint_{G_0} \left(\int_{g(x, y)}^{h(x, y)} f(x, y, z) \, dz \right) dx \, dy.$$

Zij vervolgens G_0 gegeven door

$$G_0 := \{(x, y) \mid a \leq x \leq b, k(x) \leq y \leq l(x)\},$$

dan is ook de integraal over G_0 te herleiden tot een herhaalde integraal. We vinden

$$\begin{aligned} \iiint_G f(x, y, z) \, dx \, dy \, dz &= \int_a^b \left(\int_{k(x)}^{l(x)} \left(\int_{g(x, y)}^{h(x, y)} f(x, y, z) \, dz \right) dy \right) dx = \\ &=: \int_a^b dx \int_{k(x)}^{l(x)} dy \int_{g(x, y)}^{h(x, y)} f(x, y, z) \, dz. \end{aligned}$$

Voorbeelden. 1) Bepaal $\iiint_G \frac{dx \, dy \, dz}{(x + y + z + 1)^3}$,

waarbij G een gebied is in het eerste octant dat wordt begrensd door de coördinaatvlakken en het vlak $x + y + z = 1$.

$$\begin{aligned}
 \iiint_G \frac{dx dy dz}{(x+y+z+1)^3} &= \int_0^1 dx \int_0^{1-x} dy \int_0^{1-x-y} \frac{dz}{(x+y+z+1)^3} = \\
 &= \int_0^1 dx \int_0^{1-x} dy \left[\frac{-1}{2(x+y+z+1)^2} \right]_{z=0}^{z=1-x-y} = \\
 &= -\frac{1}{8} \int_0^1 dx \int_0^{1-x} dy + \frac{1}{2} \int_0^1 dx \int_0^{1-x} \frac{dy}{(x+y+1)^2} = \\
 &= -\frac{1}{8} \int_0^1 (1-x) dx + \frac{1}{2} \int_0^1 dx \left[\frac{-1}{x+y+1} \right]_{y=0}^{y=1-x} = \\
 &= -\frac{1}{16} - \frac{1}{2} \int_0^1 \left(\frac{1}{2} - \frac{1}{x+1} \right) dx = \frac{1}{2} \log 2 - \frac{5}{16}.
 \end{aligned}$$

- 2) Gevraagd het traagheidsmoment T t.o.v. de z -as van de homogene kubus K :
 $|x| \leq h$, $|y| \leq h$, $|z| \leq h$, met massadichtheid 1.

$$\begin{aligned}
 T &= \iiint_K (x^2 + y^2) dx dy dz = \int_{-h}^h dx \int_{-h}^h (x^2 + y^2) dy \int_{-h}^h dz = \\
 &= 2h \int_{-h}^h (2hx^2 + \frac{2}{3} h^3) dx = \frac{16}{3} h^5.
 \end{aligned}$$

- 3) Het volume van de vier-dimensionale eenheidsbol
 $\{(x,y,z,t) \mid x^2 + y^2 + z^2 + t^2 \leq 1\}$ in \mathbb{R}^4 is

$$V = 2 \iiint_G \sqrt{1-x^2-y^2-z^2} dx dy dz,$$

waarbij het gebied G gegeven wordt door $x^2 + y^2 + z^2 \leq 1$.

We gebruiken bolcoördinaten en passen de substitutie $\rho = \sin \alpha$ toe:

$$V = 2 \iiint_G \sqrt{1-\rho^2} \rho^2 \sin \theta d\rho d\theta d\phi =$$

$$\begin{aligned}
 &= 2 \int_0^1 \rho^2 \sqrt{1-\rho^2} \, d\rho \int_0^\pi \sin \theta \, d\theta \int_0^{2\pi} d\varphi = \\
 &= 2 \cdot 2 \cdot 2\pi \int_0^{\pi/2} \sin^2 \alpha \cos^2 \alpha \, d\alpha = \frac{1}{2} \pi^2 .
 \end{aligned}$$

Stelling (Guldin (1577-1643)). Het volume van een omwentelingslichaam is gelijk aan het produkt van de oppervlakte van een meridiaandoorsnede en de lengte van de baan die het zwaartepunt van de meridiaandoorsnede bij de wenteling beschrijft.

Bewijs. Stel dat het omwentelingslichaam L ontstaat door wenteling om de z-as van het gebied G, gelegen in het (x,z)-vlak. Met gebruik van cylindercoördinaten vinden we voor het volume V van L:

$$V = \iiint_L dx dy dz = \iiint_L r dr d\varphi dz = 2\pi \iint_G r dr dz = 2\pi \iint_G x dx dz .$$

Voor de x-coördinaat ξ van het zwaartepunt van G geldt (zie 8.1.6, voorbeeld 3):

$$\xi \iint_G dx dz = \iint_G x dx dz ,$$

dus

$$V = 2\pi \xi \cdot \text{Opp. } G .$$

□

Voorbeelden. 4) Gevraagd het volume van het lichaam, begrensd door

$$(\sqrt{x^2 + y^2} - a)^2 + z^2 = b^2 , \quad (a > b) .$$

Het lichaam ontstaat door wenteling om de z-as van de cirkelschijf $(x-a)^2 + z^2 = b^2$ in het (x,z)-vlak; het lichaam is een torus. Op grond van de stelling van Guldin is het gevraagde volume

$$2\pi a \cdot \pi b^2 = 2\pi^2 ab^2 .$$

5) Bepaal het zwaartepunt van de helft van de torus (uit voorbeeld 4), waarvoor $y \geq 0$.

Uit symmetrie-overwegingen volgt dat het zwaartepunt op de y-as ligt.

Voor de y-coördinaat η geldt

$$\begin{aligned} \pi^2 ab^2 \eta &= \iiint_G y dx dy dz = \iiint_G r^2 \sin \varphi \, dr d\varphi dz = \\ &= \int_{a-b}^{a+b} r^2 dr \int_0^\pi \sin \varphi \, d\varphi \int_{-p}^p dz, \end{aligned}$$

waarin G de gegeven torushelft is en $p := \sqrt{b^2 - (r-a)^2}$.

Substitueer $r - a = b \cos \alpha$, $0 \leq \alpha \leq \pi$, dan is

$$\begin{aligned} \pi^2 ab^2 \eta &= 2 \cdot 2 \int_{a-b}^{a+b} r^2 \sqrt{b^2 - (r-a)^2} \, dr = \\ &= 4b^2 \int_0^\pi (a + b \cos \alpha)^2 \sin^2 \alpha \, d\alpha = 4b^2 \left(\frac{\pi}{2} a^2 + 0 + \frac{\pi}{8} b^2 \right), \end{aligned}$$

dus

$$\eta = \frac{4a^2 + b^2}{2\pi a}.$$

Voor het volume van een lichaam G in \mathbb{R}^3 geldt $V = \iiint_G dx dy dz$.

Is G gegeven als in 8.4.3, dan vinden we

8.4.4.
$$V = \iint_{G_0} h(x,y) dx dy - \iint_{G_0} g(x,y) dx dy.$$

Voorbeeld 6. Gevraagd het volume van het lichaam, gegeven door

$$0 \leq z \leq 1, \quad x^2 + y^2 \leq z^2.$$

Het lichaam wordt begrensd door het vlak $z = 1$ en de kegel $z^2 = x^2 + y^2$ met 0 als top en de z -as als as. Deze kegel snijdt het vlak $z = 1$ volgens de cirkel gegeven door $x^2 + y^2 = 1$, $z = 1$.

Neem in 8.4.4: $G_0 = \{(x,y) \mid x^2 + y^2 \leq 1\}$, $h(x,y) = 1$, $g(x,y) = \sqrt{x^2 + y^2}$, dan is het gevraagde volume

$$\begin{aligned} \iint_{G_0} 1 \, dx dy - \iint_{G_0} \sqrt{x^2 + y^2} \, dx dy &= \pi - \iint_{G_0} r^2 dr d\varphi = \\ &= \pi - \int_0^1 r^2 dr \int_0^{2\pi} d\varphi = \pi - \frac{2\pi}{3} = \frac{\pi}{3}. \end{aligned}$$

8.5. Oneigenlijke meervoudige integralen

8.5.1. Definitie. Laat G een begrensd en gesloten gebied in \mathbb{R}^2 zijn en zij de functie $f: G \rightarrow \mathbb{R}$ continu op G behalve in het punt $\underline{a} \in G$. Dan zeggen we dat de oneigenlijke integraal $\iint_G f(x,y) dx dy$ absoluut convergent is, als

$$\lim_{\delta \rightarrow 0} \iint_{G \setminus B(\underline{a}, \delta)} |f(x,y)| dx dy \text{ bestaat, waarbij } B(\underline{a}, \delta) := \{\underline{x} \in \mathbb{R}^2 \mid |\underline{x} - \underline{a}| < \delta\}$$

(zie 1.5).

Nodig en voldoende voor de absolute convergentie van $\iint_G f(x,y) dx dy$ is dat

$$I(\delta) := \iint_{G \setminus B(\underline{a}, \delta)} |f(x,y)| dx dy \text{ begrensd is, want } I(\delta) \text{ is een monotoon niet}$$

stijgende functie van δ .

Als $\iint_G f(x,y) dx dy$ absoluut convergent is, dan is de integraal tevens convergent, d.w.z. dan bestaat ook

$$\lim_{\delta \rightarrow 0} \iint_{G \setminus B(\underline{a}, \delta)} f(x,y) dx dy =: \iint_G f(x,y) dx dy .$$

Het bewijs hiervan is analoog aan het bewijs van 2.8.11.

Zonder bewijs vermelden wij de volgende resultaten.

8.5.2. Stelling. Zij $G := \{(x,y) \mid a \leq x \leq b, g(x) \leq y \leq h(x)\}$ en zij f continu op G behalve in $\underline{a} \in G$. Dan is $\iint_G f(x,y) dx dy$ absoluut convergent dan en slechts dan als

$$\int_a^b \left(\int_{g(x)}^{h(x)} |f(x,y)| dy \right) dx$$

convergent is.

In geval van absolute convergentie blijft 8.3.3 geldig, i.e.

$$\iint_G f(x,y) dx dy = \int_a^b dx \int_{g(x)}^{h(x)} f(x,y) dy .$$

Een analoog resultaat geldt als we de rol van x en y verwisselen of als we overgaan op poolcoördinaten. Als bijv. G in poolcoördinaten beschreven wordt door $\alpha \leq \varphi \leq \beta$, $g(\varphi) \leq r \leq h(\varphi)$, dan is $\iint_G f(r,\varphi) r dr d\varphi$ absoluut convergent dan en slechts dan als

$$\int_{\alpha}^{\beta} d\varphi \int_{g(\varphi)}^{h(\varphi)} |f(r,\varphi)| r dr$$

convergent is; in geval van absolute convergentie blijft 8.3.5 gelden.

Uit het voorgaande volgt dat de berekening van een oneigenlijke oppervlakte-integraal is terug te voeren tot de berekening van een herhaalde integraal (analoog als in 8.3), mits voldaan is aan de voorwaarde van absolute convergentie. In het bijzondere geval dat de integrand positief is, is deze absolute convergentie gelijkwaardig met (gewone) convergentie.

De absolute convergentie van oneigenlijke volume-integralen is op analoge wijze te behandelen. Een analogon van 8.5.2 is dan van toepassing.

Voorbeeld 1. De potentiaal in een punt $\underline{p} \in \mathbb{R}^3$ tengevolge van een geladen lichaam $G \subset \mathbb{R}^3$ met ladingsdichtheid $d(\underline{x})$ wordt gegeven door

$$V(\underline{p}) = \iiint_G \frac{d(\underline{x})}{|\underline{x} - \underline{p}|} dx dy dz .$$

Als $\underline{p} \in G$, kan deze integraal oneigenlijk zijn. We beschouwen het geval, dat $G := \{(x,y,z) \mid x^2 + y^2 + z^2 \leq 1\}$, $\underline{p} = (0,0,p)$ met $p \geq 0$, en $d(\underline{x}) = 1$ voor alle $\underline{x} = (x,y,z) \in G$. Dan vinden we met gebruik van bolcoördinaten:

$$V(\underline{p}) = \iiint_G \frac{\rho^2 \sin \theta}{|\underline{x} - \underline{p}|} d\rho d\theta d\varphi = \iiint_G \frac{\rho^2 \sin \theta}{\sqrt{p^2 + \rho^2 - 2\rho p \cos \theta}} d\rho d\theta d\varphi =$$

$$= 2\pi \int_0^1 \rho^2 d\rho \int_0^\pi \frac{\sin \theta d\theta}{\sqrt{p^2 + \rho^2 - 2p\rho \cos \theta}} .$$

De integrand is positief zodat de absolute convergentie volgt uit de (gewone) convergentie van de herhaalde integraal. De herhaalde integraal is inderdaad convergent zoals blijkt door berekening.

Substitueer $\cos \theta = t$ in de binnenste integraal I, dan vinden we

$$I = \int_{-1}^1 (p^2 + \rho^2 - 2p\rho t)^{-\frac{1}{2}} dt = -\frac{2}{2p\rho} (p^2 + \rho^2 - 2p\rho t)^{\frac{1}{2}} \Big|_{-1}^1 =$$

$$= \frac{|p+\rho| - |p-\rho|}{p\rho} = \begin{cases} 2/p & \text{als } p \geq \rho , \\ 2/\rho & \text{als } p < \rho . \end{cases}$$

De buitenste integraal blijkt eigenlijk te zijn.

Als $0 < p < 1$, splitsen we de buitenste integraal in twee stukken, waarna

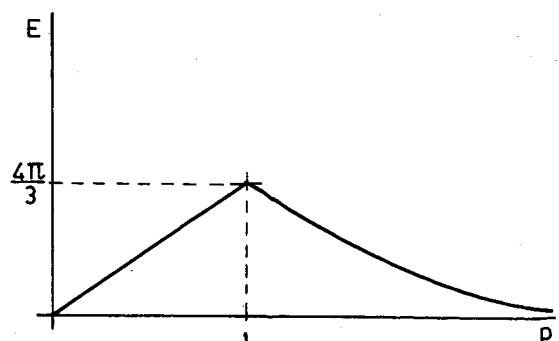
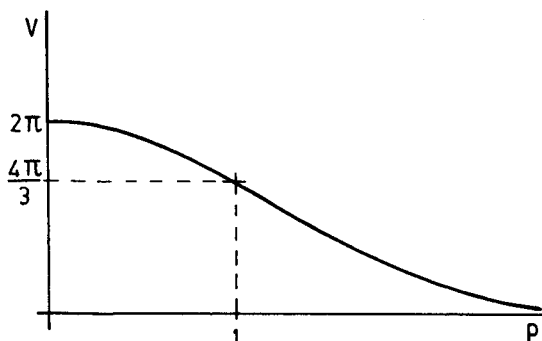
$$V(p) = 2\pi \int_0^p \rho^2 \frac{2}{p} d\rho + 2\pi \int_p^1 \rho^2 \frac{2}{\rho} d\rho = \frac{4\pi}{3} p^2 + 2\pi(1 - p^2) =$$

$$= 2\pi(1 - \frac{1}{3} p^2) .$$

Als $p \geq 1$, dan vinden we

$$V(p) = 2\pi \int_0^1 \rho^2 \frac{2}{p} d\rho = \frac{4\pi}{3p} = \frac{Q}{p} ,$$

waarbij Q de totale lading in G is.



Voor de veldsterkte $E = -\frac{dV}{dp}$ vinden we

$$E = \begin{cases} \frac{4\pi}{3} p & \text{als } 0 \leq p < 1, \\ \frac{4\pi}{3p^2} & \text{als } p \geq 1. \end{cases}$$

Als een oneigenlijke integraal niet absoluut convergent is, dan zijn verwisseling van integratievolgorde en overgang op andere coördinaten in het algemeen niet toegestaan.

Voorbeeld 2. Beschouw de herhaalde integraal

$$I_1 = \int_0^1 dy \int_0^1 \frac{x^2 - y^2}{(x^2 + y^2)^2} dx.$$

In de binnenste integraal substitueren we $x = yt$, dan vinden we

$$\int_0^1 \frac{x^2 - y^2}{(x^2 + y^2)^2} dx = \frac{1}{y} \int_0^{1/y} \frac{t^2 - 1}{(1 + t^2)^2} dt = \frac{-1}{y^2 + 1}.$$

Daaruit volgt $I_1 = -\arctan 1 = -\frac{\pi}{4}$.

Verwissel nu de integratievolgorde in I_1 en beschouw de herhaalde integraal

$$I_2 = \int_0^1 dx \int_0^1 \frac{x^2 - y^2}{(x^2 + y^2)^2} dy.$$

Analoog als voren vinden we $I_2 = +\frac{\pi}{4}$ (ga na). Blijkbaar is de integraal

$\iint_G \frac{x^2 - y^2}{(x^2 + y^2)^2} dx dy$, $G: 0 \leq x \leq 1, 0 \leq y \leq 1$, niet absoluut convergent.

- 5.3. Definitie. Laat G een niet-begrensd en gesloten gebied in \mathbb{R}^2 zijn en zij de functie $f: G \rightarrow \mathbb{R}$ continu op G . Dan zeggen we dat de oneigenlijke integraal $\iint_G f(x,y) dx dy$ absoluut convergent is, als $\lim_{R \rightarrow \infty} \iint_{G \cap B(\underline{0}, R)} |f(x,y)| dx dy$ bestaat, waarbij $B(\underline{0}, R) := \{\underline{x} \in \mathbb{R}^2 \mid |\underline{x}| \leq R\}$.

Nodig en voldoende voor de absolute convergentie van $\iint_G f(x,y) dx dy$ is dat $I(R) := \iint_{G \cap B(\underline{0}, R)} |f(x,y)| dx dy$ begrensd is, immers $I(R)$ is een monotoon niet-dalende functie van R .

Als $\iint_G f(x,y) dx dy$ absoluut convergent is, dan is de integraal tevens convergent, d.w.z. dan bestaat ook

$$\lim_{R \rightarrow \infty} \iint_{G \cap B(\underline{0}, R)} f(x,y) dx dy =: \iint_G f(x,y) dx dy .$$

Voor de hier ingevoerde oneigenlijke integraal geldt een stelling analoog aan 8.5.2. De berekening van een oppervlakte-integraal over een niet begrensd gebied is dus terug te voeren tot de berekening van een herhaalde integraal (analoog als in 8.3), mits voldaan is aan de voorwaarde van absolute convergentie. De voorgaande resultaten hebben een analoon voor volume-integralen over een niet begrensd gebied in \mathbb{R}^3 .

In de volgende voorbeelden is de integrand steeds positief zodat de absolute convergentie volgt uit de (gewone) convergentie van de herhaalde integralen.

Voorbeelden. 3) Bereken $\int_0^\infty y^5 dy \int_{y^2}^\infty \frac{dx}{1+x^8}$.

Daar de binnenste integraal te moeilijk is, schrijven we de herhaalde integraal als een oppervlakte-integraal van het niet begrensde gebied $G := \{(x,y) \mid y \geq 0, x \geq y^2\}$. Dan volgt

$$\begin{aligned} \int_0^\infty y^5 dy \int_{y^2}^\infty \frac{dx}{1+x^8} &= \iint_G \frac{y^5}{1+x^8} dx dy = \int_0^\infty \frac{dx}{1+x^8} \int_0^{\sqrt{x}} y^5 dy = \\ &= \int_0^\infty \frac{dx}{1+x^8} \cdot \frac{1}{6} x^3 = \frac{1}{24} \int_0^\infty \frac{(x^4)'}{1+x^8} dx = \frac{\pi}{48} . \end{aligned}$$

4) Bereken $\iint_G \frac{dx dy}{(x^2 + y^2)^2}$, waarin $G: 4x^2 + y^2 \geq 1$.

We gebruiken poolcoördinaten r, φ . De vergelijking van de ellips $4x^2 + y^2 = 1$ wordt dan: $r^2(4 \cos^2 \varphi + \sin^2 \varphi) = 1$. De integraal gaat nu over in

$$\begin{aligned} \iint_G \frac{dx dy}{(x^2 + y^2)^2} &= \iint_G \frac{r dr d\varphi}{r^4} = \\ &= \int_0^{2\pi} d\varphi \int_{(4 \cos^2 \varphi + \sin^2 \varphi)^{-\frac{1}{2}}}^{\infty} \frac{dr}{r^3} = \frac{1}{2} \int_0^{2\pi} d\varphi \left[-\frac{1}{r^2} \right]_{(4 \cos^2 \varphi + \sin^2 \varphi)^{-\frac{1}{2}}}^{\infty} = \\ &= \frac{1}{2} \int_0^{2\pi} (4 \cos^2 \varphi + \sin^2 \varphi) d\varphi = \frac{5\pi}{2}. \end{aligned}$$

5) Bereken $\iiint_G \frac{dx dy dz}{(x^2 + y^2 + z^2)^2}$, waarin $G: |z| \geq 3$.

1°. Gebruik cylindercoördinaten r, φ, z , dan gaat de integraal over in

$$\begin{aligned} \iiint_G \frac{r dr d\varphi dz}{(r^2 + z^2)^2} &= 2 \int_3^{\infty} dz \int_0^{\infty} \frac{r dr}{(r^2 + z^2)^2} \int_0^{2\pi} d\varphi = \\ &= 2\pi \int_3^{\infty} dz \left[\frac{-1}{r^2 + z^2} \right]_{r=0}^{r=\infty} = 2\pi \int_3^{\infty} \frac{dz}{z^2} = \frac{2\pi}{3}. \end{aligned}$$

2°. Gebruik bolcoördinaten ρ, θ, φ , dan gaat de integraal over in

$$\begin{aligned} \iiint_G \frac{\rho^2 \sin \theta d\rho d\theta d\varphi}{\rho^4} &= 2 \int_0^{\pi/2} d\theta \int_{3/\cos \theta}^{\infty} \frac{\sin \theta}{\rho^2} d\rho \int_0^{2\pi} d\varphi = \\ &= 4\pi \int_0^{\pi/2} d\theta \sin \theta \left[-\frac{1}{\rho} \right]_{3/\cos \theta}^{\infty} = \frac{4\pi}{3} \int_0^{\pi/2} \sin \theta \cos \theta d\theta = \frac{2\pi}{3}. \end{aligned}$$

23.4.75

8.6. Gammafunctie, betafunctie

In deze paragraaf bespreken we een tweetal speciale functies die veelvuldig worden gebruikt in de mathematische fysica en de statistiek.

We definiëren de gammafunctie door

$$\Gamma(x) := \int_0^{\infty} e^{-t} t^{x-1} dt, \quad (x > 0).$$

De convergentie van deze integraal moet onderzocht worden in $t = 0$ en $t = \infty$.

We splitsen het integratie-interval: $\int_0^{\infty} = \int_0^1 + \int_1^{\infty}$.

Zij $0 < a < b < \infty$ en $a \leq x \leq b$. Voor $t \in [0,1]$ geldt de ongelijkheid

$$0 \leq e^{-t} t^{x-1} \leq t^{a-1}.$$

Omdat $\int_0^1 t^{a-1} dt$ convergent is (zie 2.8, voorbeeld 6), is ook $\int_0^1 e^{-t} t^{x-1} dt$ convergent. Als $x \geq 1$, dan is de integraal over $[0,1]$ eigenlijk.

Voor $t \in [1,\infty)$ geldt

$$0 \leq e^{-t} t^{x-1} \leq e^{-t} t^{b-1} \leq Kt^{-2}$$

voor zekere $K > 0$, immers $e^{-t} t^{b+1} \rightarrow 0$ als $t \rightarrow \infty$ en dus $e^{-t} t^{b+1} \leq K$. (Een continue functie op $[1,\infty)$, die een limiet heeft als $t \rightarrow \infty$, is begrensd.)

Omdat $\int_1^{\infty} t^{-2} dt$ convergent is (zie 2.8, voorbeeld 7), is ook $\int_1^{\infty} e^{-t} t^{x-1} dt$ convergent.

Met behulp van partiële integratie leiden we af

$$\Gamma(x+1) = \int_0^{\infty} e^{-t} t^x dt = -e^{-t} t^x \Big|_0^{\infty} + x \int_0^{\infty} e^{-t} t^{x-1} dt,$$

oftewel

$$\Gamma(x+1) = x\Gamma(x), \quad (x > 0).$$

Wegens $\Gamma(1) = \int_0^{\infty} e^{-t} dt = 1$, volgt na herhaalde toepassing van 8.6.2:

$$\Gamma(n+1) = n!, \quad (n = 0,1,2,\dots).$$

Vervolgens definiëren we de betafunctie door

$$8.6.4. \quad B(p,q) := \int_0^1 t^{p-1} (1-t)^{q-1} dt, \quad (p > 0, q > 0).$$

De convergentie van deze integraal kan worden aangetoond via splitsing van het integratie-interval: $\int_0^1 = \int_0^{\frac{1}{2}} + \int_{\frac{1}{2}}^1$ (ga na).

Men kan gemakkelijk afleiden dat $B(p,q) = B(q,p)$.

Door de substitutie $t = \sin^2 \varphi$ in 8.6.4 vinden we een andere integraalvoorstelling voor de betafunctie:

$$8.6.5. \quad B(p,q) = 2 \int_0^{\pi/2} (\sin \varphi)^{2p-1} (\cos \varphi)^{2q-1} d\varphi.$$

Het verband tussen de gamma- en de betafunctie wordt gelegd door:

$$8.6.6. \quad \text{Stelling. } B(p,q) = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}.$$

Bewijs. Ga uit van

$$\Gamma(p)\Gamma(q) = \int_0^\infty e^{-t} t^{p-1} dt \int_0^\infty e^{-s} s^{q-1} ds.$$

Door de substituties $t = x^2$, $s = y^2$ gaat deze integraal over in

$$\begin{aligned} \Gamma(p)\Gamma(q) &= 4 \int_0^\infty e^{-x^2} x^{2p-1} dx \int_0^\infty e^{-y^2} y^{2q-1} dy = \\ &= 4 \iint_{\mathbb{R}_+^2} e^{-x^2-y^2} x^{2p-1} y^{2q-1} dx dy, \end{aligned}$$

waarin $\mathbb{R}_+^2 := \{(x,y) \in \mathbb{R}^2 \mid x \geq 0, y \geq 0\}$. De integraal over \mathbb{R}_+^2 convergeert absoluut vanwege de absolute convergentie van de herhaalde integraal. We mogen op poolcoördinaten overgaan, waarna volgt

$$\begin{aligned} \Gamma(p)\Gamma(q) &= 4 \iint_{\mathbb{R}_+^2} e^{-r^2} r^{2p+2q-2} (\cos \varphi)^{2p-1} (\sin \varphi)^{2q-1} r dr d\varphi = \\ &= 2 \int_0^\infty e^{-r^2} (r^2)^{p+q-1} (r^2)' dr \int_0^{\pi/2} (\cos \varphi)^{2p-1} (\sin \varphi)^{2q-1} d\varphi = \\ &= \Gamma(p+q)B(q,p) = \Gamma(p+q)B(p,q) . \end{aligned} \quad \square$$

Nemen we in 8.6.5, 8.6.6 $p = q = \frac{1}{2}$, dan vinden we

$$[\Gamma(\frac{1}{2})]^2 = B(\frac{1}{2}, \frac{1}{2}) = 2 \int_0^{\pi/2} d\varphi = \pi .$$

Wegens $\Gamma(\frac{1}{2}) > 0$ volgt hieruit

8.6.7. $\Gamma(\frac{1}{2}) = \sqrt{\pi} .$

Voorbeeld 1. Gevraagd de oppervlakte 0 van het gebied begrensd door de astroïde $x^{2/3} + y^{2/3} = 1$.

Schets de astroïde en merk op dat er symmetrie t.o.v. de x-as en t.o.v. de y-as bestaat. De in het eerste kwadrant gelegen boog van de astroïde wordt beschreven door $y = (1 - x^{2/3})^{3/2}$, $0 \leq x \leq 1$.

De gevraagde oppervlakte is

$$O = 4 \int_0^1 (1 - x^{2/3})^{3/2} dx .$$

Door middel van de substitutie $x^{2/3} = t$ is deze integraal te herleiden tot

$$\begin{aligned} O &= 6 \int_0^1 (1 - t)^{3/2} t^{1/2} dt = 6B(\frac{5}{2}, \frac{3}{2}) = 6 \frac{\Gamma(\frac{5}{2})\Gamma(\frac{3}{2})}{\Gamma(4)} = \\ &= 6 \frac{\frac{3}{2} \cdot \frac{1}{2} \Gamma(\frac{1}{2}) \cdot \frac{1}{2} \Gamma(\frac{1}{2})}{3!} = \frac{3\pi}{8} . \end{aligned}$$

We kunnen $\Gamma(n + \frac{1}{2})$ door middel van 8.6.2, 8.6.7 recursief berekenen:

$$\begin{aligned}\Gamma\left(n+\frac{1}{2}\right) &= \left(n-\frac{1}{2}\right)\left(n-\frac{3}{2}\right) \cdots \frac{3}{2} \cdot \frac{1}{2} \Gamma\left(\frac{1}{2}\right) = \frac{(2n-1)(2n-3) \cdots 3 \cdot 1}{2^n} \sqrt{\pi} = \\ &= \frac{2n(2n-1)(2n-2)(2n-3) \cdots 3 \cdot 2 \cdot 1}{2^n \cdot 2n(2n-2) \cdots 2} \sqrt{\pi},\end{aligned}$$

oftewel

8.6.8.
$$\Gamma\left(n+\frac{1}{2}\right) = \frac{(2n)!}{2^{2n} n!} \sqrt{\pi}, \quad (n = 0, 1, 2, \dots).$$

We geven enige toepassingen van deze en de voorgaande formules.

In de statistiek speelt de errorfunctie

8.6.9.
$$\operatorname{erf}(x) := \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt$$

een belangrijke rol.

Opmerking. De errorfunctie is voorbeeld van een functie die gedefinieerd is door een niet in elementaire functies uit te drukken integraal. Voor numerieke tabellen van $\operatorname{erf}(x)$, zie M. Abramowitz & I.A. Stegun, Handbook of Mathematics I Functions, 1964.

Beschouw een experiment waarvan de uitkomst $x \in \mathbb{R}$ in zekere mate door het toeval bepaald wordt. Men kan dan invoeren de zgn. kansdichtheid f , zodanig dat de kans op een uitkomst $x \in [a, b]$ gegeven wordt door

$$P([a, b]) = \int_a^b f(t) dt.$$

In vele praktische situaties blijkt de kansdichtheid f te worden gegeven door

$$f(t) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{(t-\mu)^2}{2\sigma^2}\right]$$

waarin μ en σ gegeven reële getallen zijn met $\sigma > 0$. We zeggen dan dat de uitkomst x normaal verdeeld is met gemiddelde μ en standaardafwijking σ . De grafiek van $f(t)$ heet de kromme van Gauss; deze kromme is ééntoppig en symmetrisch t.o.v. $t = \mu$.

In geval van een normale verdeling is $P([a,b])$ uit te drukken in de error-functie. Door middel van de substitutie $t - \mu = \sigma\sqrt{2} s$ volgt namelijk

$$\begin{aligned}
 P([a,b]) &= \frac{1}{\sigma\sqrt{2\pi}} \int_a^b \exp\left[-\frac{(t-\mu)^2}{2\sigma^2}\right] dt = \frac{1}{\sqrt{\pi}} \int_{\frac{a-\mu}{\sigma\sqrt{2}}}^{\frac{b-\mu}{\sigma\sqrt{2}}} e^{-s^2} ds = \\
 &= \frac{1}{\sqrt{\pi}} \int_0^{\frac{b-\mu}{\sigma\sqrt{2}}} e^{-s^2} ds - \frac{1}{\sqrt{\pi}} \int_0^{\frac{a-\mu}{\sigma\sqrt{2}}} e^{-s^2} ds = \frac{1}{2} \operatorname{erf}\left(\frac{b-\mu}{\sigma\sqrt{2}}\right) - \frac{1}{2} \operatorname{erf}\left(\frac{a-\mu}{\sigma\sqrt{2}}\right) .
 \end{aligned}$$

De kans $P((-\infty, \infty))$ op een uitkomst $x \in (-\infty, \infty)$ dient 1 te zijn. Inderdaad blijkt via de substituties $t - \mu = \sigma\sqrt{2} s$, $s^2 = y$, dat

$$\begin{aligned}
 P((-\infty, \infty)) &= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} \exp\left[-\frac{(t-\mu)^2}{2\sigma^2}\right] dt = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} e^{-s^2} ds = \\
 &= \frac{2}{\sqrt{\pi}} \int_0^{\infty} e^{-s^2} ds = \frac{1}{\sqrt{\pi}} \int_0^{\infty} e^{-y} y^{-\frac{1}{2}} dy = \frac{1}{\sqrt{\pi}} \Gamma\left(\frac{1}{2}\right) = 1
 \end{aligned}$$

op grond van 8.6.7.

De volgende integralen zijn door de substitutie $t^2 = s$ te herleiden tot de gammafunctie:

$$\int_0^{\infty} e^{-t^2} t^{2n+1} dt = \frac{1}{2} \int_0^{\infty} e^{-s} s^n ds = \frac{1}{2} \Gamma(n+1) = \frac{1}{2} n! , \quad (n=0,1,2,\dots) ,$$

$$\int_0^{\infty} e^{-t^2} t^{2n} dt = \frac{1}{2} \int_0^{\infty} e^{-s} s^{n-\frac{1}{2}} ds = \frac{1}{2} \Gamma\left(n+\frac{1}{2}\right) = \frac{(2n)! \sqrt{\pi}}{2^{2n+1} n!} ,$$

(n=0,1,2,...) .

Overeenkomstig 8.6.5 geldt

$$I_{mn} := \int_0^{\pi/2} \cos^m \varphi \sin^n \varphi d\varphi = \frac{1}{2} B\left(\frac{m+1}{2}, \frac{n+1}{2}\right) .$$

Als m, n niet negatieve gehele getallen zijn, is I_{mn} expliciet te bepalen. Immers, volgens 8.6.6 is $B(\frac{m+1}{2}, \frac{n+1}{2}) = \frac{\Gamma(\frac{m+1}{2})\Gamma(\frac{n+1}{2})}{\Gamma(\frac{m+n+2}{2})}$ en de hierin optredende waarden van de gammafunctie zijn te berekenen met behulp van 8.6.3 of 8.6.8.

Voorbeelden.

$$2) \int_0^{\pi/2} \sin^7 \varphi \cos^3 \varphi \, d\varphi = \frac{1}{2} B(4, 2) = \frac{1}{2} \frac{\Gamma(4)\Gamma(2)}{\Gamma(6)} = \frac{1}{2} \frac{3!1!}{5!} = \frac{1}{40} .$$

$$3) \int_0^{\pi/2} \sin^8 \varphi \cos^4 \varphi \, d\varphi = \frac{1}{2} B(\frac{9}{2}, \frac{5}{2}) = \frac{1}{2} \frac{\Gamma(\frac{9}{2})\Gamma(\frac{5}{2})}{\Gamma(7)} =$$

$$= \frac{1}{2} \frac{8!\sqrt{\pi}}{2^8 4!} \frac{4!\sqrt{\pi}}{2^4 2!} \frac{1}{6!} = \frac{7\pi}{2048} .$$

4) Gevraagd het traagheidsmoment T t.o.v. de z -as van het homogene lichaam G , dat begrensd wordt door het vlak $z = x$ en de omwentelingsparaboloïde $z = x^2 + y^2$; de massadichtheid van G is 1. We gebruiken cylindercoördinaten, dan wordt het vlak gegeven door $z = r \cos \varphi$ en de paraboloïde door $z = r^2$. De projectie op het (x,y) -vlak van de snijkromme van vlak en paraboloïde wordt beschreven door $r = \cos \varphi, z = 0$.

$$T = \iiint_G (x^2 + y^2) \, dx \, dy \, dz = \iiint_G r^3 \, dr \, d\varphi \, dz = 2 \int_0^{\pi/2} d\varphi \int_0^{\cos \varphi} r^3 \, dr \int_{r^2}^{r \cos \varphi} dz =$$

$$= 2 \int_0^{\pi/2} d\varphi \int_0^{\cos \varphi} (r^4 \cos \varphi - r^5) \, dr = \frac{1}{15} \int_0^{\pi/2} \cos^6 \varphi \, d\varphi = \frac{\pi}{96} .$$

8.7. Integralen met een parameter

De functie

$$8.7.1. \quad F(a) := \int_0^{\infty} e^{-ax} \cos x \, dx, \quad (a > 0)$$

is voorbeeld van een functie die gedefinieerd is door middel van een integraal; de variabele a treedt op als parameter in de integraal. In 2.9.2, voorbeeld 1, is berekend: $F(a) = a/(a^2 + 1)$. We willen $F(a)$ differentiëren. Neem aan dat de differentiatie naar a in 8.7.1 "onder het integraalteken" mag worden uitgevoerd, dan vinden we

$$F'(a) = - \int_0^{\infty} x e^{-ax} \cos x \, dx$$

en dus

$$\int_0^{\infty} x e^{-ax} \cos x \, dx = - \frac{d}{da} \left(\frac{a}{a^2 + 1} \right) = \frac{a^2 - 1}{(a^2 + 1)^2}.$$

Men kan verifiëren dat het laatste resultaat correct is, zodat de verwisseling van differentiatie en integratie hier inderdaad geoorloofd is.

In deze paragraaf zullen we een algemene stelling afleiden voor het differentiëren onder het integraalteken van een integraal met een parameter.

8.7.2. Definitie. Zij $f(x,t)$ gedefinieerd op $(a,b) \times [c,d]$ met eventueel $a = -\infty$ en/of $b = \infty$, en zij $f(x,t)$ continu in x op (a,b) voor alle $t \in [c,d]$. Indien er een functie $g(x)$ gedefinieerd op (a,b) bestaat zodat $|f(x,t)| \leq g(x)$ voor alle $x \in (a,b)$, $t \in [c,d]$ en $\int_a^b g(x) dx$ convergent is, dan heet $\int_a^b f(x,t) dx$ gemajoreerd convergent voor $t \in [c,d]$. De functie $g(x)$ heet een majorant van $f(x,t)$; $g(x)$ majoreert $f(x,t)$.

De definitie 8.7.2 is vooral van betekenis in de volgende twee gevallen:

(i) $a = -\infty$ en/of $b = \infty$; (ii) a, b zijn eindig en $|f(x,t)| \rightarrow \infty$ als $x \uparrow a$ of $x \uparrow b$. In beide gevallen is $\int_a^b f(x,t) dx$ oneigenlijk. Uit 2.8.11, 2.8.12,

2.8.13 volgt dan: als $\int_a^b f(x,t)dx$ gemajoreerd convergent is voor $t \in [c,d]$, dan is de integraal absoluut convergent en dus convergent voor alle $t \in [c,d]$.

In geval a, b eindig zijn en $f(x,t)$ continu is op $[a,b] \times [c,d]$, dan is $f(x,t)$ daar begrensd (6.7.1). De integraal $\int_a^b f(x,t)dx$ is dan eigenlijk en zeker gemajoreerd convergent; als majorant van $f(x,t)$ kan men bijv. nemen $g(x) = M := \max\{|f(x,t)| \mid a \leq x \leq b, c \leq t \leq d\}$.

De volgende stelling, die we niet bewijzen, is fundamenteel.

8.7.3. Stelling (gemajoreerde-convergentiestelling). Zij $\int_a^b f(x,t)dx$ gemajoreerd convergent voor $t \in [c,d]$, zij $t_0 \in [c,d]$ en zij $\lim_{t \rightarrow t_0} f(x,t) = h(x)$ voor alle $x \in (a,b)$, waarbij $h(x)$ continu is op (a,b) .

Dan is $\int_a^b h(x)dx$ convergent en er geldt: $\lim_{t \rightarrow t_0} \int_a^b f(x,t)dx = \int_a^b h(x)dx$.

De stelling drukt uit dat men de limiet onder het integraalteken mag nemen (of: dat men limiet en integraal mag verwisselen), mits voldaan is aan de voorwaarde van gemajoreerde convergentie. De stelling geldt ook voor een limiet waarbij $t \rightarrow \infty$, aangenomen dat $\int_a^b f(x,t)dx$ gemajoreerd convergent is voor $t \in [c, \infty)$. Voor een integraal die in plaats van één parameter meerdere parameters bevat, kan men een stelling analoog aan 8.7.3 formuleren.

Indien de convergentie van $\int_a^b f(x,t)dx$ niet gemajoreerd is, is de verwisseling van limiet en integraal in het algemeen niet geoorloofd zoals blijkt uit het volgende voorbeeld.

Voorbeeld 1. Beschouw de functie $F(t) = \int_0^\infty te^{-tx} dx$, ($t \geq 0$).

Voor alle $t \geq 0$ is de integraal absoluut convergent. Voorts is $\lim_{t \rightarrow 0} te^{-tx} = 0$

voor alle $x \geq 0$. Er geldt echter niet $\lim_{t \rightarrow 0} F(t) = 0$, immers met behulp van

de substitutie $tx = y$ volgt eenvoudig $F(t) = \int_0^\infty e^{-y} dy = 1$ voor alle $t > 0$.

De convergentie van $\int_0^{\infty} te^{-tx} dx$ is niet gemajoreerd. Als er nl. een majorant $g(x)$ bestaat met $te^{-tx} \leq g(x)$ voor alle $x \geq 0, t \geq 0$, dan volgt na substitutie van $t = 1/x: g(x) \geq e^{-1}/x$ voor $x > 0$. De integraal $\int_0^{\infty} g(x)dx$ is dan echter divergent.

We geven een aantal toepassingen van 8.7.3.

8.7.4. Stelling. Zij $f(x,t)$ continu op $(a,b) \times [c,d]$ met eventueel $a = -\infty$ en/of $b = \infty$, en zij $\int_a^b f(x,t)dx$ gemajoreerd convergent voor $t \in [c,d]$, dan is $F(t) := \int_a^b f(x,t)dx$ continu op $[c,d]$.

8.7.5. Stelling. Zij $f(x,t)$ gedefinieerd op $(a,b) \times [c,d]$ met eventueel $a = -\infty$ en/of $b = \infty$, en laat $f(x,t)$ voldoen aan de volgende onderstellingen:

- (i) $f(x,t)$ is continu op $(a,b) \times [c,d]$;
- (ii) $f(x,t)$ is differentieerbaar naar t en $f_t(x,t)$ is continu in x op (a,b) voor alle $t \in (c,d)$;
- (iii) $F(t) := \int_a^b f(x,t)dx$ is convergent voor alle $t \in [c,d]$;
- (iv) $G(t) := \int_a^b f_t(x,t)dx$ is gemajoreerd convergent voor $t \in (c,d)$.

Dan is $F(t)$ differentieerbaar op (c,d) met afgeleide $F'(t) = G(t)$.

Bewijs. Zij $t_0 \in (c,d)$. Beschouw het quotiënt

$$\frac{F(t_0 + \tau) - F(t_0)}{\tau} = \int_a^b \frac{f(x, t_0 + \tau) - f(x, t_0)}{\tau} dx, \quad \begin{matrix} (\tau \neq 0, \\ c - t_0 \leq \tau \leq d - t_0) \end{matrix}$$

Voer in de functie $\varphi(x, \tau)$ gedefinieerd door

$$\varphi(x, \tau) := \begin{cases} \frac{f(x, t_0 + \tau) - f(x, t_0)}{\tau} & \text{als } \tau \neq 0, \tau \in [c - t_0, d - t_0], \\ f_t(x, t_0) & \text{als } \tau = 0, \end{cases}$$

dan is

$$\text{dan } \varphi(x, 0) = f_t(x, t_0)$$

$$\frac{F(t_0 + \tau) - F(t_0)}{\tau} = \int_a^b \varphi(x, \tau) dx, \quad G(t_0) = \int_a^b \varphi(x, 0) dx.$$

Uit de veronderstelling (iv) volgt dat er een majorant $g(x)$ bestaat zodanig dat $|f_t(x, t)| \leq g(x)$ voor alle $x \in (a, b)$, $t \in (c, d)$ en $\int_a^b g(x) dx$ convergent is. De functie $g(x)$ is tevens majorant van $\varphi(x, \tau)$. Op grond van de middelwaardestelling (2.6.7) bestaat er namelijk een getal $\theta \in (0, 1)$ zodat $\varphi(x, \tau) = f_t(x, t_0 + \theta\tau)$, waaruit volgt $|\varphi(x, \tau)| \leq g(x)$ voor alle $x \in (a, b)$, $\tau \in [c - t_0, d - t_0]$. De integraal $\int_a^b \varphi(x, \tau) dx$ is dus gemajoreerd convergent voor $\tau \in [c - t_0, d - t_0]$.

Uit de definitie van $\varphi(x, \tau)$ volgt dat $\lim_{\tau \rightarrow 0} \varphi(x, \tau) = \varphi(x, 0)$ voor alle $x \in (a, b)$, waarbij $\varphi(x, 0)$ continu is op (a, b) op grond van de onderstelling (ii). Maak nu gebruik van 8.7.3, dan is

$$\lim_{\tau \rightarrow 0} \int_a^b \varphi(x, \tau) dx = \int_a^b \varphi(x, 0) dx,$$

oftewel

$$\lim_{\tau \rightarrow 0} \frac{F(t_0 + \tau) - F(t_0)}{\tau} = G(t_0).$$

De functie F is dus differentieerbaar in t_0 met afgeleide $F'(t_0) = G(t_0)$. \square

De stelling 8.7.5 drukt uit dat men onder het integraalteken mag differentiëren, mits de gedifferentieerde integraal gemajoreerd convergent is. Beschouw nu nogmaals de integraal 8.7.1. Zij $a_0 > 0$, kies dan $\varepsilon > 0$ zodanig dat $a_0 - \varepsilon > 0$. Op het interval $(a_0 - \varepsilon, a_0 + \varepsilon)$ wordt de gedifferentieerde integraal $x e^{-ax} \cos x$ gemajoreerd door $x e^{-(a_0 - \varepsilon)x}$. Op grond van 8.7.5 is $F(a)$ dan differentieerbaar op $(a_0 - \varepsilon, a_0 + \varepsilon)$ en dus in a_0 , en de afgeleide $F'(a)$ wordt verkregen door differentiatie onder het integraalteken. Merk op: voor differentieerbaarheid van $F(a)$ in a_0 is het reeds voldoende dat de gedifferentieerde integraal gemajoreerd convergent is op een omgeving van a_0 .

Met gebruik van de opmerking die voorafgaat aan 8.7.3, vinden we het volgende resultaat als specialisatie van 8.7.4 en 8.7.5.

8.7.6. Stelling. Zij $f(x,t)$ continu op $[a,b] \times [c,d]$ met a en b eindig, dan is

$$F(t) := \int_a^b f(x,t) dx \text{ continu op } [c,d].$$

Als bovendien $f(x,t)$ differentieerbaar is naar t en $f_t(x,t)$ continu is op $[a,b] \times (c,d)$, dan is $F(t)$ differentieerbaar op (c,d) met afgeleide

$$F'(t) = \int_a^b f_t(x,t) dx.$$

In het geval van stelling 8.7.6 zijn $\int_a^b f(x,t) dx$, $\int_a^b f_t(x,t) dx$ eigenlijke integralen en de gemajoreerde convergentie hoeft dan niet afzonderlijk geverifieerd te worden. Een analoge stelling geldt voor een integraal die van meerdere parameters afhangt.

Voorbeelden. 2) Beschouw de integraal

$$F(t) := \int_0^{h(t)} f(x,t) dx, \quad (t \in [c,d])$$

waarbij $h(t): [c,d] \rightarrow [0,1]$ en $f(x,t)$ gedefinieerd is op $[0,1] \times [c,d]$. Zij $h(t)$ continu op $[c,d]$ en $f(x,t)$ continu op $[0,1] \times [c,d]$, dan is $F(t)$ continu op $[c,d]$. Immers, door middel van de substitutie $x = h(t)y$ volgt

$$F(t) = h(t) \int_0^1 f(h(t)y,t) dy,$$

waarna 8.7.6 kan worden toegepast.

Als bovendien $h(t)$ differentieerbaar is op (c,d) , $f(x,t)$ differentieerbaar is naar t en $f_t(x,t)$ continu is op $[0,1] \times (c,d)$, dan is $F(t)$ differentieerbaar op (c,d) . De afgeleide wordt verkregen met behulp van de kettingregel.

Schrijf daartoe $F(t) = H(h(t),t)$, waarbij $H(s,t) := \int_0^s f(x,t) dt$ een integraal met twee parameters is. Onder toepassing van 2.8.4, 8.7.6 volgt dan

$$F'(t) = H_s(h(t), t)h'(t) + H_t(h(t), t) = f(h(t), t)h'(t) + \int_0^{h(t)} f_t(x, t) dx ,$$

geldig voor alle $t \in (c, d)$.

Evenzo geldt

$$\frac{d}{dt} \int_{g(t)}^{h(t)} f(x, t) dx = f(h(t), t)h'(t) - f(g(t), t)g'(t) + \int_{g(t)}^{h(t)} f_t(x, t) dx$$

onder analoge onderstellingen als boven.

- 3) Beschouw de gedwongen ongedempte trilling onder invloed van een uitwendige kracht, beschreven door de vergelijking (zie 5.5)

$$\ddot{x}(t) + \omega_0^2 x(t) = \omega_0 f(t) .$$

Laat de beginvoorwaarden zijn: $x(0) = \dot{x}(0) = 0$.

We zullen verifiëren dat de oplossing wordt gegeven door

$$x(t) = \int_0^t f(\tau) \sin \omega_0(t - \tau) d\tau .$$

Bepaal daartoe $\dot{x}(t)$, $\ddot{x}(t)$ door differentiatie onder het integraalteken:

$$\dot{x}(t) = f(t) \sin \omega_0(t - t) + \omega_0 \int_0^t f(\tau) \cos \omega_0(t - \tau) d\tau =$$

$$= \omega_0 \int_0^t f(\tau) \cos \omega_0(t - \tau) d\tau ,$$

$$\ddot{x}(t) = \omega_0 f(t) \cos \omega_0(t - t) - \omega_0^2 \int_0^t f(\tau) \sin \omega_0(t - \tau) d\tau =$$

$$= \omega_0 f(t) - \omega_0^2 \int_0^t f(\tau) \sin \omega_0(t - \tau) d\tau .$$

Uit het laatste resultaat volgt: $\ddot{x}(t) = \omega_0 f(t) - \omega_0^2 x(t)$, zodat $x(t)$ inderdaad aan de gegeven differentiaalvergelijking voldoet. Voorts is

$$x(0) = \int_0^0 f(\tau) \sin \omega_0(0 - \tau) d\tau = 0$$

en evenzo $\dot{x}(0) = 0$, zodat $x(t)$ ook aan de beginvoorwaarden voldoet.

- 4) We beschouwen het volgende probleem uit de statica. Een staaf met lengte ℓ is geplaatst langs het interval $[0, \ell]$ van de x -as. Het uiteinde $x = 0$ van de staaf is ingeklemd en de staaf ondervindt een dwarsbelasting $q(x)$.

De dwarskracht Q en het moment M in een punt x zijn dan achtereenvolgens

$$Q(x) = \int_x^\ell q(y) dy, \quad M(x) = \int_x^\ell (y - x)q(y) dy.$$

Differentieer $Q(x)$, $M(x)$, dan volgt

$$Q'(x) = -q(x), \quad M'(x) = (x - x)q(x) - \int_x^\ell q(y) dy = - \int_x^\ell q(y) dy$$

en dus $M''(x) = q(x)$.

- 5) In 8.6 is aangetoond dat de integraal

$$\Gamma(x) := \int_0^\infty e^{-t} t^{x-1} dt$$

gemajoreerd convergent is voor $x \in [a, b]$, waarbij $0 < a < b < \infty$. Met behulp van 8.7.4 volgt dan dat $\Gamma(x)$ continu is op $[a, b]$. Omdat a en b willekeurig zijn (afgezien van de voorwaarde $0 < a < b < \infty$), is $\Gamma(x)$ dus continu op $(0, \infty)$. Evenzo kan men aantonen dat $B(p, q)$ continu is voor $p > 0$, $q > 0$. De gammafunctie is ook differentieerbaar op $(0, \infty)$ en de afgeleide wordt verkregen door differentiatie onder het integraalteken:

$$\Gamma'(x) = \int_0^\infty e^{-t} t^{x-1} \log t dt, \quad (x > 0).$$

Ga daartoe na dat de gedifferentieerde integraal gemajoreerd convergent is voor $x \in [a, b]$, waarbij $0 < a < b < \infty$. Vervolgens kan 8.7.5 worden toegepast.