Hoofdstuk 2

Lineaire afbeeldingen

2.1 Inleiding

Een *afbeelding* f van een verzameling V naar een verzameling W is een regel die aan ieder element v van V een element $f(v)$ van W toevoegt m.a.w. een generalisatie van het bekende functie-begrip. Als ook $g : W \to Z$ een afbeelding is, dan is de *samenstelling* van f en g de afbeelding, genoteerd $g \circ f$, van V naar Z gedefinieerd d.m.v.

$$(g \circ f)(v) := g(f(v)) \text{ voor alle } v \text{ in } V.$$

Afbeeldingen kunnen voor verschillende doeleinden gebruikt worden. Ze kunnen als object opzich bestudeerd worden, zoals in de quantummechanica waar iedere meetbare fysische grootheid door een welbepaalde lineaire afbeelding wordt voorgesteld. Het bestuderen van zo'n afbeelding, bijvoorbeeld of hij diagonaliseerbaar is, geeft dan informatie over de bijbehorende fysische grootheid. Ook kunnen afbeeldingen als dynamische systemen worden opgevat d.w.z. je onderzoekt bij een gegeven afbeelding f en een gegeven punt v_0 het gedrag van $f^n(v_0)$ als n groot wordt.

Wat we in dit hoofdstuk zullen doen is afbeeldingen gebruiken om verschillende vectorruimten met elkaar te vergelijken. Net zoals we in de verzamelingstheorie kunnen laten zien dat de verzamelingen $\mathbb{N}$, $\mathbb{Z}$ en $\mathbb{Q}$ "in wezen hetzelfde zijn", d.w.z. er bestaat een bijectie (zie definitie hieronder) tussen al deze drie verzamelingen, zo kunnen we ook laten zien dat veel van de ogenschijnlijk verschillende vectorruimten die we tot nu toe gezien hebben "in wezen hetzelfde zijn".

Om vectorruimten zinvol te kunnen vergelijken, moeten we niet alleen kijken of ze als verzamelingen "hetzelfde" zijn, maar moeten we ook rekening houden met de optelstructuur en de scalairvermenigvuldiging. Afbeeldingen die met deze twee extra structuren rekening houden zijn de zgn. lineaire afbeeldingen (definitie 2.2.1). Voordat we deze definitie geven, willen we eerst nog wat bekende begrippen uit de verzamelingstheorie in herinnering roepen.

Definitie 2.1.1 Zij $f : V \to W$ een willekeurige afbeelding tussen twee verzamelingen. Als $v \in V$, dan heet $w := f(v)$ het *beeld van v onder f* en v heet *een origineel van w onder f*.

Een beeld kan meerdere originelen hebben: bijvoorbeeld als $f(x) = x^2$ van $\mathbb{R}$ naar $\mathbb{R}$, dan zien we dat 2 en -2 originelen van 4 zijn. Aan de andere kant hoeft niet iedere $w \in W$ een origineel in V te hebben; zo heeft bij de afbeelding $f(x) = x^2$ van $\mathbb{R}$ naar $\mathbb{R}$ het getal -1 geen originelen in $\mathbb{R}$ want $x^2 \geq 0$ voor alle x in $\mathbb{R}$.

Definitie 2.1.2 Zij $f : V \to W$ een afbeelding tussen twee verzamelingen.

i) Als iedere $w \in W$ ten hoogste één origineel in V heeft, heet f *één-éénduidig* of *injectief* m.a.w. als $f(v_1) = f(v_2)$ dan moet $v_1 = v_2$.

ii) Als iedere $w \in W$ tenminste één origineel in V heeft, heet f *surjectief* of *op*.

iii) Een afbeelding die zowel injectief als surjectief is, heet *bijectief*. M.a.w. een afbeelding $f : V \to W$ is bijectief (ook wel een *bijectie* genoemd) als er voor iedere $w \in W$ precies één origineel $v \in V$ bestaat. De punten uit V corresponderen dan één-éénduidig met de punten van W.

2.2 Definitie en fundamentele eigenschappen

In deze paragraaf zijn V en W steeds vectorruimten.

Definitie 2.2.1 Een afbeelding $f : V \to W$ heet een *lineaire afbeelding* als voor alle v_1, v_2, v in V en alle c in $\mathbb{R}$ geldt:

i) $f(v_1 + v_2) = f(v_1) + f(v_2)$

ii) $f(cv) = cf(v)$.

Het idee achter deze definitie is dat een lineaire afbeelding een afbeelding is die de structuur van de vectorruimten bewaart, d.w.z. die rekening met de bewerkingen houdt. Het maakt namelijk volgens deze definitie niets uit of we twee vectoren eerst bij elkaar optellen en vervolgens afbeelden, of eerst de vectoren apart afbeelden en dan de beelden bij elkaar optellen. Hetzelfde geldt voor de scalairvermenigvuldiging.

Voorbeeld 2.2.2 Zij A een $m \times n$ matrix. Definieer $f_A : \mathbb{R}^n \to \mathbb{R}^m$ d.m.v. $f_A(x) := Ax$. Laat zien dat f_A een lineaire afbeelding is.

Oplossing: Zij x en y in $\mathbb{R}^n$. Dan geldt volgens stelling 2.2.13, LA1, dat $f_A(x + y) = A(x + y) = Ax + Ay = f_A(x) + f_A(y)$. Ook zie je meteen dat $A(cx) = cAx$ en dus $f_A(cx) = cf_A(x)$. Dus f_A is lineair.

Voorbeeld 2.2.3 Zij V de vectorruimte bestaande uit alle reëel-waardige differentieerbare functies op $\mathbb{R}$ en W de vectorruimte van alle reëel-waardige functies op $\mathbb{R}$. Zij $D : V \to W$ de afbeelding gedefinieerd d.m.v. $D(f) := f'$ (de afgeleide van f). Dan volgt uit de bekende regels voor differentiëren dat D een lineaire afbeelding is, immers

$$D(f + g) = (f + g)' = f' + g' = Df + Dg \quad \text{en} \quad D(cf) = (cf)' = cf' = cDf.$$

Voorbeeld 2.2.4 Zij V de vectorruimte der reëel-waardige continue functies op $[0, 1]$. Dan is de afbeelding $I : V \to \mathbb{R}$ gedefinieerd d.m.v.

$$I(f) := \int_0^1 f(x)\, dx$$

een lineaire afbeelding, immers

$$I(f + g) = \int_0^1 (f + g)(x)\, dx = \int_0^1 (f(x) + g(x))\, dx = \int_0^1 f(x)\, dx + \int_0^1 g(x)\, dx = I(f) + I(g).$$

Verder

$$I(cf) = \int_0^1 (cf)(x)\,dx = \int_0^1 cf(x)\,dx = c\int_0^1 f(x)\,dx = cI(f).$$

Voorbeeld 2.2.5 De afbeelding $d : M_2(\mathbb{R}) \to \mathbb{R}$ gedefinieerd d.m.v. $d(A) := \det A$ is geen lineaire afbeelding. Immers $\det(3I_2) = 9$ terwijl $3.\det I_2 = 3 \neq 9$.

Opgave 2.2.6 Zij V de vectorruimte der reëel-waardige functies op $\mathbb{R}$. Laat zien dat de afbeelding $\delta : V \to R$ gedefinieerd d.m.v. $\delta(f) := f(0)$ een lineaire afbeelding.

Opgave 2.2.7 Laat zien dat de volgende afbeeldingen lineair zijn.
i) De *nulafbeelding* $O : V \to W$ gedefinieerd d.m.v. $O(v) = 0$ voor alle $v \in V$.
ii) De *identiteit* $1_V : V \to V$ gedefinieerd d.m.v. $1_V(v) = v$ voor alle $v \in V$.

Stelling 2.2.8 Zij $f : V \to W$ een lineaire afbeelding. Dan geldt
i) $f(0_V) = 0_W$ waarbij 0_V (resp. 0_W) de nulvector in V (resp. W) is.
ii) $f(-v) = -f(v)$, voor alle v in V.
iii) $f(c_1v_1 + \cdots + c_nv_n) = c_1f(v_1) + \cdots + c_nf(v_n)$ voor alle $n \geq 1$, alle v_j in V en alle c_j in $\mathbb{R}$.

Bewijs. i) Volgens 1.2.17 i) geldt $0_V = 0.0_V$. Dus $f(0_V) = 0.f(0_V) = 0_W$, weer vanwege 1.2.17 i).
ii) Volgens 1.2.17 ii) geldt: $-v = (-1)v$. Dus $f(-v) = f((-1)v) = (-1)f(v) = -f(v)$, vanwege 1.2.17 ii).
iii) Bewijs dit zelf met inductie naar n. □

We gaan nu, zoals aangekondigd, vectorruimten met elkaar vergelijken.

Definitie 2.2.9 Een lineaire afbeelding $f : V \to W$ heet een *isomorfisme* als f een bijectie is. We zeggen dat twee vectorruimten V en W *isomorf* zijn als er een isomorfisme $f : V \to W$ bestaat. Notatie: $V \cong W$.

Isomorfie (= gelijke structuur) van twee vectorruimten V en W betekent dat beide ruimten dezelfde vectorruimte eigenschappen hebben m.a.w. wat voor V geldt, geldt ook voor W en omgekeerd. De volgende stelling is hiervan een illustratie.

Stelling 2.2.10 Zij $f : V \to W$ een isomorfisme. Dan geldt
1) Als $(v_1, \ldots, v_n)$ een onafhankelijk stelsel in V is, dan is $(f(v_1), \ldots, f(v_n))$ een onafhankelijk stelsel in W.
2) Als $(v_1, \ldots, v_n)$ een volledig stelsel voor V is, dan is $(f(v_1), \ldots, f(v_n))$ een volledig stelsel voor W.
3) Als $(v_1, \ldots, v_n)$ een basis voor V is, dan is $(f(v_1), \ldots, f(v_n))$ een basis voor W.
4) $\dim V = \dim W$.

Bewijs. 1) Stel $c_1f(v_1)+\cdots+c_nf(v_n) = 0$. Dan volgt uit 2.2.8 iii) dat $f(c_1v_1+\cdots+c_nv_n) = 0 = f(0)$ (volgens 2.2.8 i)). Uit de injectiviteit van f volgt dan $c_1v_1 + \cdots + c_nv_n = 0$ en de onafhankelijkheid van $(v_1, \ldots, v_n)$ geeft dan $c_1 = \cdots = c_n = 0$. Dus is $(f(v_1), \ldots, f(v_n))$ onafhankelijk.
2) Zij $w \in W$. We moeten inzien dat er $c_1, \ldots, c_n$ in $\mathbb{R}$ bestaan met $w = c_1f(v_1)+\cdots+c_nf(v_n)$. We weten dat er een $v \in V$ bestaat met $w = f(v)$ (want f is i.h.b. surjectief). Omdat

$(v_1, \dots, v_n)$ een volledig stelsel voor V is, bestaan er $c_1, \dots, c_n$ in $\mathbb{R}$ met $v = c_1v_1 + \cdots + c_nv_n$. Dus volgt m.b.v. 2.2.8 iii) dat $w = f(v) = f(c_1v_1 + \cdots + c_nv_n) = c_1f(v_1) + \cdots + c_nf(v_n)$, dus is w inderdaad een lineaire combinatie van $f(v_1), \dots, f(v_n)$.
Eigenschap 3) volgt onmiddellijk uit 1) en 2) en eigenschap 4) volgt meteen uit 3). $\square$

Stelling 2.2.10 zegt i.h.b. dat isomorfe vectorruimte dezelfde dimensie hebben. We tonen nu (voor eindig dimensionale reële vectorruimten) de omkering aan.

Stelling 2.2.11 (Classificatie stelling) Zij V een n-dimensionale vectorruimte. Dan geldt $V \cong \mathbb{R}^n$.

Bewijs. Zij $(v_1, \dots, v_n)$ een basis voor V. Definieer $f : \mathbb{R}^n \to V$ d.m.v.

$$f\left(\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}\right) := x_1v_1 + \cdots + x_nv_n.$$

Men gaat eenvoudig na dat f lineair is. De basisvectoren v_i worden als beelden van de standaard basis $(e_1, \dots, e_n)$ van $\mathbb{R}^n$ (zie 1.3.15) verkregen en omdat $(v_1, \dots, v_n)$ een volledig stelsel voor V is (het is immers een basis voor V), is f surjectief.

Dat f injectief is volgt uit het feit dat iedere vector een *eenduidige* lineaire combinatie van een basis is (lemma 1.3.16). Dus f is lineair, surjectief en injectief, dus een isomorfisme. Dus $V \cong \mathbb{R}^n$. $\square$

Opmerking 2.2.12 Voor een vector $v = x_1v_1 + \cdots + x_nv_n \in V$ noemt men $\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ ook de *coördinaatvector* van v (met betrekking tot de basis $(v_1, \dots, v_n)$ van V).

Voorbeeld 2.2.13 Geef een isomorfisme tussen $\mathbb{R}^4$ en $M_2(\mathbb{R})$ aan.

Oplossing. Volgens 1.3.14 vormt (v_1, v_2, v_3, v_4) gedefinieerd in 1.3.4 een basis van $M_2(\mathbb{R})$. Dus is volgens 2.2.11 de afbeelding $f : \mathbb{R}^4 \to M_2(\mathbb{R})$ gedefinieerd d.m.v. $f\left(\begin{pmatrix} x_1 \\ \vdots \\ x_4 \end{pmatrix}\right) = x_1v_1 + \cdots + x_nv_4$ een isomorfisme tussen $\mathbb{R}^4$ en $M_2(\mathbb{R})$. De waarden der v_j's invullend zien we dat $f\left(\begin{pmatrix} x_1 \\ \vdots \\ x_4 \end{pmatrix}\right) = \begin{pmatrix} x_1 & x_2 \\ x_3 & x_4 \end{pmatrix}$ een isomorfisme tussen $\mathbb{R}^4$ en $M_2(\mathbb{R})$ is.

Ook zonder gebruik te maken van 2.2.11 kan men natuurlijk rechtstreeks inzien dat deze f een isomorfisme is.

Opgave 2.2.14 Bewijs dat de vectorruimte F gedefinieerd in 1.2.19 isomorf is met $\mathbb{R}^2$ en geef een isomorfisme expliciet aan.

Tot slot van deze paragraaf geven we nog een stelling die vooral van theoretisch belang is. Hiervoor hebben we twee nieuwe begrippen nodig.

Definitie 2.2.15 Zij $f : V \to W$ een lineaire afbeelding. Dan definiëren we twee lineaire deelruimten van V resp. W door:

$$\begin{aligned} \ker f &:= \{v \in V \mid f(v) = 0\} \text{ heet de } \textit{kern} \text{ van } f; \\ \operatorname{Im} f &:= \{f(v) \mid v \in V\} = f(V) \text{ heet het } \textit{beeld} \text{ van } f. \end{aligned}$$

Opgave 2.2.16 Bewijs dat $\ker f$ en $\operatorname{Im} f$ inderdaad lineaire deelruimten zijn van V resp. W.

Opgave 2.2.17 Zij $f : V \to W$ een lineaire afbeelding. Bewijs dat
i) f is surjectief d.e.s.d.a. $\operatorname{Im} f = W$.
ii) f is injectief d.e.s.d.a. $\ker f = \{0\}$.
[Aanwijzing: gebruik bij ii) dat $f(0) = 0$. Dus als bijvoorbeeld $f(v) = 0$ en f is injectief dan volgt uit $f(v) = 0 = f(0)$ dat $v = 0$.]

Stelling 2.2.18 Zij $f : V \to W$ een lineaire afbeelding. Neem aan dat $\ker f$ en $\operatorname{Im} f$ eindig dimensionaal zijn. Dan is ook V eindig dimensionaal en er geldt

$$\dim V = \dim \ker f + \dim \operatorname{Im} f.$$

Bewijs. Omdat in $\operatorname{Im} f$ iedere vector van de vorm $f(v)$ is en $\operatorname{Im} f$ eindig dimensionaal is, bestaan er $v_1, \dots, v_r$ in V zodanig dat $(f(v_1), \dots, f(v_r))$ een basis van $\operatorname{Im} f$ is. Zij nu $(v'_1, \dots, v'_d)$ een basis van $\ker f$. We zullen nu gaan bewijzen

$$(v_1, \dots v_r, v'_1, \dots, v'_d) \text{ is een basis van } V. \tag{2.1}$$

Uit (2.1) volgt dan $\dim V = r + d = \dim \operatorname{Im} f + \dim \ker f$.
Om (2.1) te bewijzen moeten we de volledigheid en de onafhankelijkheid van het stelsel $(v_1, \dots, v_r, v'_1, \dots, v'_d)$ aantonen.
1) Volledigheid: zij $v \in V$. Dan $f(v) \in \operatorname{Im} f$, dus bestaan er $c_1, \dots, c_r$ in $\mathbb{R}$ met $f(v) = c_1 f(v_1) + \cdots + c_r f(v_r) = f(c_1 v_1 + \cdots + c_r v_r)$. Dan is $0 = f(v) - f(c_1 v_1 + \cdots + c_r v_r) = f(v - (c_1 v_1 + \cdots + c_r v_r))$ m.a.w. $v' := v - (c_1 v_1 + \cdots + c_r v_r) \in \ker f$. Omdat $(v'_1, \dots, v'_d)$ een basis van $\ker f$ is, bestaan er $c'_1, \dots, c'_d \in \mathbb{R}$ met $v' = c'_1 v'_1 + \cdots + c'_d v'_d$. Dus is $v = c_1 v_1 + \cdots + c_r v_r + v' = c_1 v_1 + \cdots + c_r v_r + c'_1 v'_1 + \cdots + c'_d v'_d$ en dus is $(v_1, \dots, v_r, v'_1, \dots, v'_d)$ een volledig stelsel voor V.
2) Onafhankelijkheid: stel

$$c_1 v_1 + \cdots + c_r v_r + c'_1 v'_1 + \cdots + c'_d v'_d = 0 \tag{2.2}$$

Laat f werken. Uit $f(0) = 0$ en 2.2.8 iii) volgt dan $c_1 f(v_1) + \cdots + c_r f(v_r) + c'_1 f(v'_1) + \cdots + c'_d f(v'_d) = 0$. Omdat iedere $v'_i \in \ker f$ m.a.w. $f(v'_i) = 0$ volgt $c_1 f(v_1) + \cdots + c_r f(v_r) = 0$. Omdat $(f(v_1), \dots, f(v_r))$ een basis voor $\operatorname{Im} f$ is, dus i.h.b. onafhankelijk, volgt $c_1 = \cdots = c_r = 0$. Uit (2.2) volgt dan $c'_1 v'_1 + \cdots + c'_d v'_d = 0$. Echter $(v'_1, \dots, v'_d)$ is een basis voor $\ker f$, dus onafhankelijk en dus geldt ook $c'_1 = \cdots = c'_d = 0$. Dus totaal zien we: uit (2.2) volgt $c_1 = \cdots = c_r = c'_1 = \cdots = c'_d = 0$ m.a.w. $(v_1, \dots, v_r, v'_1, \dots, v'_d)$ is een onafhankelijk stelsel in V. □

Gevolg 2.2.19 Zij $f : V \to W$ een lineaire afbeelding en zij $\dim V = \dim W$ eindig. Dan geldt f is injectief d.e.s.d.a. f is surjectief d.e.s.d.a. f is een isomorfisme.

Bewijs. Als f injectief is dan is (vanwege 2.2.17) $\dim\ker f = 0$ en dus (vanwege 2.2.18) $\dim\operatorname{Im} f = \dim V = \dim W$. Dus (vanwege 1.4.3) $\operatorname{Im} f = W$, m.a.w. f is surjectief en dus een isomorfisme. Teruglezend zien we dat als f surjectief is, hij ook injectief en dus een isomorfisme is. □

Een isomorfisme van V naar zichzelf heet ook wel een *inverteerbare* (lineaire) afbeelding. De reden hiervoor ligt in het volgende eenvoudige feit.

Lemma 2.2.20 Zij $f : V \to V$ een lineaire afbeelding. Bewijs dat de volgende twee uitspraken gelijkwaardig zijn:

i) f is bijectief.

ii) Er bestaat een lineaire afbeelding $g : V \to V$ met $f \circ g = g \circ f = 1_V$, d.w.z. f is inverteerbaar.

Merk op dat we niet veronderstellen dat V eindig dimensionaal is.

Bewijs. i) $\Rightarrow$ ii): We kunnen de gezochte afbeelding g expliciet construeren: Omdat f surjectief is, laat zich iedere $w \in V$ schrijven als $w = f(v)$ en omdat f injectief is, is deze v eenduidig. We definiëren nu $g(w) := v$ voor deze v, dan is g een welgevormde afbeelding van V naar V. Volgens de definitie van g geldt $f(g(w)) = f(v) = w$ en $g(f(v)) = g(w) = v$, dus is inderdaad $f \circ g = g \circ f = 1_V$.

We moeten nog laten zien dat g lineair is. Zij $w_1, w_2 \in V$ en $v_1, v_2 \in V$ met $f(v_1) = w_1$ en $f(v_2) = w_2$. Wegens de lineariteit van f is dan $f(v_1 + v_2) = w_1 + w_2$. Volgens onze definitie van g is dan $g(w_1) = v_1$, $g(w_2) = v_2$ en $g(w_1 + w_2) = v_1 + v_2$, dus is inderdaad $g(w_1 + w_2) = g(w_1) + g(w_2)$. De redenering voor de scalairvermenigvuldiging werkt analoog (en wordt aan de lezer overgelaten).

ii) $\Rightarrow$ i): We moeten laten zien dat f injectief en surjectief is. Stel dat f niet injectief is, dan bestaan $v_1 \neq v_2$ in V met $f(v_1) = f(v_2) = w$. Maar wegens $g \circ f = 1_V$ is dan $g(f(v_1)) = g(w) = v_1$ en $g(f(v_2)) = g(w) = v_2$ en dus $v_1 = v_2$, tegenspraak.

Stel nu dat f niet surjectief is en zij $w \in V$ met $w \notin \operatorname{Im} f$. Maar wegens $f \circ g = 1_V$ is $f(g(w)) = w$, dus is w het beeld van $g(w)$ onder f, tegenspraak. □

Verder volgt uit 2.2.18 en 2.2.19:

Gevolg 2.2.21 Zij V een eindig dimensionale vectorruimte en $f : V \to V$ een lineaire afbeelding. Dan zijn gelijkwaardig:

i) f is bijectief (of inverteerbaar).

ii) f is injectief.

iii) als $f(v) = 0$, dan $v = 0$.

iv) f is surjectief

Bewijs. i) $\Rightarrow$ ii) geldt wegens de definitie van bijectiviteit en ii) $\Rightarrow$ iii) wegens de definitie van injectiviteit. Verder zegt iii) dat $\dim\ker f = 0$, dus is volgens 2.2.18 $\dim\operatorname{Im} f = \dim V$ en dus f surjectief. Tenslotte hadden we iv) $\Rightarrow$ i) al in 2.2.19 ingezien. □

De volgende opgave laat zien dat 2.2.21 niet waar is als V oneindig dimensionaal is.

Opgave 2.2.22 Zij V de verzameling der oneindige rijtjes $(a_1, a_2, \ldots)$ van reële getallen.

i) Laat zien dat V een vectorruimte is.

ii) Laat zien dat de afbeelding $l : V \to V$ gedefinieerd d.m.v.

$$l((a_1, a_2, \ldots)) = (0, a_1, a_2, \ldots)$$

een lineaire afbeelding is.

iii) Laat zien dat l injectief is.

iv) Laat zien dat l niet surjectief is.

v) Leid uit iii) en iv) af dat V oneindig dimensionaal is.

2.3 Lineaire afbeeldingen en matrices

Zij A een $m \times n$ matrix. De afbeelding $f_A : \mathbb{R}^n \to \mathbb{R}^m$ gedefinieerd d.m.v. $f_A(x) := Ax$ is volgens 2.2.2 een lineaire afbeelding. We kunnen nu verschillende uitspraken over het stelsel $Ax = b$ beschrijven in termen van de afbeelding f_A: zo zegt de surjectiviteit van f_A dat er bij iedere $b \in \mathbb{R}^m$ tenminste één $x \in \mathbb{R}^n$ bestaat met $Ax = b$, terwijl de injectiviteit van f_A zegt dat er bij iedere $b \in \mathbb{R}^m$ ten hoogste één $x \in \mathbb{R}^n$ bestaat met $Ax = b$. Verder is de oplossingsruimte van het stelsel $Ax = 0$ d.w.z. de verzameling van alle $x \in \mathbb{R}^n$ met $Ax = 0$, welke we in 1.2.14 $\text{Nul}(A)$ genoemd hebben en vaak als $N(A)$ schrijven, precies de kern van f_A m.a.w.

$$N(A) = \ker f_A. \tag{2.3}$$

Het beeld van f_A bestaat uit de verzameling van alle vectoren $f_A(x)$ d.w.z. de verzameling van alle Ax met $x \in \mathbb{R}^n$. Als we met $A_1, \ldots, A_n$ de kolommen van A aanduiden dan volgt uit de bekende formule

$$Ax = x_1 A_1 + \cdots + x_n A_n \tag{2.4}$$

dat $\operatorname{Im} f_A = \{x_1 A_1 + \cdots + x_n A_n \mid x_1, \ldots, x_n \in \mathbb{R}\}$ m.a.w. $\operatorname{Im} f_A$ is de verzameling van alle lineaire combinaties van de kolommen van A, of anders gezegd $\operatorname{Im} f_A$ is het opspansel $\langle A_1, \ldots, A_n \rangle$ van de kolommen van A. Deze ruimte heet de *kolommenruimte* van A, genoteerd $\text{Col}(A)$. Dus

$$\operatorname{Im} f_A = \text{Col}(A). \tag{2.5}$$

Omdat volgens 2.2.17 $\operatorname{Im} f_A$ een lineaire deelruimte van $\mathbb{R}^m$ is, is dus $\text{Col}(A)$ een vectorruimte die een dimensie $\leq m$ heeft.

Definitie 2.3.1 Zij A een $m \times n$ matrix, dan definieert men

$$\text{rang}\, A := \dim \text{Col}(A).$$

Stelling 2.3.2 Zij A een $m \times n$ matrix. Dan geldt $\dim N(A) = n - \text{rang}\, A$.

Bewijs. We passen 2.2.18 toe op de lineaire afbeelding $f_A : \mathbb{R}^n \to \mathbb{R}^m$. We vinden dan: $n = \dim \ker f_A + \dim \operatorname{Im} f_A$. Uit (2.5) en 2.3.1 volgt $\dim \operatorname{Im} f_A = \text{rang}\, A$ en uit (2.3) dat $\dim \ker f_A = \dim N(A)$. Dus $n = \dim N(A) + \text{rang}\, A$, waaruit de stelling volgt. □

Gevolg 2.3.3 Zij A een $m \times n$ matrix. Dan verandert de rang van A niet bij elementaire rij- en elementaire kolomoperaties.

Bewijs. Omdat $\operatorname{rang} A = \dim \operatorname{Im} f_A = \dim\langle A_1, \ldots, A_n\rangle$ verandert de rang van A niet bij elementaire kolomoperaties. Dit is duidelijk voor het verruilen van twee kolommen en voor het vermenigvuldigen van een kolom met een factor $\neq 0$. Maar iedere lineaire combinatie $v = x_1A_1 + \cdots + x_nA_n$ is ook een lineaire combinatie van $A_1, \ldots, A_{i-1}, A_i + cA_j, A_{i+1}, \ldots, A_n$ en omgekeerd, want $v = x_1A_1 + \cdots + x_i(A_i + cA_j) + \cdots + (x_j - cx_i)A_j + \cdots + x_nA_n$.

Aan de andere kant verandert $N(A) = \ker f_A$ niet bij elementaire rijtransformaties, dus verandert i.h.b. $\dim N(A)$ niet en dus ook $\operatorname{rang} A = n - \dim N(A)$ niet. $\square$

Stelling 2.3.4 Zij A een $m \times n$ matrix, dan geldt $\operatorname{rang} A = \operatorname{rang} A^t$.

Bewijs. Dit bewijzen we met inductie over het aantal kolommen van A. Zij $n = 1$. Voor $A = 0$ is $\operatorname{rang} A = 0$ en ook $\operatorname{rang} A^t = 0$. Voor $A \neq 0$ is $\operatorname{rang} A = 1$ en omdat minstens een van de kolommen van A^t niet nul is, is ook $\operatorname{rang} A^t = 1$.

Stel nu dat de bewering bewezen is voor $k < n$ en zij A een $m \times n$ matrix. We noteren met B de deelmatrix van A die de kolommen 2 t/m n bevat. Als de eerste kolom van A de nulvector is, geldt

$$A = \left(\begin{array}{c|c} 0 & \\ \vdots & B \\ 0 & \end{array}\right) \quad \text{en} \quad A^t = \left(\begin{array}{c} 0 \ \cdots \ 0 \\ \hline \\ B^t \\ \\ \end{array}\right)$$

en het is duidelijk dat $\operatorname{rang} A = \operatorname{rang} B$. Volgens de inductieaanname is dan $\operatorname{rang} B = \operatorname{rang} B^t$. Maar B^t verschilt van A^t alleen maar door aan iedere kolom van B^t een eerste component 0 toe te voegen. Omdat dit niets aan de dimensie van de kolommenruimte verandert, geldt $\operatorname{rang} B^t = \operatorname{rang} A^t$.

Stel nu dat de eerste kolom van A niet nul is. Dan kunnen we d.m.v. elementaire rijoperaties de eerste kolom tot $e_1 = (1\ 0\ \ldots\ 0)^t$ vegen en vervolgens de eerste rij door elementaire kolomoperaties tot $(1\ 0\ \ldots\ 0)$. Omdat volgens 2.3.3 deze operaties de rang van A niet veranderen, hebben we nu een matrix A' met $\operatorname{rang} A' = \operatorname{rang} A$. Maar rijoperaties toegepast op A zijn hetzelfde als kolomoperaties toegepast op A^t en andersom, dus geldt ook $\operatorname{rang}(A')^t = \operatorname{rang} A^t$. Nu is dus

$$A' = \left(\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & B & \\ 0 & & & \end{array}\right) \quad \text{en} \quad (A')^t = \left(\begin{array}{c|ccc} 1 & 0 & \cdots & 0 \\ \hline 0 & & & \\ \vdots & & B^t & \\ 0 & & & \end{array}\right)$$

en het is duidelijk dat $\operatorname{rang} A' = \operatorname{rang} B + 1$ en $\operatorname{rang}(A')^t = \operatorname{rang} B^t + 1$. Maar volgens de inductieaanname is $\operatorname{rang} B = \operatorname{rang} B^t$ en dus ook $\operatorname{rang} A' = \operatorname{rang}(A')^t$ en dus $\operatorname{rang} A = \operatorname{rang} A^t$. $\square$

Hoe nauw de samenhang tussen lineaire afbeeldingen en matrices is, zegt de volgende stelling.

Stelling 2.3.5 Zij $f : \mathbb{R}^n \to \mathbb{R}^m$ een lineaire afbeelding. Dan geldt:

i) Er bestaat precies één $m \times n$ matrix A met $f = f_A$.

ii) Zij $g : \mathbb{R}^m \to \mathbb{R}^l$ een tweede lineaire afbeelding en zij B de $l \times m$ matrix met $g = f_B$. Voor de samenstelling $g \circ f : \mathbb{R}^n \to \mathbb{R}^l$ geldt dan $g \circ f = f_{BA}$, d.w.z. het matrix product weerspiegelt de samenstelling van de lineaire afbeeldingen.

iii) Als $n = m$ is, geldt: f is inverteerbaar d.e.s.d.a. A inverteerbaar is.

Bewijs. i) Zij $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$. Dan $x = x_1e_1 + \cdots + x_ne_n$, waarbij e_i de i-de standaard basis vector in $\mathbb{R}^n$ is. Volgens 2.2.8 geldt

$$f(x) = x_1f(e_1) + \cdots + x_nf(e_n). \tag{2.6}$$

Iedere $f(e_i)$ is een kolom in $\mathbb{R}^m$. Zij nu A de $m \times n$ matrix met als i-de kolom $f(e_i)$, m.a.w. $A_i = f(e_i)$. Uit (2.4) en (2.6) volgt dat

$$f(x) = x_1A_1 + \cdots + x_nA_n = Ax = f_A(x).$$

Dit geldt voor alle $x \in \mathbb{R}^n$, dus $f = f_A$. Omdat $f_A(e_i) = f(e_i)$ moet zijn, ligt de i-de kolom A_i van A door $f(e_i)$ vast, dus is de matrix A eenduidig.
ii) Het is genoeg als we dit voor een vector e_j uit de standaardbasis van $\mathbb{R}^n$ kunnen aantonen. Er geldt $f(e_j) = a_{1j}e'_1 + \cdots + a_{mj}e'_m = \sum_{i=1}^m a_{ij}e'_i$, waarbij $(e'_1, \ldots, e'_m)$ de standaardbasis van $\mathbb{R}^m$ is. Dan is $g(f(e_j)) = \sum_{i=1}^m g(a_{ij}e'_i) = \sum_{i=1}^m a_{ij}g(e'_i) = \sum_{i=1}^m a_{ij}(\sum_{k=1}^l b_{ki}e''_k)$ met $(e''_1, \ldots, e''_l)$ de standaardbasis van $\mathbb{R}^l$. Hieruit volgt $(g \circ f)(e_j) = \sum_{k=1}^l(\sum_{i=1}^m b_{ki}a_{ij})e''_k = \sum_{k=1}^l(BA)_{kj}.e''_k$, dus geeft de j-de kolom van de matrix BA juist de coördinaten van $(g \circ f)(e_j)$.
iii) Neem nu aan dat $n = m$. Stel dat A inverteerbaar is. Dan bestaat er een $n \times n$ matrix B met $AB = BA = I_n$ en dus $A(Bx) = B(Ax) = x$ voor alle $x \in \mathbb{R}^n$ m.a.w. $f_A \circ f_B = f_B \circ f_A = 1_{\mathbb{R}^n}$. Dus $f = f_A$ is inverteerbaar.

Omgekeerd, stel dat $f = f_A$ inverteerbaar is. Dan bestaat er een lineaire afbeelding $g : \mathbb{R}^n \to \mathbb{R}^n$ met $f \circ g = g \circ f = 1_{\mathbb{R}^n}$. Volgens i) bestaat er een $n \times n$ matrix B met $g = f_B$. Dus $f_A \circ f_B = f_B \circ f_A = 1_{\mathbb{R}^n}$. Laat nu deze relatie werken op e_i. Dit geeft $A(Be_i) = B(Ae_i) = e_i$ ofwel $(AB)e_i = (BA)e_i = e_i$. M.a.w.

$$i\text{-de kolom van } AB \ = \ i\text{-de kolom van } BA \ = \ i\text{-de kolom van } I_n.$$

Omdat dit geldt voor iedere i volgt $AB = BA = I_n$, m.a.w. A is inverteerbaar. □

We kunnen nu een zeer kort bewijs van 2.4.6 uit LA1 geven, d.w.z. we bewijzen:

Stelling 2.3.6 Zij A een $n \times n$ matrix. Als uit $Ax = 0$ volgt dat $x = 0$, dan is A inverteerbaar.

Bewijs. Zij f_A de bij A behorende lineaire afbeelding van $\mathbb{R}^n$ naar zichzelf. Dan volgt uit het gegeven: als $f_A(x) = 0$, dan $x = 0$. Vanwege 2.2.21 is dan f_A inverteerbaar. Dus A is inverteerbaar (vanwege 2.3.5). □

Opgave 2.3.7 Zij A een $n \times n$ matrix. Laat zien dat de volgende uitspraken gelijkwaardig zijn.

i) A is inverteerbaar.
ii) rang $A = n$.
iii) De kolommen van A zijn onafhankelijk.
iv) De rijen van A zijn onafhankelijk.
v) Als $Ax = 0$, dan $x = 0$.
vi) Bij iedere $b \in \mathbb{R}^n$ bestaat er een $x \in \mathbb{R}^n$ met $Ax = b$.

vii) Er bestaat een $n \times n$ matrix B met $AB = I_n$.
viii) Er bestaat een $n \times n$ matrix B met $BA = I_n$.
ix) Nul is geen eigenwaarde van A.
x) $\det A \neq 0$.

2.4 Basistransformaties

We hebben in de vorige sectie gekeken naar de matrix van een lineaire afbeelding van $\mathbb{R}^n$ naar $\mathbb{R}^m$. Als we naar lineaire afbeeldingen tussen willekeurige (maar wel eindig dimensionale) vectorruimten V en W kijken, zullen we gebruik maken van de classificatie stelling 2.2.11, die zegt dat $V \cong \mathbb{R}^n$ en $W \cong \mathbb{R}^m$ voor $\dim V = n$ en $\dim W = m$. Het cruciale hulpmiddel hierbij zijn de *coördinaatvectoren* die door middel van het isomorfisme $\mathbb{R}^n \to V$ de vectoren in V beschrijven als vectoren in $\mathbb{R}^n$.

Notatie: Voor een basis $B = (v_1, \ldots, v_n)$ van V en $v = x_1v_1 + \cdots x_nv_n$ noteren we de coördinaatvector $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ van v met $x = _{B}v$.

Definitie 2.4.1 Zij $f : V \to W$ een lineaire afbeelding en laten $B = (v_1, \ldots, v_n)$ en $C = (w_1, \ldots, w_m)$ bases van V en W zijn. Dan heet de $m \times n$ matrix A met

$$f(v_j) = a_{1j}w_1 + a_{2j}w_2 + \cdots + a_{mj}w_m$$

de *matrix van de lineaire afbeelding* f m.b.t. de bases B en C. Deze wordt genoteerd met $A = _{C}f_{B}$.

Lemma 2.4.2 Zij $f : V \to W$ een lineaire afbeelding en zij $A = _{C}f_{B} \in M_{m,n}(\mathbb{R})$ de matrix van f m.b.t. de bases $B = (v_1, \ldots, v_n)$ en $C = (w_1, \ldots, w_m)$. Zij $v \in V$ en zij $x = _{B}v$ de coördinaatvector van v t.o.v. de basis B. Dan heeft $f(v)$ t.o.v. de basis C de coördinaatvector Ax d.w.z. $_{C}f(v) = _{C}f_{B} \cdot _{B}v$.

Bewijs. We laten dit eerst voor de basisvectoren van V zien. Omdat (volgens de definitie van A) $f(v_j) = a_{1j}w_1 + \cdots + a_{mj}w_m$, is $\begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix}$ de coördinaatvector van $f(v_j)$ en dit is juist de j-de kolom A_j van A. Aan de andere kant is de coördinaatvector van v_j de vector e_j uit de standaardbasis, en Ae_j is ook de j-de kolom van A, dus klopt de uitspraak voor de basisvectoren.

Nu geldt voor $v \in V$ met coördinaatvector x dat $f(v) = f(x_1v_1 + \cdots + x_nv_n) = x_1f(v_1) + \cdots + x_nf(v_n)$ en dit heeft $x_1A_1 + \cdots + x_nA_n$ als coördinaatvector, dus Ax. $\square$

Opmerking 2.4.3 Voor $\mathbb{R}^n$ en $\mathbb{R}^m$ met de standaardbases E en E' zijn de coördinaatvectoren hetzelfde als de vectoren zelf. Daarom is in dit geval de matrix $A = _{E'}f_{E}$ van de lineaire afbeelding f juist de matrix A met $f = f_A$.

Algemeen bevat de j-de kolom van de matrix van een lineaire afbeelding de coördinaatvector van het beeld van de j-de basisvector.

Voorbeeld 2.4.4 Zij $V = \langle \sin(t), \cos(t) \rangle$ en zij $D : V \to V$ de afgeleide, d.w.z. $D(f) := f'$. Dan heeft D m.b.t. de basis $B = (\sin(t), \cos(t))$ de matrix $A = {}_B D_B = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$, want $\sin(t)' = 0.\sin(t) + 1.\cos(t)$ en $\cos(t)' = -1.\sin(t) + 0.\cos(t)$.

Voorbeeld 2.4.5 Zij $T : \mathrm{Pol}(3) \to \mathrm{Pol}(3)$ gegeven door $p(t) \to p(t+1)$. Dan heeft T m.b.t. de basis $B = (1, t, t^2, t^3)$ de matrix $A = {}_B T_B = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 2 & 3 \\ 0 & 0 & 1 & 3 \\ 0 & 0 & 0 & 1 \end{pmatrix}$.

Voorbeeld 2.4.6 Zij $I : \mathrm{Pol}(3) \to \mathrm{Pol}(4)$ gegeven door $I(p) := \int_0^x p(t)\,dt$. Bepaal de matrix A van I t.o.v. de bases $(1, t, t^2, t^3)$ en $(1, x, x^2, x^3, x^4)$ van Pol(3) en Pol(4).

Oplossing. Er geldt $\int_0^x t^k\,dt = \frac{1}{k+1} t^{k+1}\big|_0^x = \frac{1}{k+1} x^{k+1}$. Dus is

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

Voorbeeld 2.4.7 Zij $f : \mathbb{R}^2 \to \mathbb{R}^2$ de rotatie rond de oorsprong met hoek φ. Bepaal de matrix van f m.b.t. de standaardbasis van $\mathbb{R}^2$.

Oplossing. Uit de elementaire meetkunde weten we dat $f(\begin{pmatrix} 1 \\ 0 \end{pmatrix}) = \begin{pmatrix} \cos(\varphi) \\ \sin(\varphi) \end{pmatrix}$ en $f(\begin{pmatrix} 0 \\ 1 \end{pmatrix}) = \begin{pmatrix} -\sin(\varphi) \\ \cos(\varphi) \end{pmatrix}$. De matrix van f m.b.t. de standaardbasis is dus $\begin{pmatrix} \cos(\varphi) & -\sin(\varphi) \\ \sin(\varphi) & \cos(\varphi) \end{pmatrix}$.

Voorbeeld 2.4.8 Zij $f : \mathbb{R}^2 \to \mathbb{R}^2$ de spiegeling in de diagonaal $x = y$, d.w.z.

$$f(\begin{pmatrix} x \\ y \end{pmatrix}) = \begin{pmatrix} y \\ x \end{pmatrix}.$$

M.b.t. de standaardbasis $\left(\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}\right)$ heeft f de matrix $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$.

Maar m.b.t. de basis $\left(\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix}\right)$ met een vector op de spiegelingsas en een loodrecht erop heeft f de matrix $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$.

Stelling 2.4.9 Laten $f : U \to V$ en $g : V \to W$ lineaire afbeeldingen zijn met matrices ${}_C f_B$ en ${}_D g_C$ m.b.t. de bases $B = (u_1, \ldots, u_n)$, $C = (v_1, \ldots, v_m)$ en $D = (w_1, \ldots, w_k)$ voor U, V en W, respectievelijk. Dan heeft $g \circ f : U \to W$ m.b.t. de bases B en D de matrix ${}_D(g \circ f)_B = {}_D g_C \cdot {}_C f_B$.

Bewijs. Volgens 2.4.2 is de j-de kolom van ${}_C f_B$ de coördinaatvector van $f(u_j)$ t.o.v. de basis C. Nog een keer volgens 2.4.2 is dan de j-kolom van ${}_D g_C \cdot {}_C f_B$ de coördinaatvector van $g(f(u_j))$ t.o.v. de basis D. □

Opmerking 2.4.10 De notatie van de matrix van een lineaire afbeelding met twee indices voor de bases heeft het effect, dat we bij de samenstelling van twee lineaire afbeeldingen de binnenste bases kunnen 'schrappen', deze bases moeten namelijk sowieso hetzelfde zijn: de basis waarin we de beelden van de eerste afbeelding uitdrukken moet gelijk zijn aan de basis voor de originelen van de tweede afbeelding.

Let op de volgorde van de bases: omdat we samengestelde afbeeldingen van rechts naar links lezen, staat de *basis voor de originelen rechts* en de *basis voor de beelden links*.

Omdat de matrix van een lineaire afbeelding m.b.v. de coördinaatvectoren gedefinieerd is, hangt de matrix van de keuze van de bases voor V en W af. We zullen nu nagaan hoe de matrix verandert, als we een of beide bases wijzigen. Hiervoor is de volgende observatie belangrijk. We hebben gezien dat een inverteerbare $n \times n$ matrix T bij een bijectieve lineaire afbeelding $\mathbb{R}^n \to \mathbb{R}^n$ hoort, i.h.b. vormen de kolommen van T een basis van $\mathbb{R}^n$. Met onze nieuwe zichtwijze op de matrix van een lineaire afbeelding kunnen we een inverteerbare matrix echter ook anders opvatten.

Lemma 2.4.11 Zij V een vectorruimte met basis $C = (v_1, \ldots, v_n)$ en zij T een inverteerbare $n \times n$ matrix. Zij $u_j \in V$ de vector met als coördinaatvector ${}_C u_j$ t.o.v. de basis C de j-de kolom T_j van T. Dan is T de matrix van de identiteit $1 = 1_V$ m.b.t. de bases $B = (u_1, \ldots, u_n)$ en C, d.w.z. $T = {}_C 1_B$.

Bewijs. Het beeld van de j-de basisvector uit de basis $B = (u_1, \ldots, u_n)$ onder de identiteit is nog steeds u_j en t.o.v. de basis C heeft u_j de coördinaatvector T_j, dus is T_j de j-de kolom van ${}_C 1_B$. □

Gevolg 2.4.12 Laten B en C twee bases van de n-dimensionale vectorruimte V zijn en zij $T = {}_C 1_B$. Dan geldt voor de inverse matrix T^{-1} van T dat $T^{-1} = {}_B 1_C$.

Bewijs. Zij $S = {}_B 1_C$. Volgens 2.4.9 geldt

$$TS = {}_C 1_B \cdot {}_B 1_C = {}_C 1_C = I_n \quad \text{en} \quad ST = {}_B 1_C \cdot {}_C 1_B = {}_B 1_B = I_n,$$

dus is $S = T^{-1}$. □

Opmerking 2.4.13 Omdat de matrix $T = {}_C 1_B$ de vectoren van de basis B als coördinaatvectoren in de matrix C uitdrukt, noemt men T ook een *basistransformatie*. Vaak is een van de twee bases B en C de standaardbasis $E = (e_1, \ldots, e_n)$ van $\mathbb{R}^n$. In het geval $C = E$ is dan de basistransformatie bijzonder eenvoudig, want de vectoren in de basis B zijn gelijk aan hun coördinaatvectoren t.o.v. de standaardbasis E, d.w.z. $T = {}_E 1_B$ heeft gewoon de vectoren in B als kolommen.

Voor het omgekeerde geval is de matrix $T = {}_C 1_E$ de inverse matrix van de matrix met de vectoren in C als kolommen.

Lemma 2.4.14 Zij $f : V \to W$ een lineaire afbeelding, laten B en B' bases van V en C en C' bases van W zijn. Verder zij $A = {}_C f_B$ de matrix van de lineaire afbeelding f m.b.t. de 'oude' bases B en C, en laten $T = {}_B 1_{B'}$ en $S = {}_C 1_{C'}$ de basistransformaties zijn de die nieuwe in de oude bases uitdrukken.

Dan is de matrix A' van de lineaire afbeelding f m.b.t. de nieuwe bases B' en C' gegeven door

$${}_{C'} f_{B'} = A' = S^{-1} A T.$$

Bewijs. Er geldt ${}_{C'}f_{B'} = {}_{C'}1_C \cdot {}_C f_B \cdot {}_B 1_{B'}$ en omdat $S = {}_C 1_{C'}$ is ${}_{C'}1_C = S^{-1}$. □

Een belangrijk speciaal geval van lemma 2.4.14 is een lineaire afbeelding $f : V \to V$ die op een andere basis van V wordt getransformeerd.

Stelling 2.4.15 Zij $f : V \to V$ een lineaire afbeelding en laten B en B' bases van V zijn. Verder zij $A = {}_B f_B$ de matrix van de lineaire afbeelding f m.b.t. de 'oude' basis B en zij $T = {}_B 1_{B'}$ de basistransformatie die de nieuwe in de oude basis uitdrukt.

Dan is de matrix A' van de lineaire afbeelding f m.b.t. de nieuwe basis B' gegeven door

$$ {}_{B'}f_{B'} = A' = T^{-1}AT. $$

Bewijs. Dit is de uitspraak van 2.4.14 met $V = W$, $C = B$ en $C' = B'$. □

De uitspraak van deze stelling zijn we in LA1 stelling 4.2.6 al een keer tegen gekomen. Zij namelijk A de matrix van de lineaire afbeelding $f = f_A$ m.b.t. de standaardbasis van $\mathbb{R}^n$ en stel A heeft n verschillende eigenwaarden $c_1, \ldots, c_n$ met corresponderende eigenvectoren $v_1, \ldots, v_n$. Dan geldt voor de matrix T met de eigenvectoren v_i als kolommen dat $T^{-1}AT = D$, waarbij D de diagonaalmatrix is met de eigenwaarden c_i op de diagonaal. Maar D is juist de matrix van f m.b.t. de basis $(v_1, \ldots, v_n)$, want T is de basistransformatie die de basis uit eigenvectoren in de standaardbasis uitdrukt.

Voorbeeld 2.4.16 Zij f de lineaire afbeelding met matrix $A = \frac{1}{5}\begin{pmatrix} -4 & 3 \\ 3 & 4 \end{pmatrix}$ m.b.t. de standaardbasis van $\mathbb{R}^2$.

Met de methodes uit LA1 gaat men eenvoudig na dat A eigenwaarden 1 en -1 heeft. De eigenvectoren voor de eigenwaarde 1 liggen in de kern van $I_2 - A = \frac{1}{5}\begin{pmatrix} 9 & -3 \\ -3 & 1 \end{pmatrix}$, dit zijn dus de vectoren $\begin{pmatrix} x \\ 3x \end{pmatrix}$ met $x \neq 0$. Analoog liggen de eigenvectoren voor de eigenwaarde -1 in de kern van $-I_2 - A = \frac{1}{5}\begin{pmatrix} -1 & -3 \\ -3 & -9 \end{pmatrix}$, dit zijn dus de vectoren $\begin{pmatrix} 3x \\ -x \end{pmatrix}$ met $x \neq 0$.

Een basis uit eigenvectoren is dus (bijvoorbeeld) $\left(\begin{pmatrix} 1 \\ 3 \end{pmatrix}, \begin{pmatrix} 3 \\ -1 \end{pmatrix}\right)$ en met $T = \begin{pmatrix} 1 & 3 \\ 3 & -1 \end{pmatrix}$ geldt $T^{-1}AT = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$.

Deze afbeelding is dus een spiegeling in de lijn langs de vector $\begin{pmatrix} 1 \\ 3 \end{pmatrix}$.

Historische opmerkingen

Het was Oliver Heaviside (1850-1925) die het begrip lineaire vector operator invoerde in een van zijn werken over elektro-magnetisme in 1885. Hij definieerde het gebruikmakend van coördinaten: de vector $\vec{B}$ komt van de vector $\vec{H}$ d.m.v. een lineaire vector operator als, wanneer $\vec{B}$ componenten B_1, B_2, B_3 en $\vec{H}$ componenten H_1, H_2, H_3 hebben, er getallen μ_{ij}

zijn voor $i, j = 1, 2, 3$ zodanig dat

$$\begin{aligned}
B_1 &= \mu_{11}H_1 + \mu_{12}H_2 + \mu_{13}H_3 \\
B_2 &= \mu_{21}H_1 + \mu_{22}H_2 + \mu_{23}H_3 \\
B_3 &= \mu_{31}H_1 + \mu_{32}H_2 + \mu_{33}H_3
\end{aligned}$$

In zijn colleges aan de Yale Universiteit, die gepubliceerd werden in 1901, noemde J. Willard Gibbs dezelfde transformatie een lineaire vectorfunctie. Maar hij definieerde dit ook abstracter als een continue functie zodanig dat

$$f(\vec{v} + \vec{w}) = f(\vec{v}) + f(\vec{w}).$$

Exact dezelfde definitie als onze definitie 2.2.1 werd gegeven door Hermann Weyl in zijn boek 'Raum-Zeit-Materie' in 1918.