

STICHTING
MATHEMATISCH CENTRUM

2e BOERHAAVESTRAAT 49

A M S T E R D A M

STATISTISCHE AFDELING

Rapport S 265 (C 13)

Leergang Besliskunde

Hoofdstuk XIII

Kwadratische Programmering

door

S.C. van Westrhenen

november 1960

1. Inleiding

Het lineaire programmeringsmodel, zoals beschreven in hoofdstuk XII van deze leergang, zal in veel gevallen een te eenvoudig model zijn, dat of in het geheel niet toepasbaar is op, of slechts een eerste benadering is van het praktijkprobleem (bijv. de te optimaliseren winstfuncties mogen in het beschouwde gebied niet als lineair beschouwd worden, zie bijv. productieproblemen).

Een geringe generalisatie, namelijk door alle of enkele van de te optimaliseren lineaire functies (lineair in de variabelen waarover men vrij kan beschikken) te vervangen door kwadratische functies, maakt het toepassingsgebied van het programmeringsmodel veel ruimer.

De meeste van de tot nu toe bekende algorithmen eisen van de te optimaliseren functies convexiteit (zie definitie 3.1). Als deze voorwaarde ontbreekt, dan is men er in het algemeen niet zeker van, of de gevonden waarde inderdaad een absoluut extremum is. Máár enkele algorithmen kunnen in dit geval zodanig gewijzigd worden, dat de uitkomst minstens een lokaal extremum oplevert (o.a. het in paragraaf 5 te bespreken algoritme van BEALE).

Het lineaire en het kwadratische programmeringsmodel kunnen als specialisaties van het algemene programmeringsmodel beschouwd worden nl.:

Minimaliseer $f(x_1, \dots, x_n)$ onder de nevenvoorwaarden (1.1)

$$\begin{aligned} g_k(x_1, \dots, x_n) &\geq 0 & k &= 1, 2, \dots, m \\ x_j &\geq 0 & j &= 1, 2, \dots, n. \end{aligned}$$

Het programmeringsmodel is:

a. lineair als $f(x_1, \dots, x_n)$ en $g_k(x_1, \dots, x_n)$ lineaire functies van $x_1, \dots, x_n$ zijn. Voorbeeld voor $m=3$ en $n=4$:

Minimaliseer $f(x_1, x_2, x_3, x_4) = 5x_1 + 7x_2 + 3x_3 + 5x_4$ (1.2)
onder de nevenvoorwaarden

$$\begin{aligned} g_1(x_1, x_2, x_3, x_4) &= 9x_1 + 7x_2 && \geq 8 \\ g_2(x_1, x_2, x_3, x_4) &= x_2 + x_3 + x_4 && \geq 9 \\ g_3(x_1, x_2, x_3, x_4) &= x_1 + 7x_4 && \geq 13 \\ x_1, x_2, x_3, x_4 &\geq 0 \end{aligned}$$

b. kwadratisch als $f(x_1, \dots, x_n)$ en $g_k(x_1, \dots, x_n)$ resp. kwadratisch en lineair zijn in $x_1, \dots, x_n$. Voorbeeld voor $m=2$, $n=3$:

Minimaliseer $f(x_1, x_2, x_3) = 6x_1 + 7x_2 + x_3 + 5x_1^2 + 3x_2^2 + 7x_3^2 + 8x_1x_2$ (1.3)
 onder de nevenvoorwaarden:

$$\begin{aligned} g_1(x_1, x_2, x_3) &= 6x_1 + 3x_2 + x_3 \geq 10 \\ g_2(x_1, x_2, x_3) &= 2x_1 + 3x_2 + 9x_3 \geq 12 \\ x_1, x_2, x_3 &\geq 0. \end{aligned}$$

Men spreekt ook wel van kwadratische programmering als behalve $f(x_1, \dots, x_n)$ ook één of meerdere functies $g_k(x_1, \dots, x_n)$ kwadratisch in de variabelen x_1 t/m x_n zijn.¹⁾ Aangezien echter vrijwel alle tot nu bekende algoritmen alleen maar van toepassing zijn op het eerstgenoemde geval, zullen we het begrip kwadratische programmering alleen in de eerste betekenis gebruiken.

Het stelsel (1.1) kan verkort opgeschreven worden als:

Minimaliseer $f(X)$

onder de nevenvoorwaarden:

$$\begin{aligned} g_k(X) &\geq 0 \quad k=1, 2, \dots, m \\ X &\geq 0, \end{aligned}$$

waarin X de vector $(x_1, \dots, x_n)$ voorstelt.

In dit hoofdstuk zullen dus slechts programmeringsproblemen ter sprake komen waarin

- de te minimaliseren functie $f(X)$ kwadratisch is in de variabelen, waarover men vrij kan beschikken en waarin $f(X)$ bovendien (strikt) convex is in X (zie def. 3.1),
- de nevenvoorwaarden of restricties lineair zijn in de onder a) genoemde variabelen,
- de onder a) genoemde variabelen niet negatief zijn.

In formule luidt dit:

$$\text{Minimaliseer } f(X) = \sum_{i=1}^n p_i x_i + \frac{1}{2} \sum_{i,j=1}^n c_{ij} x_i x_j \quad (c_{ij} = c_{ji}) \quad (1.4)$$

$$\text{onder de restricties } \sum_{j=1}^n a_{ij} x_j = a_{i0} \quad i = 1, 2, \dots, m \quad (1.5)$$

$$x_j \geq 0, \quad a_{i0} \geq 0 \quad j = 1, 2, \dots, n, \quad (1.6)$$

of in vectornotatie:

1) Zie bijv. G. ZOUTENDIJK, Methods of feasible directions, 1960, dissertatie - Elsevier-Amsterdam, Londen etc.

Minimaliseer $f(X) = P \cdot X + \frac{1}{2} X \cdot C \cdot X$

onder de restricties $A \cdot X = A_0$

$$X \geq 0, \quad A_0 \geq 0,$$

waarin: A_0 - m-vector ¹⁾
 P - n-vector
 X - n-vector
 A - mxn matrix
 C - nxn symmetrische (semi-) positief
 definitie matrix.

Een matrix C is symmetrisch als voor ieder element c_{ij} geldt:

$c_{ij} = c_{ji}$. A_1 is bijv. een symmetrische matrix (zie verder par. 5 van hoofdstuk VIII).

De gespiegelde matrix van A , die ontstaat door de elementen a_{ij} van A te vervangen door a_{ji} , wordt aangegeven door A' . A_3 is bijv. de gespiegelde van A_2 en omgekeerd. A_1 is de gespiegelde van zichzelf. Dit laatste geldt voor iedere symmetrische matrix.

$$A_1 = \begin{pmatrix} 1 & 4 & 6 \\ 4 & 2 & 5 \\ 6 & 5 & 3 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{pmatrix}.$$

Om in het algorithm van BEALE een vereenvoudiging in de notatie te verkrijgen - zie formule (4.17) - schrijft men de functie $f(X)$ ook wel in homogene coördinaten. In plaats van f_1 schrijft men dan f_2 .

$$f_1 = \sum_{i=1}^n p_i x_i + \frac{1}{2} \sum_{i,j=1}^n c_{ij} x_i x_j = 0 + \frac{1}{2} p_1 x_1 + \dots + \frac{1}{2} p_n x_n$$

$$\frac{1}{2} p_1 x_1 + \frac{1}{2} c_{11} x_1^2 + \dots + \frac{1}{2} c_{1n} x_1 x_n \quad (1.7)$$

$$\begin{aligned} & \vdots \\ & \frac{1}{2} p_n x_n + \frac{1}{2} c_{n1} x_n x_1 + \dots + \frac{1}{2} c_{nn} x_n^2 \\ f_2 = & \bar{c}_{00} x_0^2 + \bar{c}_{01} x_0 x_1 + \dots + \bar{c}_{0n} x_0 x_n \quad (18) \\ & \bar{c}_{10} x_1 x_0 + \bar{c}_{11} x_1^2 + \dots + \bar{c}_{1n} x_1 x_n \\ & \vdots \\ & \bar{c}_{n0} x_n x_0 + \bar{c}_{n1} x_n x_1 + \dots + \bar{c}_{nn} x_n^2 \end{aligned}$$

 1) In de notatie zal geen verschil tussen kolom- en rij-vectoren gemaakt worden.

waarin $\bar{c}_{00}=0$ $\bar{c}_{j0}=\bar{c}_{0j}=\frac{1}{2}p_j$ $\bar{c}_{ij}=\frac{1}{2}c_{ij}$
 $i, j=1, 2, \dots, n$ en $x_0=1$.

Met behulp van vector- en matrixnotatie luidt f_2 :

$$f_2(X) = X \cdot \bar{C} \cdot X,$$

waarin X nu de vector $(x_0, x_1, \dots, x_n)$ voorstelt met $x_0=1$.

Numeriek voorbeeld:

$$\begin{aligned} f_1 &= 2x_1 + 8x_2 + 4x_3 + x_1^2 + 4x_1x_2 + 2x_1x_3 + 4x_2^2 - 2x_2x_3 + 2x_3^2 \quad (1.10) \\ &= 2x_1 + 8x_2 + 4x_3 + \frac{1}{2}(2x_1^2 + 8x_1x_2 + 4x_1x_3 + 8x_2^2 - 4x_2x_3 + 4x_3^2), \end{aligned}$$

of

$$\begin{aligned} f_1 &= \begin{aligned} &x_1 &+ &4x_2 &+ &2x_3 \\ &x_1 &+ &x_1^2 &+ &2x_1x_2 &+ &x_1x_3 \\ &4x_2 &+ &2x_2x_1 &+ &4x_2^2 &- &x_2x_3 \\ &2x_3 &+ &x_3x_1 &- &x_2x_3 &+ &2x_3^2 \end{aligned} \quad (1.11) \end{aligned}$$

$$\begin{aligned} f_2 &= 0 \cdot x_0^2 + 1 \cdot x_0x_1 + 4 \cdot x_0x_2 + 2x_0x_3 \quad x_0=1 \quad (1.12) \\ &1 \cdot x_1x_0 + 1 \cdot x_1^2 + 2 \cdot x_1x_2 + x_1x_3 \\ &4 \cdot x_2x_0 + 2 \cdot x_2x_1 + 4x_2^2 - x_2x_3 \\ &2 \cdot x_3x_0 + 1 \cdot x_3x_1 - x_2x_3 + 2x_3^2. \end{aligned}$$

In matrixnotatie luidt f_2 dus:

$$f_2 = (x_0, x_1, x_2, x_3) \begin{pmatrix} 0 & 1 & 4 & 2 \\ 1 & 1 & 2 & 1 \\ 4 & 2 & 4 & -1 \\ 2 & 1 & -1 & 2 \end{pmatrix} \begin{pmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \end{pmatrix} \quad x_0=1 \quad (1.13)$$

2. Voorbeelden van kwadratische programmeringsproblemen

Het kwadratische algoritme kan bijvoorbeeld toegepast worden op de volgende problemen:

a) Produktieproblemen.¹⁾

We beschouwen eerst een produktiemodel van een automobielenfabriek

 1) Het hier genoemde voorbeeld is ontleend aan: R. DORFMAN, P.A. SAMUELSEN en R.M. SOLOW - "Linear Programming and Economic Analysis" - MacGraw-Hill New York 1958.

die produceert voor een markt met volledige mededinging (d.w.z. de verkoopprijs van een automobiel is onafhankelijk van de grootte van de eigen produktie). Dit markttype geeft bij het hier beschouwde produktiemodel aanleiding tot een lineair programmeringsmodel.

Door nu de marktform te veranderen in een monopolistisch markttype (d.w.z. de verkoopprijs is afhankelijk van de grootte van de eigen produktie) blijkt dat het programmeringsmodel, onder bepaalde voorwaarden, kwadratisch wordt.

De genoemde fabriek heeft vier afdelingen, waarvan de grondstoffen onbeperkt toegeleverd kunnen worden. De produktie bestaat uit luxe auto's en trucks, waarop, bij de markt met volledige mededinging, resp. een netto winst per stuk gemaakt wordt van f 300.- en f 250.-.

De maximumcapaciteit per maand van iedere afdeling is in tabel 2.I aangegeven.

Tabel 2.I

Produktiecapaciteiten per maand van de afdelingen van
aan autofabriek

	Afdeling	Luxe auto's	Trucks
I	Metaalstamping	25.000	35.000
II	Motor assemblage	33.333	16.667
III	Auto assemblage	22.500	-
IV	Truck assemblage	-	15.000

Uit deze tabel blijkt dat er maximaal 22.500 luxe auto's of 15.000 trucks per maand afgeleverd kunnen worden.

Voor de produktie van één luxe wagen (truck) per maand heeft men $\frac{1}{25.000}$ ($\frac{1}{35.000}$) deel van de totale produktiecapaciteit van afdeling I nodig. Analooq vindt men voor afdeling II de breuken $\frac{1}{33.333}$ en $\frac{1}{16.667}$. Als nu gevraagd wordt, onder bovenstaande restricties, naar de produktiecombinatie, waarbij de winst optimaal is, dan kan men het volgende lineaire programmeringsmodel opstellen (x_1 stelt het aantal geproduceerde auto's en x_2 het aantal geproduceerde trucks voor):

$$\frac{1}{25.000} \cdot x_1 + \frac{1}{35.000} \cdot x_2 = \text{deel gebruikt v.d. produktiecap. van afd. I} \leq 1$$

$$\frac{1}{33.333} \cdot x_1 + \frac{1}{16.667} \cdot x_2 = \text{" " " " " " " " II} \leq 1$$

$$x_1 = \text{produktie van afdeling III} \leq 22.500$$

$$x_2 = \text{" " " " IV} \leq 15.000.$$

De totale winst bij een produktie van x_1 luxe auto's en x_2 trucks bedraagt: $300x_1 + 250x_2$ gulden. Bovendien geldt $x_1, x_2 \geq 0$.

In formulevorm luidt dit probleem:

$$\text{Maximaliseer } f = 300x_1 + 250x_2 \quad (2.1)$$

$$\text{onder de restricties: } \frac{x_1}{25.000} + \frac{x_2}{35.000} \leq 1 \quad (2.2)$$

$$\frac{x_1}{33.333} + \frac{x_2}{16.667} \leq 1$$

$$x_1 \leq 22.500$$

$$x_2 \leq 15.000$$

$$x_1, x_2 \geq 0.$$

De nevenvoorwaarden (2.2) kunnen grafisch uitgezet worden. Uit fig. 2.1 blijkt dat de toegelaten oplossingen binnen het gebied OABCDE liggen. De iso-winstlijnen - $300x_1 + 250x_2 = \text{constant}$ - vormen een stelsel evenwijdige lijnen. De iso-winstlijn door E geeft de maximale winst, en punt C de optimale produktiecombinatie aan.

AUTO'S
(IN DUIZENDTALLEN)

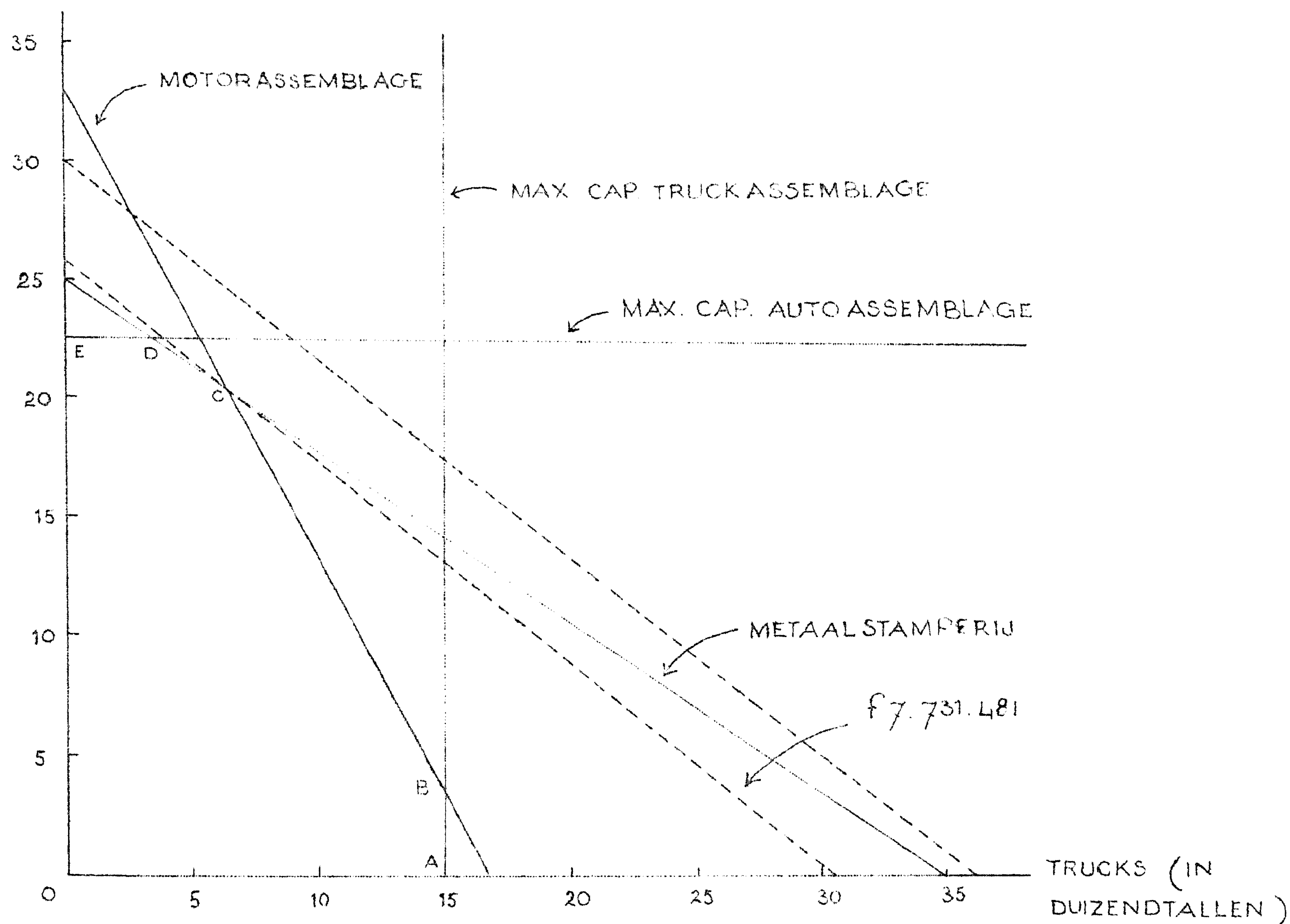


fig. 2.1

Het gebied van toegelaten produktiecombinaties en de iso-winstlijnen (---) voor een markt met volledige mededinging

Door berekening, bijv. door middel van de simplexmethode, vindt men dat de optimale waarde van $f(=7.731.481)$ aangenomen wordt voor $x_1=20.370$ en $x_2=6.481$.

Vervolgens beschouwen we een markt, waarop de winst v_1 per luxe auto afhankelijk is van het aantal geproduceerde (=verkochte) luxe auto's, terwijl de winst v_2 per truck dezelfde blijft, dus onafhankelijk is van het verkochte aantal trucks. Stel bijvoorbeeld:

$$v_1 = 625 - \frac{x_1}{60} \quad \text{en} \quad v_2 = 250. \quad (2.3)$$

De winstfunctie krijgt dan de vorm:

$$f = v_1 x_1 + v_2 x_2 = (625 - \frac{x_1}{60})x_1 + 250x_2 \quad (2.4)$$

De iso-winst lijnen zijn nu kromme lijnen waarvan de vergelijking luidt:

$$-\frac{1}{60} x_1^2 + 625x_1 + 250x_2 = \text{const.} \quad (2.5)$$

Uit fig. 2.2 blijkt - de restricties zijn dezelfde gebleven - dat de optimale iso-winstlijn BC raakt in P en dat P dus de optimale productiecombinatie is. Het optimale produktieschema luidt nu $x_1=15.000$ $x_2=9.167$, terwijl de optimale winst 7.916.667 gulden bedraagt.

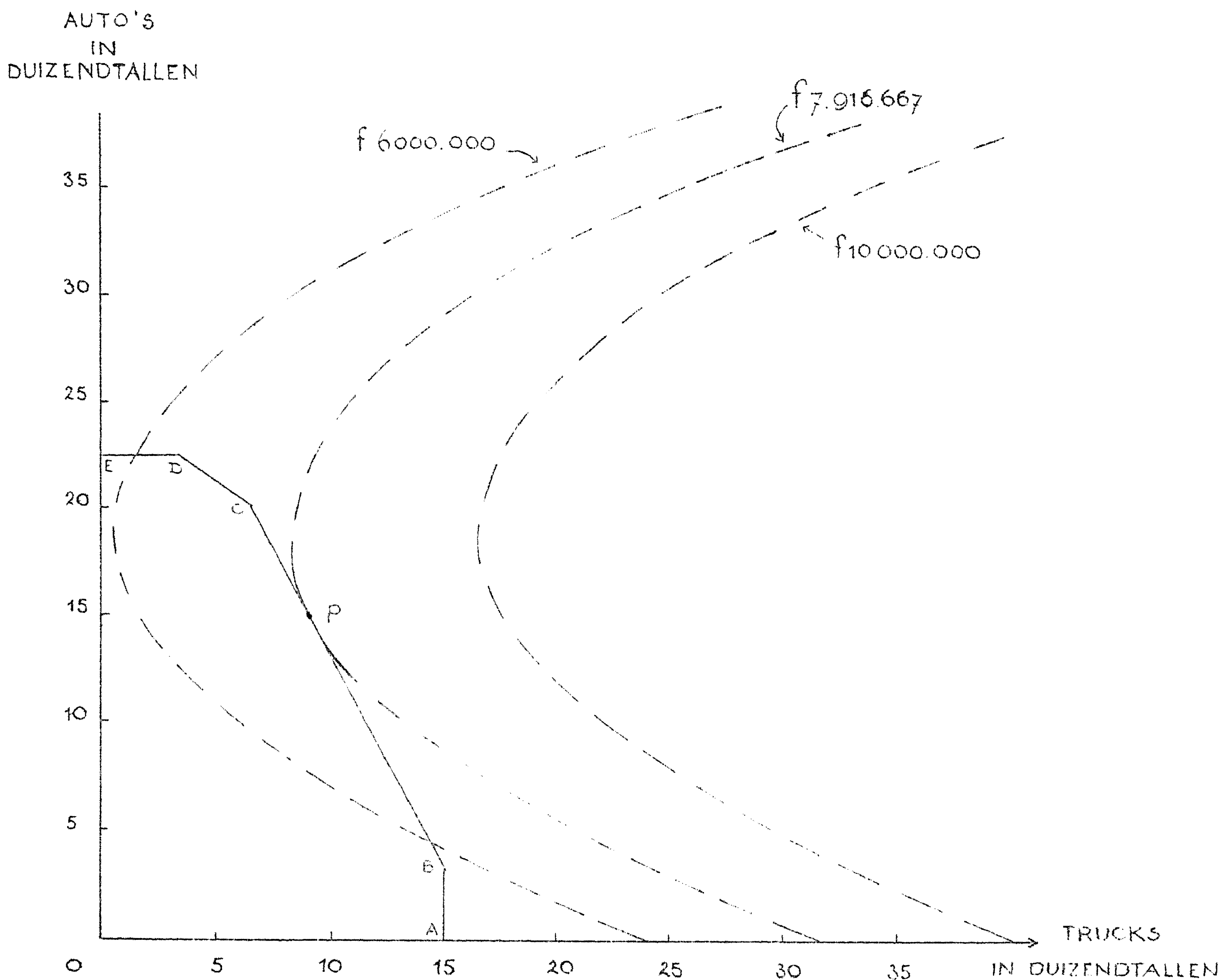


fig. 2.2

Het gebied van toegelaten productiecombinaties en de iso-winstlijnen (---) voor een deels monopolistische markt

Opmerking

De optimale produktiecombinatie van het kwadratische model heeft - in tegenstelling tot het lineaire model - niet samen te vallen met een hoekpunt van het convexe gebied OABCDE.

b. Beleggingsproblemen

Een voetbaltotowinnaar wil zijn prijs ad f 1.000.000.- beleggen in vier ondernemingen A, B, C en D (aangenomen wordt dat de prijs belastingvrij is).

De jaarlijkse dividenduitkeringen per 100 gld. kunnen worden voorgesteld door vier stochastische, onderling onafhankelijk verdeelde variabelen $\underline{\alpha}_1$, $\underline{\alpha}_2$, $\underline{\alpha}_3$ en $\underline{\alpha}_4$ met verwachtingen en spreidingen van resp. 4, 5, 6, 3 en 1, 2, 2, 1 gulden.

De belegger wil minstens 4,5% rente van zijn kapitaal maken, zo weinig mogelijk risico lopen en in de bedrijven B en C niet meer dan 60% van zijn kapitaal investeren. Zo weinig mogelijk risico lopen zullen wij hier interpreteren als het zo klein mogelijk maken van de variantie van de totale jaarlijkse dividenduitkering (vgl. ook voorbeeld 1.4 met hoofdstuk VII). De totowinnaar vraagt zich nu af, hoeveel gulden hij in elk bedrijf moet beleggen om dit doel te bereiken.

Stel hij belegt x_1 coupures van honderd gulden in bedrijf A, x_2 coupures van honderd gulden in bedrijf B etc. De totale dividenduitkering wordt dan voorgesteld door de stochastische variabele $\underline{d}$.

$$\underline{d} = \underline{\alpha}_1 x_1 + \underline{\alpha}_2 x_2 + \underline{\alpha}_3 x_3 + \underline{\alpha}_4 x_4 \quad (2.6)$$

Volgens stelling 4.3 van hoofdstuk III is de verwachting van $\underline{d}$:

$$E\underline{d} = x_1 E\underline{\alpha}_1 + x_2 E\underline{\alpha}_2 + x_3 E\underline{\alpha}_3 + x_4 E\underline{\alpha}_4$$

of in getallen:

$$E\underline{d} = 4x_1 + 5x_2 + 6x_3 + 3x_4 \quad (2.7)$$

De eis dat minstens 4,5% rente gemaakt moet worden betekent dus:

$$4x_1 + 5x_2 + 6x_3 + 3x_4 \geq 45.000. \quad (2.8)$$

Uit stelling 4.4 van hoofdstuk III volgt, dat de variantie van $\underline{d}$ is:

$$\sigma^2(\underline{d}) = \sigma^2(\underline{\alpha}_1 x_1) + \sigma^2(\underline{\alpha}_2 x_2) + \sigma^2(\underline{\alpha}_3 x_3) + \sigma^2(\underline{\alpha}_4 x_4)$$

of

$$\sigma^2(\underline{d}) = x_1^2 \sigma^2(\underline{\alpha}_1) + x_2^2 \sigma^2(\underline{\alpha}_2) + x_3^2 \sigma^2(\underline{\alpha}_3) + x_4^2 \sigma^2(\underline{\alpha}_4).$$

Vullen wij de opgegeven varianties in, dan is dus:

$$\sigma^2(\underline{d}) = x_1^2 + 4x_2^2 + 4x_3^2 + x_4^2. \quad (2.9)$$

Het totale kapitaal bestaat uit 10.000 coupures van honderd gulden, dus:

$$x_1 + x_2 + x_3 + x_4 = 10.000. \quad (2.10)$$

In B en C mag niet meer dan 60% belegd worden, dus:

$$x_2 + x_3 \leq 6.000. \quad (2.11)$$

Het kwadratische programmeringsprobleem luidt dus:

$$\text{minimaliseer } f = x_1^2 + 4x_2^2 + 4x_3^2 + x_4^2 \quad (2.12)$$

$$\text{onder de restricties: } x_1 + x_2 + x_3 + x_4 = 10.000 \quad (2.10)$$

$$x_2 + x_3 \leq 6.000 \quad (2.11)$$

$$4x_1 + 5x_2 + 6x_3 + 3x_4 \geq 45.000 \quad (2.8)$$

$$x_1, x_2, x_3, x_4 \geq 0.$$

Opmerking

Het probleem kan gemakkelijk realistischer gemaakt worden door bijv. koerswaarden in te voeren en de voorwaarde van onafhankelijkheid van $\underline{\alpha}_1, \dots, \underline{\alpha}_4$ te laten vallen. Dit heeft alleen tot gevolg dat de constanten van het probleem en de functie f (de totale variantie) iets veranderen.

c. Regressie-analyse

Een rijksbureau voor verkeersveiligheid wil een tabel publiceren, die bij gegeven aequidistante snelheden de lengte van de remweg van een luxe auto uit de zogenaamde middenklasse aangeeft. Door middel van proeven, verricht met auto's uit de Rijksautomobielen-centrale, heeft men het volgende waarnemingsmateriaal verzameld: $(x_1, y_1), \dots, (x_n, y_n)$, waarin x_i de snelheid in km/uur en y_i de gemeten remweg in m aangeeft. De beambte, belast met de samenstelling van de tabel, probeert door middel van een lineair verband tussen de verwachting van de remweg bij gegeven snelheid en de snelheid de tabel op te stellen. De voorwaardelijke verwachting van y hangt dus als volgt van x af:

$$\mathcal{L}(\underline{y}|\underline{x}) = \alpha + \beta x \quad (2.13)$$

Met behulp van de methode der kleinste kwadraten kunnen nu schattingen voor α en β gevonden worden door

$$f = \sum_{i=1}^n (y_i - \alpha - \beta x_i)^2 \quad (2.14)$$

te minimaliseren als functie van α en β (vgl. ook voorbeeld 2.4 uit hoofdstuk V).

Het ligt voor de hand aan te nemen dat de remweg niet afneemt bij toenemende snelheden. De richtingscoëfficiënt β zal dus moeten voldoen aan

$$\beta \geq 0. \quad (2.15)$$

Na bepaling van α en β kan (2.13) voor de samenstelling van de gevraagde tabel gebruikt worden. α en β volgen uit het kwadratische programmeringsprobleem: minimaliseer (2.17) onder de voorwaarde (2.15).

f kan ook als volgt geschreven worden:

$$\begin{aligned} f &= \sum_{i=1}^n (y_i^2 - 2\alpha y_i - 2\beta x_i y_i + \alpha^2 + 2\alpha\beta x_i + \beta^2 x_i^2) \\ &= n\alpha^2 + \beta^2 \sum_{i=1}^n x_i^2 + 2\alpha\beta \sum_{i=1}^n x_i - 2\alpha \sum_{i=1}^n y_i - 2\beta \sum_{i=1}^n x_i y_i + \sum_{i=1}^n y_i^2 \\ &= n\alpha^2 + A\beta^2 + 2B\alpha\beta + 2C\alpha + 2D\beta + E. \end{aligned}$$

Probleem (2.14), (2.15) luidt dan:

$$\text{Minimaliseer } f = n\alpha^2 + A\beta^2 + 2B\alpha\beta + 2C\alpha + 2D\beta + E \quad (2.16)$$

$$\text{onder de restrictie: } \beta \geq 0. \quad (2.17)$$

3. Definities en hulpstellingen

Een functie van één variabele noemt men convex (concaaf) als iedere koorde van de grafiek $y=f(x)$ boven (onder) de grafiek ligt ten opzichte van de x -as (het samenvallen van de koorde met de grafiek is ook toegestaan). Als voorbeeld kunnen de bekende parabolen uit de schoolwiskunde dienen.

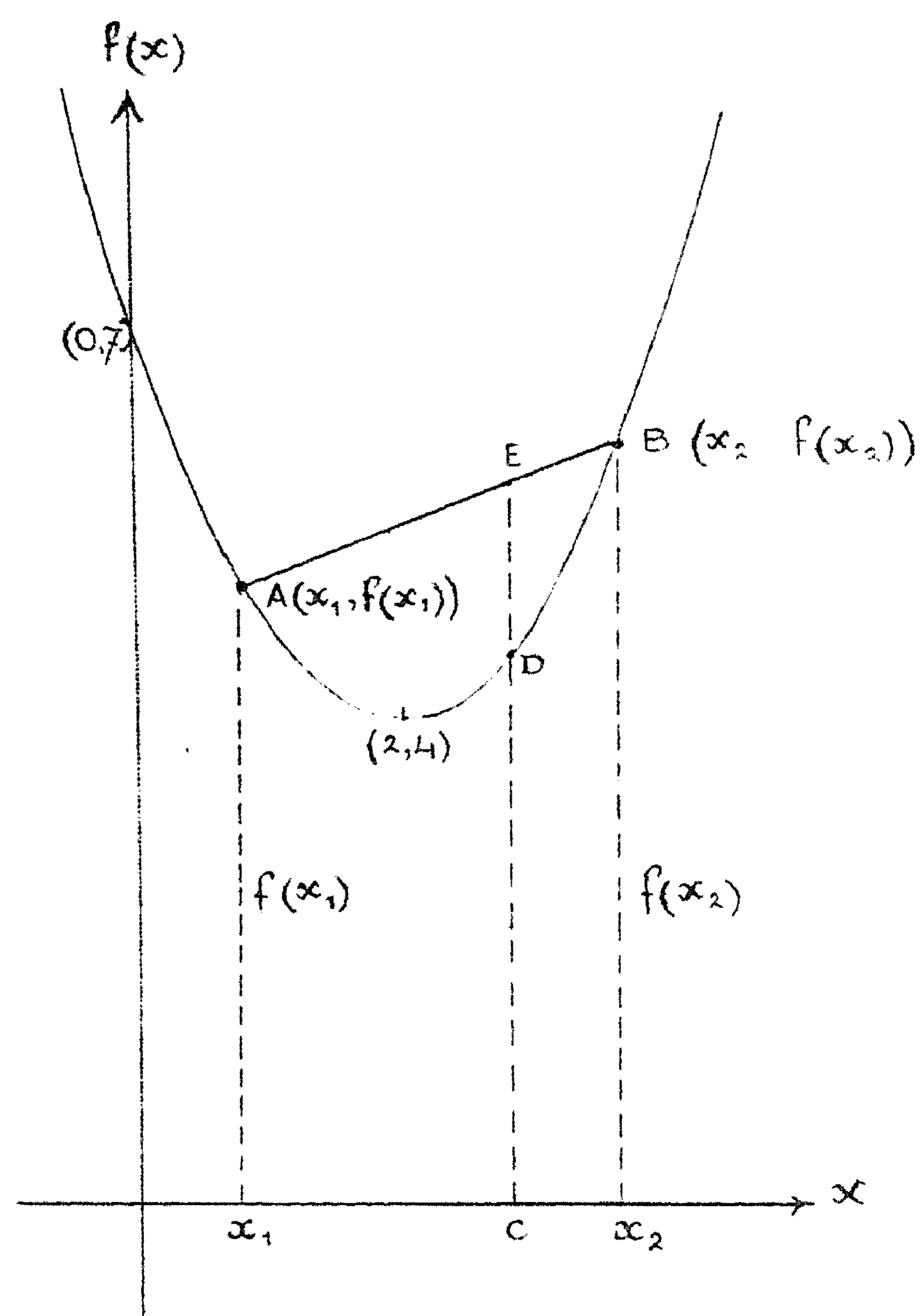


fig. 3.1

$$f(x) = \frac{3}{4}x^2 - 3x + 7$$

$f(x)$ is convex

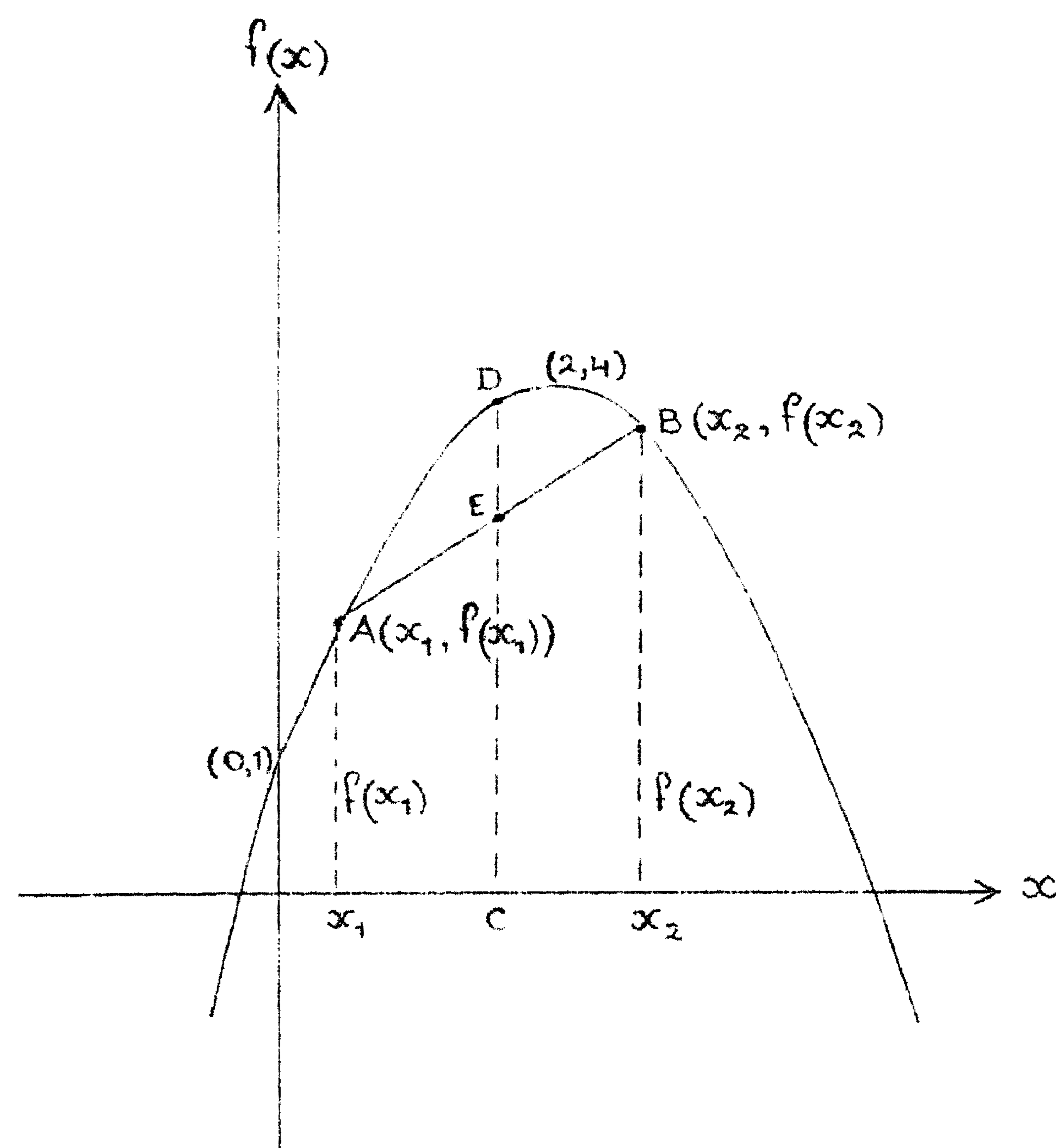


fig. 3.2

$$f(x) = -\frac{3}{4}x^2 + 3x + 1$$

$f(x)$ is concaaf

Ieder punt E op de koorde AB ($E \neq A$ of B) heeft tot coördinaten $(\overline{1-\lambda}x_1 + \lambda x_2, \overline{1-\lambda}f(x_1) + \lambda f(x_2))$ ¹⁾, $0 < \lambda < 1$. Aangezien in fig. 3.1 $EC \geq DC$, $DC = f(\overline{1-\lambda}x_1 + \lambda x_2)$ en $EC = \overline{1-\lambda}f(x_1) + \lambda f(x_2)$, geldt voor $0 < \lambda < 1$ en $x_1 \neq x_2$

$$(1-\lambda)f(x_1) + \lambda f(x_2) \geq f(\overline{1-\lambda}x_1 + \lambda x_2) . \quad (3.1)$$

Uit fig. 3.1 blijkt, dat de meetkundige definitie van het begrip convexiteit en de inhoud van formule (3.1) equivalent zijn.

Analoog volgt uit fig. 3.2 voor $0 < \lambda < 1$ en $x_1 \neq x_2$

$$(1-\lambda)f(x_1) + \lambda f(x_2) \leq f(\overline{1-\lambda}x_1 + \lambda x_2) . \quad (3.2)$$

Bij functies met meer variabelen definieert men de convexiteit (concaviteit) meestal met behulp van een formule analoog aan (3.1) resp. (3.2). Bijvoorbeeld $z = f(x, y)$ is convex als voor $0 < \lambda < 1$ en

1) Om haken uit te sparen wordt $(1-\lambda)$ soms geschreven als $\overline{1-\lambda}$.

$(x_1, y_1) \neq (x_2, y_2)$ geldt

$$(1-\lambda)f(x_1, y_1) + \lambda f(x_2, y_2) \geq f(\overline{1-\lambda}x_1 + \lambda x_2, \overline{1-\lambda}y_1 + \lambda y_2) . \quad (3.3)$$

Meetkundig drukt (3.3) uit dat ieder punt E van de verbindings-
koorde AB - $A(x_1, y_1, f(x_1, y_1))$, $B(x_2, y_2, f(x_2, y_2))$ - boven of op het
oppervlak $z = f(x, y)$ ligt t.o.v. het XY-vlak.

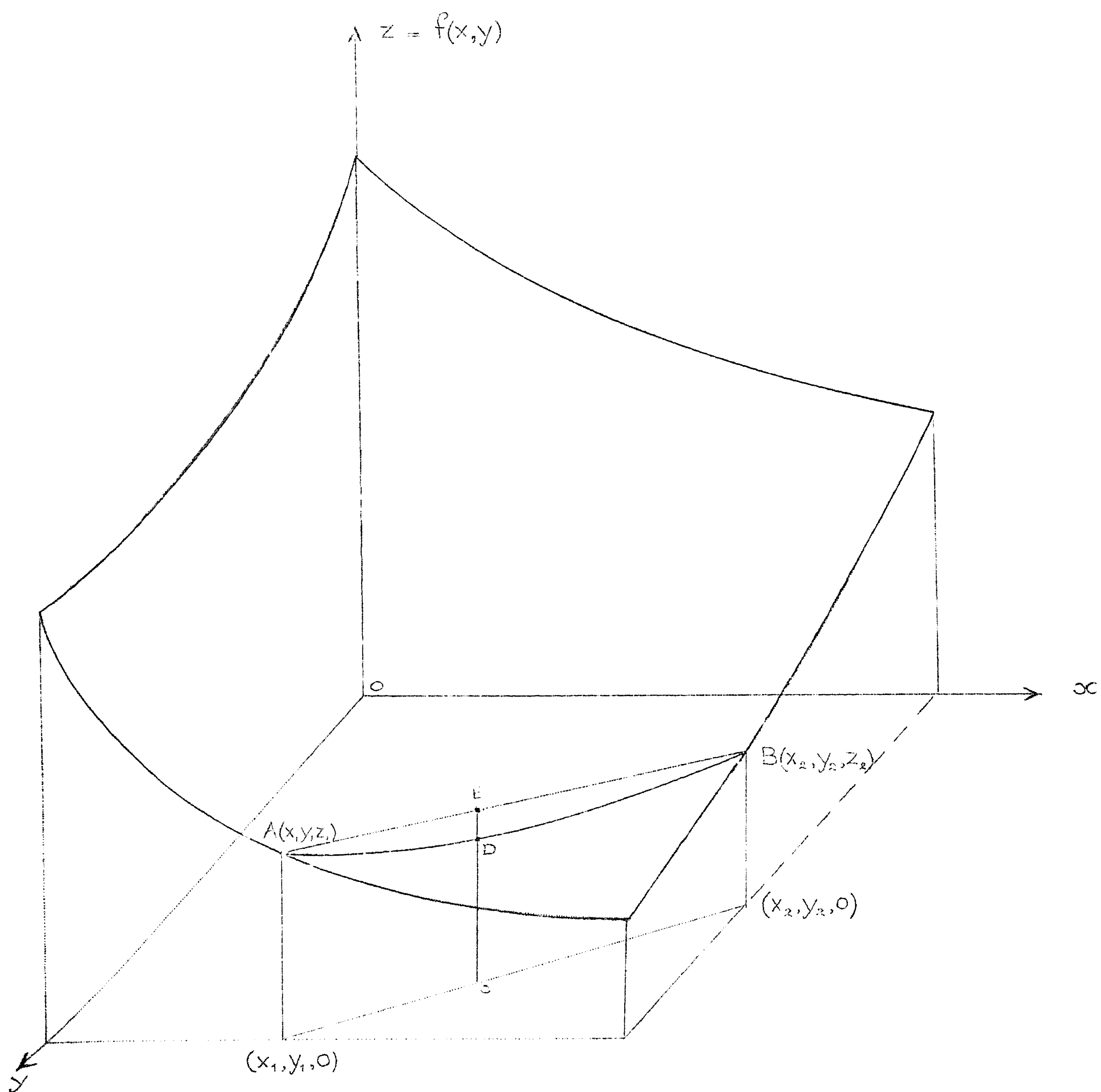


fig. 3.3

Een convexe functie van twee variabelen

De algemene definitie luidt nu:

Definitie 3.1. Een functie $f(X)$ gedefinieerd voor alle waarden van $x = (x_1, \dots, x_n)$ is:

$$\text{convex als } f(\overline{1-\lambda}X + \lambda\bar{X}) \leq (1-\lambda)f(X) + \lambda f(\bar{X}) \quad (3.4)$$

$$\text{concaaf als } f(\overline{1-\lambda}X + \lambda\bar{X}) \geq (1-\lambda)f(X) + \lambda f(\bar{X}) \quad (3.5)$$

geldt voor iedere $X = (x_1, \dots, x_n)$ en $\bar{X} = (\bar{x}_1, \dots, \bar{x}_n)$, $X \neq \bar{X}$ en $0 < \lambda < 1$.

De functie $f(X)$ is strikt convex (strikt concaaf) als onder bovenstaande voorwaarden alleen het ongelijktteken geldt.

Verder volgt uit bovenstaande definitie:

- a) Een lineaire functie is convex en concaaf.
- b) De som van twee convexe (concave) functies is weer convex (concaaf).
- c) Als $f(X)$ (strikt) convex is, dan is $-f(X)$ (strikt) concaaf en vice versa.

Het bewijs van deze eigenschappen kan eenvoudig door substitutie in (3.4) resp. (3.5) geleverd worden.

Definitie 3.2. Een symmetrische $n \times n$ -matrix is positief definit als voor iedere $X \neq 0$ geldt:
 $X.C.X > 0$,
 respectievelijk semi-definit positief als voor iedere $X \neq 0$ geldt $X.C.X \geq 0$.

Voor het positief definit zijn van een matrix bestaan diverse criteria, die hier echter niet besproken zullen worden.

Definitie 3.3. Een symmetrische $n \times n$ -matrix is indefinit als er minstens één paar vectoren X en Y , $X \neq Y$, bestaat zodat $X.C.X > 0$ en $Y.C.Y < 0$.

Voorbeeld 3.1:

$$\mathcal{A}_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix} \quad \mathcal{A}_2 = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \quad \mathcal{A}_3 = \begin{pmatrix} -1 & -1 \\ -1 & -1 \end{pmatrix} \quad \mathcal{A}_4 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$\mathcal{A}_1$ is positief definit, want $X \mathcal{A}_1 X = (x_1 x_2 x_3) \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = x_1^2 + 2x_2^2 + 3x_3^2$ en dus is $X \mathcal{A}_1 X > 0$. Voor iedere $X = (x_1, x_2, x_3) \neq 0$.

$\mathcal{A}_2$ is semi-definit positief, want $X \mathcal{A}_2 X = (x_1 x_2 x_3) \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = (x_1 + x_2 + x_3)^2$ en dus is $X \mathcal{A}_2 X \geq 0$; immers een kwadraat is altijd niet negatief, terwijl de waarde nul ook aangenomen kan worden, bijv. voor $X = (0, 2, -2)$.

$\mathcal{A}_3$ is semi-definit negatief, want $X \mathcal{A}_3 X = (x_1 x_2) \begin{pmatrix} -1 & -1 \\ -1 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = -(x_1 + x_2)^2$ en dus is $X \mathcal{A}_3 X \leq 0$; $X \mathcal{A}_3 X = 0$ voor bijv. $X = (3, -3)$.

$\mathcal{A}_4$ is indefinit, want $X \mathcal{A}_4 X = (x_1 x_2 x_3) \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = x_1^2 - x_2^2 + x_3^2$;

$X \mathcal{A}_4 X < 0$ voor bijv. $X = (0, 4, 0)$, $X \mathcal{A}_4 X > 0$ voor bijv.

$X = (2, 0, -3)$ en $X \mathcal{A}_4 X = 0$ voor bijv. $X = (3, -3, 0)$.

Stelling 3.1. De functie $g(X) = X \cdot C \cdot X$, waarin C een symmetrische $n \times n$ -matrix voorstelt, is strikt convex als C positief definit is. Is C semi-definit positief, dan is $g(X)$ convex.

Bewijs

Stel $X \neq Y$ en $0 < \lambda < 1$. Nu geldt:

$$\frac{g((1-\lambda)X + \lambda Y) - (1-\lambda)g(X) - \lambda g(Y)}{(1-\lambda)g(X) + \lambda g(Y)} = \frac{(1-\lambda)^2 X \cdot C \cdot X + \lambda(1-\lambda) X \cdot C \cdot Y + \lambda(1-\lambda) Y \cdot C \cdot X + \lambda^2 Y \cdot C \cdot Y}{(1-\lambda) X \cdot C \cdot X + \lambda Y \cdot C \cdot Y}$$

$$\begin{aligned} g((1-\lambda)X + \lambda Y) - (1-\lambda)g(X) - \lambda g(Y) &= -\lambda(1-\lambda) X \cdot C \cdot X + \lambda(1-\lambda) X \cdot C \cdot Y + \lambda(1-\lambda) Y \cdot C \cdot X - \lambda(1-\lambda) Y \cdot C \cdot Y \\ &= -\lambda(1-\lambda) X \cdot C \cdot (X - Y) + \lambda(1-\lambda) Y \cdot C \cdot (X - Y) \\ &= -\lambda(1-\lambda) (X - Y) \cdot C \cdot (X - Y) < 0 \end{aligned}$$

Immers $0 < \lambda < 1$, $\lambda(1-\lambda) > 0$ en $(X - Y) \cdot C \cdot (X - Y) > 0$ (want $X \cdot C \cdot X > 0$, X willekeurig).

Voor het geval C semi-definit positief is, volgt het gestelde uit bovenstaande redenering, als men onder het ongelijkteken een $=$ teken plaatst.

Voorbeeld 3.2:

1) Neem $C = \left(\frac{1}{4}\right)$, $g(X) = x \cdot \frac{1}{4}x = \frac{1}{4}x^2$. C is positief definit, terwijl uit fig. 3.4 blijkt dat $g(X)$ convex is. Hetzelfde voor:

2) $C = \begin{pmatrix} \frac{1}{4} & 0 \\ 0 & \frac{1}{4} \end{pmatrix}$, $g(X) = \frac{1}{4}x^2 + \frac{1}{4}y^2$; zie fig. (3.5).

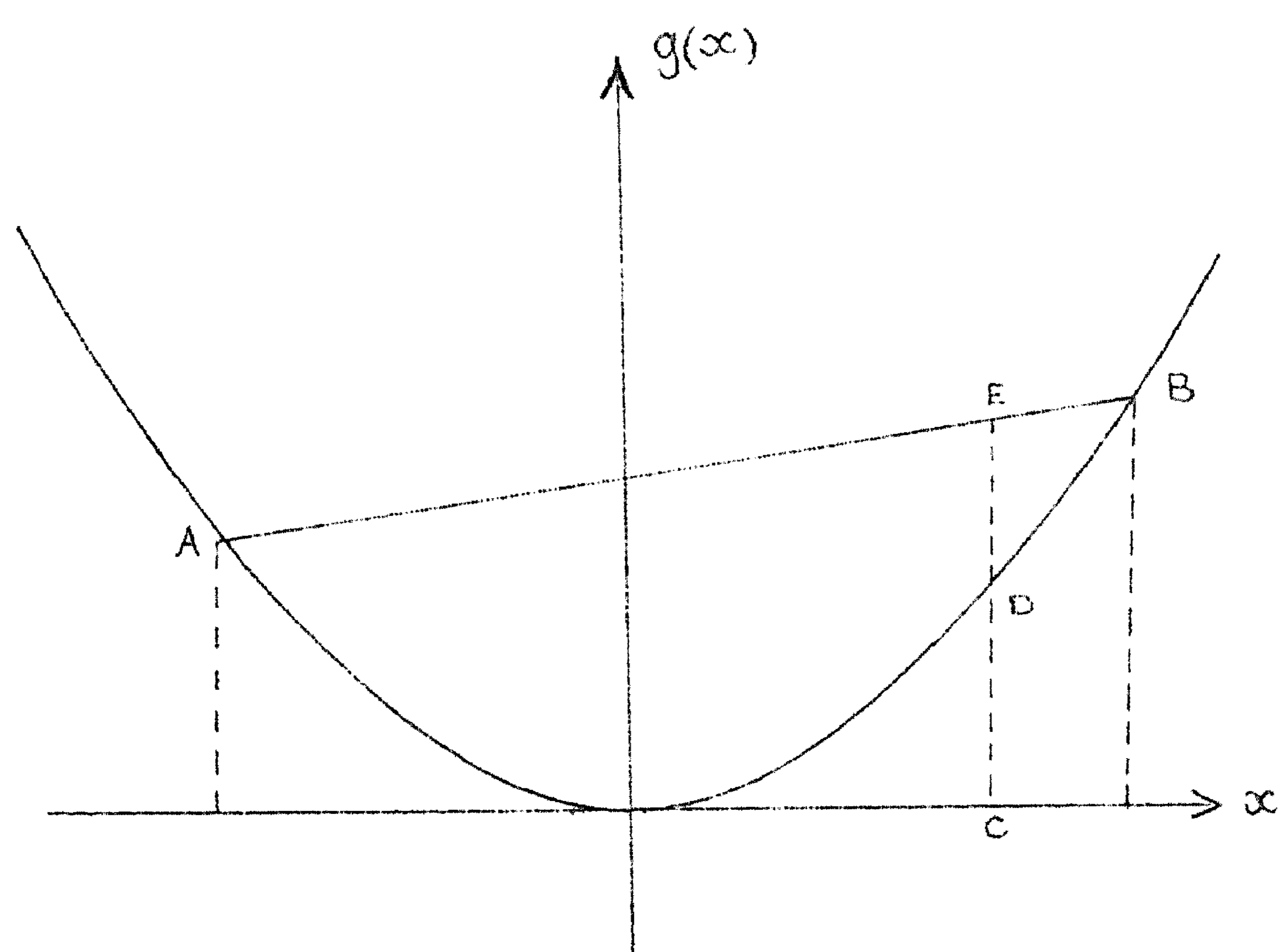


fig. 3.4

$$g(x) = \frac{1}{4}x^2$$

$g(x)$ is convex

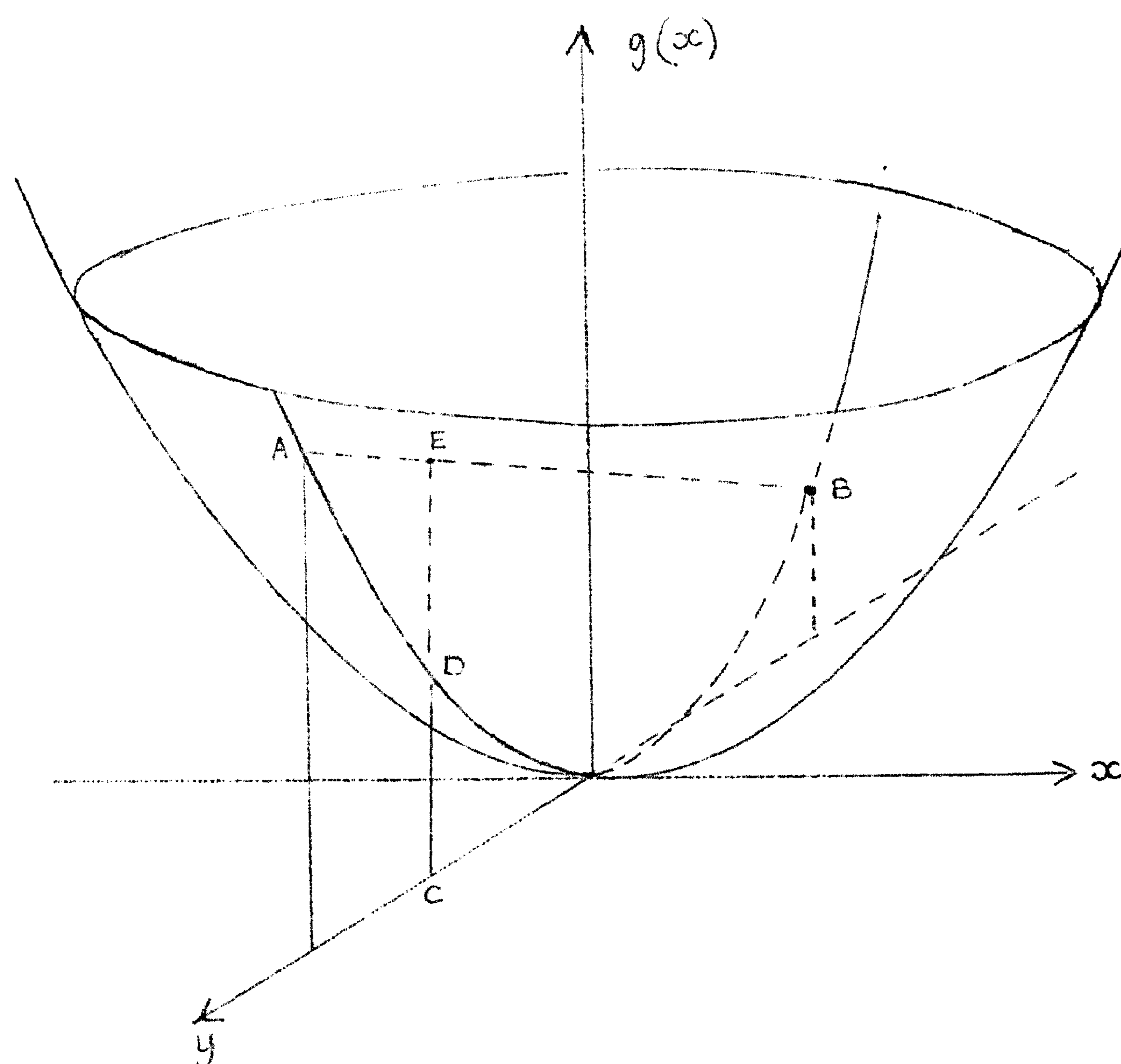


fig. 3.5

$$g(x) = \frac{1}{4}x^2 + \frac{1}{4}y^2$$

$g(x)$ is convex

De letters A, B, C, D en E hebben dezelfde betekenis als in fig. 3.1 t/m 3.3.

Neemt men in het eerste voorbeeld i.p.v. $g(X) = \frac{1}{4}x^2$ $h(X) = -\frac{1}{4}x^2$, dan ziet men gemakkelijk in, dat $h(X)$ concaaf en de bijbehorende "matrix" $\left(-\frac{1}{4}\right)$ negatief definit is.

Stelling 3.2. Een lokaal minimum van een convexe functie $f(X)$ gedefinieerd voor alle waarden van $X = (x_1, \dots, x_n)$ is een absoluut minimum.

Toelichting

Dat een lokaal minimum niet steeds een absoluut minimum is, valt gemakkelijk in te zien voor een functie van één variabele; zie fig. 3.6. Hierin is A een lokaal minimum, maar geen absoluut minimum, want er bestaat nog een ander lokaal minimum B, dat lager is;

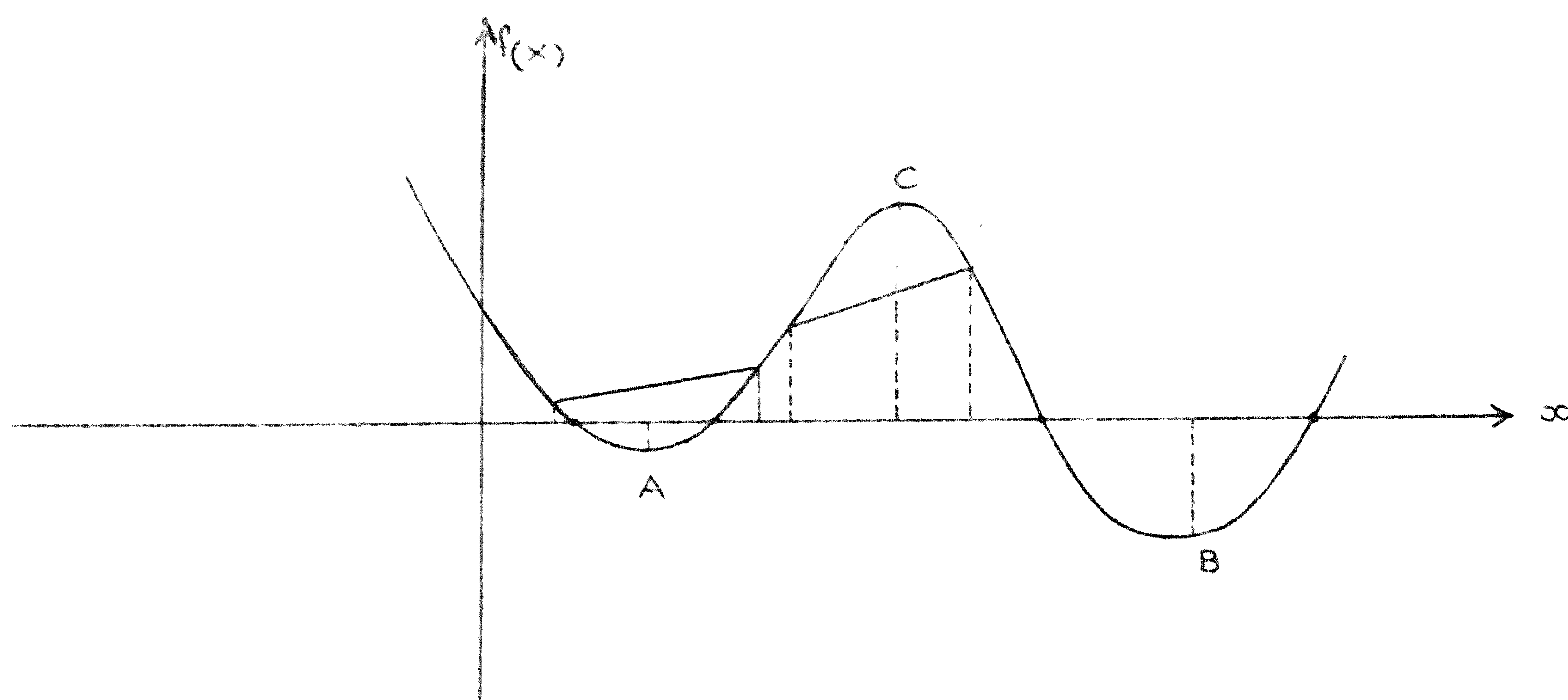


fig. 3.6

$f(x)$ is soms convex, soms concaaf

B is wel een absoluut minimum. Uit fig. 3.6 blijkt ook dat $f(x)$ niet overal convex is, zodat deze functie niet voldoet aan de voorwaarde van de stelling. Functies, die wel voldoen aan de voorwaarde van de stelling zijn die uit de figuren 3.1 en 3.3. Het bewijs van de stelling luidt:

Stel $f(x)$ neemt een lokaal minimum m aan in $\bar{x}$. Voor $m = -\infty$ is de stelling evident. In het geval $m > -\infty$ is er bij een willekeurige x altijd een λ , $0 < \lambda < 1$, te vinden, zodat geldt (want $\bar{x}$ is een lokaal minimum):

$$m \leq f((1-\lambda)\bar{x} + \lambda x) \leq (1-\lambda)f(\bar{x}) + \lambda f(x) = (1-\lambda)m + \lambda f(x)$$

$$\text{of } m = f(\bar{x}) \leq f(x) \text{ voor willekeurige } x.$$

Bovenstaand bewijs gaat ook op, als $\bar{X}$ in een convex gebied R (zie blz. 14 en 15, hoofdstuk XII) ligt. Immers als $\bar{X}$ en X in R liggen, dan ligt $(1-\lambda)\bar{X} + \lambda X$, $0 < \lambda < 1$, ook in R . Hieruit volgt, dat het locale minimum $f(\bar{X})$ in een convex gebied het absolute minimum is.

4. Het algoritme van BEALE

4.1 Inleiding

Het minimum van een functie van meerdere variabelen kan - als er geen bijzondere complicaties optreden - met behulp van de differentiaalrekening vrij eenvoudig berekend worden (zie par. 6 hoofdstuk IV). Als de variabelen echter alleen maar niet-negatieve waarden mogen aannemen en bovendien aan een aantal lineaire restricties moeten voldoen, dan kan het voorkomen, dat het bovengenoemde minimum aangenomen wordt in een punt buiten het toegelaten gebied. Het algoritme van BEALE geeft nu een methode aan, hoe in dit geval een toegelaten minimum berekend kan worden. Er wordt echter steeds verondersteld dat de functie convex is (vgl. definitie 3.1).¹⁾ Laten we, om de gedachten te bepalen, aannemen dat we het minimum moeten berekenen van een convexe differentieerbare functie van twee variabelen; dus van een functie $y=f(x_1, x_2)$.

Als $\frac{\partial f}{\partial x_1}$ en $\frac{\partial f}{\partial x_2}$ beide nul zijn voor $X = (0,0)$, dan neemt $f(x_1, x_2)$ zijn minimum aan voor $X = (0,0)$ (zie ook figuur 4.1).²⁾

Zijn echter beide - of één der - partiële afgeleiden niet nul, dan kan $f(0,0)$ verminderd worden door een variatie vanuit nul van de x_1 , waarvoor geldt: $\frac{\partial f}{\partial x_1} \neq 0$, hetgeen gemakkelijk blijkt uit fig. 4.1. Ook als x_1 en x_2 niet-negatieve variabelen zijn, kan f in dit geval nog verminderd worden door een variatie van x_2 van uit nul.

Bovenstaande gedachtengang, waarbij wij eerst onderzoeken of de waarde van de functie in 0 verminderd kan worden en zo ja deze vermindering trachten te bereiken door slechts één variabele te ver-

1) Voor een toelichting van de noodzakelijkheid van deze voorwaarde zie de opmerkingen 4.1 en 4.2 aan het einde van deze inleiding.

2) Het op blz. 39 van hoofdstuk II genoemde voorbeeld betrof een functie, die niet convex was in het punt $X = (0,0)$; zie ook opmerking 4.1.

anderen, zal nu uitgebreid en toegelicht worden met behulp van enkele eenvoudige voorbeelden.

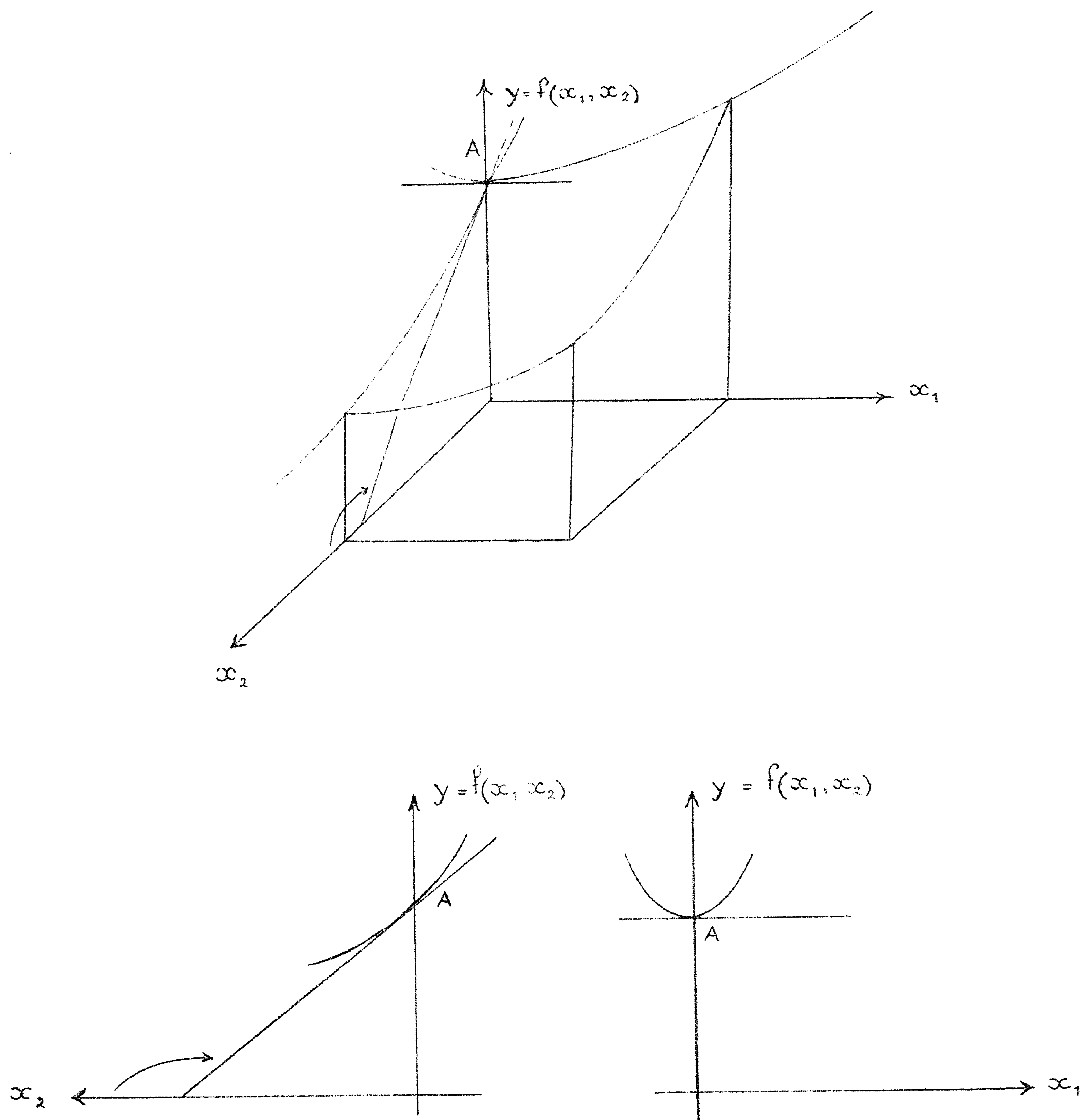


fig. 4.1

$f(x_1, x_2)$ kan door een kleine variatie van x_2 ($x_1=0$)
verminderd worden

Voorbeeld 4.1

Bepaal het minimum van $y = f(x_1, x_2) = (x_1 - 16)^2 + (x_2 + 4)^2$.

Met behulp van de differentiaalrekening kan het minimum zonder meer (zie hoofdstuk II blz. 39) uit:

$$\frac{1}{2} \frac{\partial f}{\partial x_1} = x_1 - 16 = 0 \quad (4.1)$$

$$\frac{1}{2} \frac{\partial f}{\partial x_2} = x_2 + 4 = 0 \quad (4.2)$$

berekend worden. Het minimum nul wordt dus voor $x_1=16$ en $x_2=-4$ aangenomen. Aangezien f bestaat uit een som van twee kwadraten is dit antwoord ook zo wel in te zien.

De berekening met behulp van bovengenoemd principe verloopt als volgt: bepaal eerst de waarde van de functie en de partiële afgeleiden in $X=(0,0)$: $f(0,0)=272$, $\frac{1}{2} \frac{\partial f}{\partial x_1} = -16$, $\frac{1}{2} \frac{\partial f}{\partial x_2} = +4$. Zolang nu u_1 ,

$$u_1 = \frac{1}{2} \frac{\partial f}{\partial x_1} = x_1 - 16, \quad (4.3)$$

negatief blijft, kan f verminderd worden door een toename van x_1 vanuit nul, waarbij $x_2=0$ gehouden wordt; x_1 kan groeien tot 16, want $u_1=0$ voor $x_1=16$. Vervolgens elimineren wij x_1 uit f met behulp van (4.3). Uit (4.3) volgt:

$$x_1 = u_1 + 16, \quad (4.4)$$

zodat substitutie in y leidt tot:

$$f' = u_1^2 + (x_2 + 4)^2. \quad (4.5)$$

De functie $f(x_1, x_2)$ neemt voor $x_1=16$ en $x_2=0$ dezelfde waarde aan, als de functie $f(u_1, x_2)$ voor $u_1=x_2=0$, nl. 16. Door de substitutie van (4.4) in f kan nu op f' weer hetzelfde procédé toegepast worden, als reeds op f is toegepast. Wij berekenen de waarde van f' en de partiële afgeleiden $\frac{\partial f'}{\partial u_1}$ en $\frac{\partial f'}{\partial x_2}$ in het punt $u_1=0$, $x_2=0$ en vinden resp. $f'(0,0)=16$, $\frac{1}{2} \left(\frac{\partial f'}{\partial u_1} \right)_{u_1=0, x_2=0} = 0$, $\frac{1}{2} \left(\frac{\partial f'}{\partial x_2} \right)_{u_1=0, x_2=0} = 4$. f kan dus vermin-

derd worden door een negatieve variatie van x_2 vanuit nul tot -4, want:

$$u_2 = \frac{1}{2} \frac{\partial f'}{\partial x_2} = x_2 + 4 \quad (4.6)$$

en dus blijft u_2 positief zolang $x_2 > -4$ is.

Na substitutie van $x_2 = u_2 - 4$ in f' gaat f' over in

$$f'' = u_1^2 + u_2^2. \quad (4.7)$$

De waarde van f'' voor $u_1=u_2=0$ is nul, terwijl ook de partiële afgeleiden van f'' naar u_1 en u_2 in dit punt nul zijn. f'' kan dus niet meer verminderd worden en dus volgt uit (4.4) en (4.6), dat $\min f''(u_1, u_2) = \min f(x_1, x_2) = 0$ voor $u_1=u_2 = 0$ of $x_1=16$, $x_2=-4$.

Voorbeeld 4.2

Bepaal het minimum van $y = f(x_1, x_2) = (x_1-16)^2 + (x_2+4)^2$ onder de restricties: $x_1, x_2 \geq 0$.

De methode uit de differentiaalrekening kan niet toegepast worden, aangezien (zie voorbeeld 4.1) de hierna gevonden oplossing nu ontoelaatbaar is. ($x_2=-4$ is negatief!)

De redenering volgens de methode van BEALE loopt tot $f' = u_1^2 + (x_2+4)^2$ - formule (4.5) - parallel aan voorbeeld 4.1. f' kan echter nu alleen verminderd worden door een negatieve variatie van x_2 , die echter hier niet toelaatbaar is. Het minimum van f is dus onder de restricties $x_1, x_2 \geq 0$ gelijk aan 16 en wordt aangenomen voor $x_1=16$ ($u_1=0$) en $x_2=0$.

Voorbeeld 4.3

Bepaal het minimum van $f = (x_1-16)^2 + (x_2+4)^2$ onder de restricties: $x_1+x_2 \leq 8$
 $x_1, x_2 \geq 0$.

Dit voorbeeld kan ook geschreven worden als:

Minimaliseer $f(x_1, x_2) = (x_1-16)^2 + (x_2+4)^2$ (4.8)

onder de restricties: $x_3 = 8 - x_1 - x_2$ (4.9)

$x_1, x_2, x_3 \geq 0$. (4.10)

De multiplicatorenmethode van Lagrange (zie hoofdstuk II, blz. 40) kan hier niet toegepast worden, aangezien de oplossing niet voldoet aan (4.10). Dit blijkt gemakkelijk uit:

$$F = (x_1-16)^2 + (x_2+4)^2 + \lambda(x_1+x_2+x_3-8) \quad \text{en:} \quad (4.11)$$

$$\left. \begin{aligned} \frac{\partial F}{\partial x_1} &= 2(x_1-16) + \lambda = 0 \\ \frac{\partial F}{\partial x_2} &= 2(x_2+4) + \lambda = 0 \\ \frac{\partial F}{\partial x_3} &= \lambda = 0 \\ \frac{\partial F}{\partial \lambda} &= x_1 + x_2 + x_3 - 8 = 0 \end{aligned} \right\} \quad (4.12)$$

De oplossing van (4.12) luidt: $x_1=16$, $x_2=-4$ en $x_3=-4$ en voldoet niet aan (4.10).

Uit voorbeeld 4.1 volgt dat f alleen verminderd kan worden door een (positieve) variatie van x_1 vanuit nul. Bij het laten toenemen van x_1 moet er op gelet worden dat:

$$1^\circ \quad u_1 = \frac{1}{2} \frac{\partial f}{\partial x_1} = x_1 - 16 \quad \text{niet positief wordt.}$$

$$2^\circ \quad x_3 = 8 - x_1 \quad (x_2=0 \text{ bij de variatie!}) \quad \text{niet negatief wordt.}$$

Men ziet gemakkelijk in, dat x_3 eerder nul wordt - nl. voor $x_1=8$ - dan u_1 . Daarom lossen we nu x_1 op uit (4.9); dus:

$$x_1 = 8 - x_2 - x_3 \quad (4.13)$$

en elimineren x_1 uit (4.8). Aangezien de grootte en het teken van de variabele u_1 niet van belang zijn voor het probleem, kan u_1 verder weggelaten worden. Het probleem (4.8), (4.9) en (4.10) kan dan vervangen worden door:

$$\text{Minimaliseer } f' = (x_2 + x_3 + 8)^2 + (x_2 + 4)^2 \quad (4.14)$$

$$\text{onder de restricties: } x_1 = 8 - x_2 - x_3 \quad (4.15)$$

$$x_1, x_2, x_3 \geq 0. \quad (4.16)$$

In het punt $x_2=x_3=0$ is de functiewaarde van f' gelijk aan 80, terwijl

$$\frac{1}{2} \left(\frac{\partial f'}{\partial x_2} \right)_{x_2=x_3=0} = 12 \quad \text{en} \quad \frac{1}{2} \left(\frac{\partial f'}{\partial x_3} \right)_{x_2=x_3=0} = 8 \quad \text{is. Beide partiële af-}$$

geleiden zijn dus positief, zodat f' niet verminderd kan worden door variatie van x_2 of x_3 (beide variabelen zijn niet-negatief!). Het minimum van f wordt dus bereikt voor $x_1=8$, $x_2=x_3=0$ en bedraagt 80.

Bovenstaande berekeningen zijn gemakkelijk meetkundig te interpreteren, als men opmerkt, dat de krommen $(x_1-16)^2 + (x_2+4)^2 = \text{constant} = C > 0$ cirkels voorstellen met middelpunt $(16, -4)$ en straal $R=\sqrt{C}$.

De meetkundige formulering van de drie voorbeelden luidt dan: Tracht een cirkel met middelpunt $(16, -4)$ en een zo klein mogelijke straal te vinden, waarvan minstens één der punten in een vooraf bepaald toegelaten gebied moet liggen.

In voorbeeld 4.1 - zie figuur 4.2 - wordt de x_1 eerst tot 16 ge-

1) $x_1=x_2=0$ en $x_3=8$ levert direct een toegelaten oplossing van stelsel (4.9) en (4.10) op.

varieerd, waardoor wij in het punt E komen, en daarna constant gehouden. Vervolgens wordt x_2 gevarieerd tot -4, waardoor het punt F bereikt wordt.

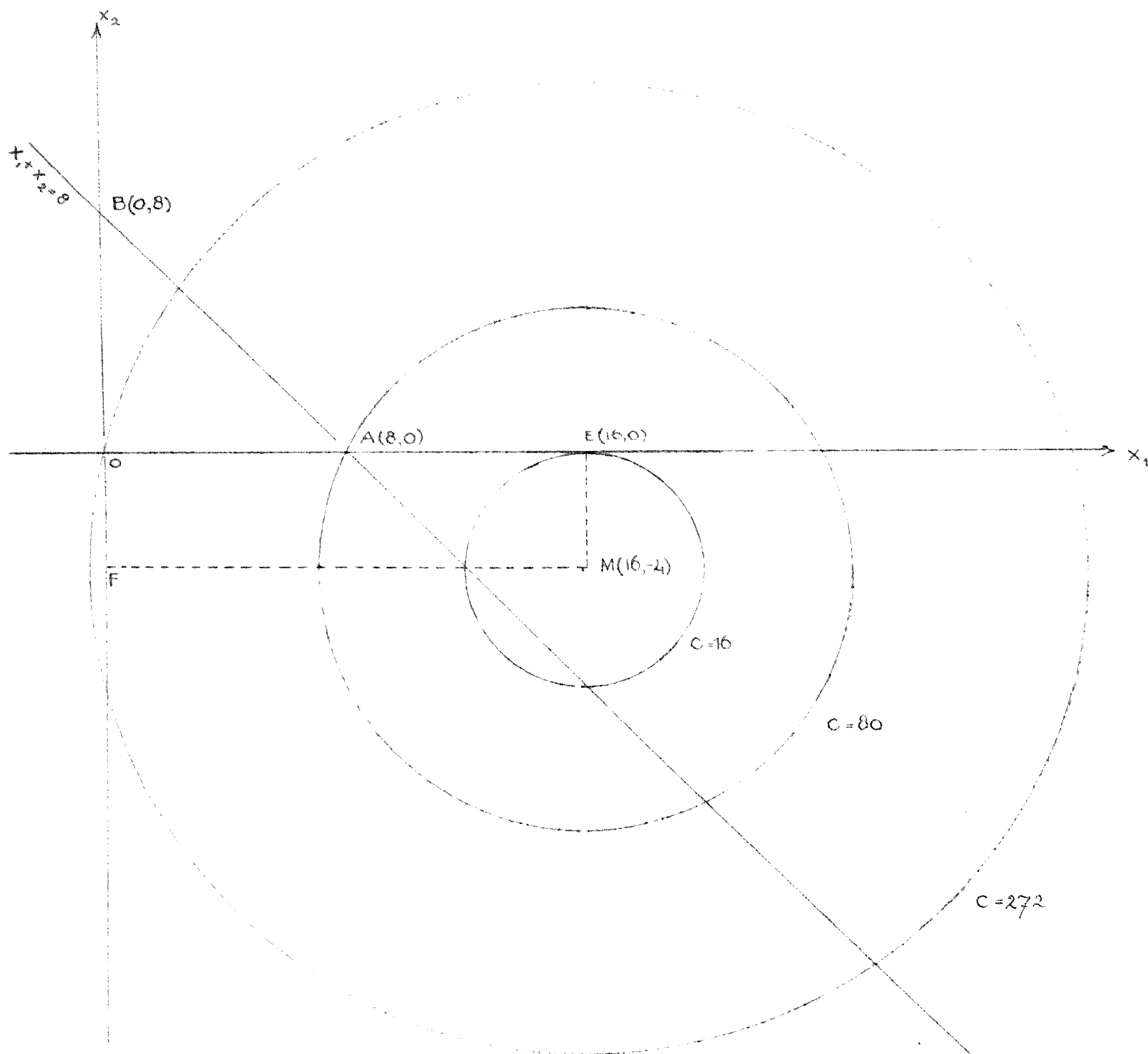


fig. 4.2

Voorbeeld 4.1.	toegelaten gebied:	gehele vlak	; minimum $f=0$
" 4.2.	"	kwadrant $x_1 O x_2$	; minimum $f=16$
" 4.3.	"	OAB	; minimum $f=80$

Het minimum wordt dus bereikt via O, E, M en de oplossing is een nulcirkel met middelpunt $(16, -4)$ (een nulcirkel is een cirkel met straal nul). In voorbeeld 4.2 wordt x_1 gevarieerd tot 16. Verdere vermindering van f is niet mogelijk, zoals meetkundig gemakkelijk is in te zien. In voorbeeld 4.3 kan bij variatie van x_1 niet verder gegaan worden dan tot punt A. Een grotere variatie zou tot gevolg hebben, dat men buiten het toegelaten gebied OAB komt.

Opmerkingen

4.1 Als de differentieerbare functie f niet convex is, en alle partiële afgeleiden in $X=(0,0)$ nul zijn, dan kan f soms toch wel verminderd worden door een kleine variatie van één der variabelen. Hieruit volgt dat in het algemeen het nul zijn van de partiële afgeleiden in 0 wel een noodzakelijke maar geen voldoende voorwaarde is voor het minimaal zijn van f in 0. Zie voor verdere toelichting fig. 4.3 en paragraaf 6 van hoofdstuk II.

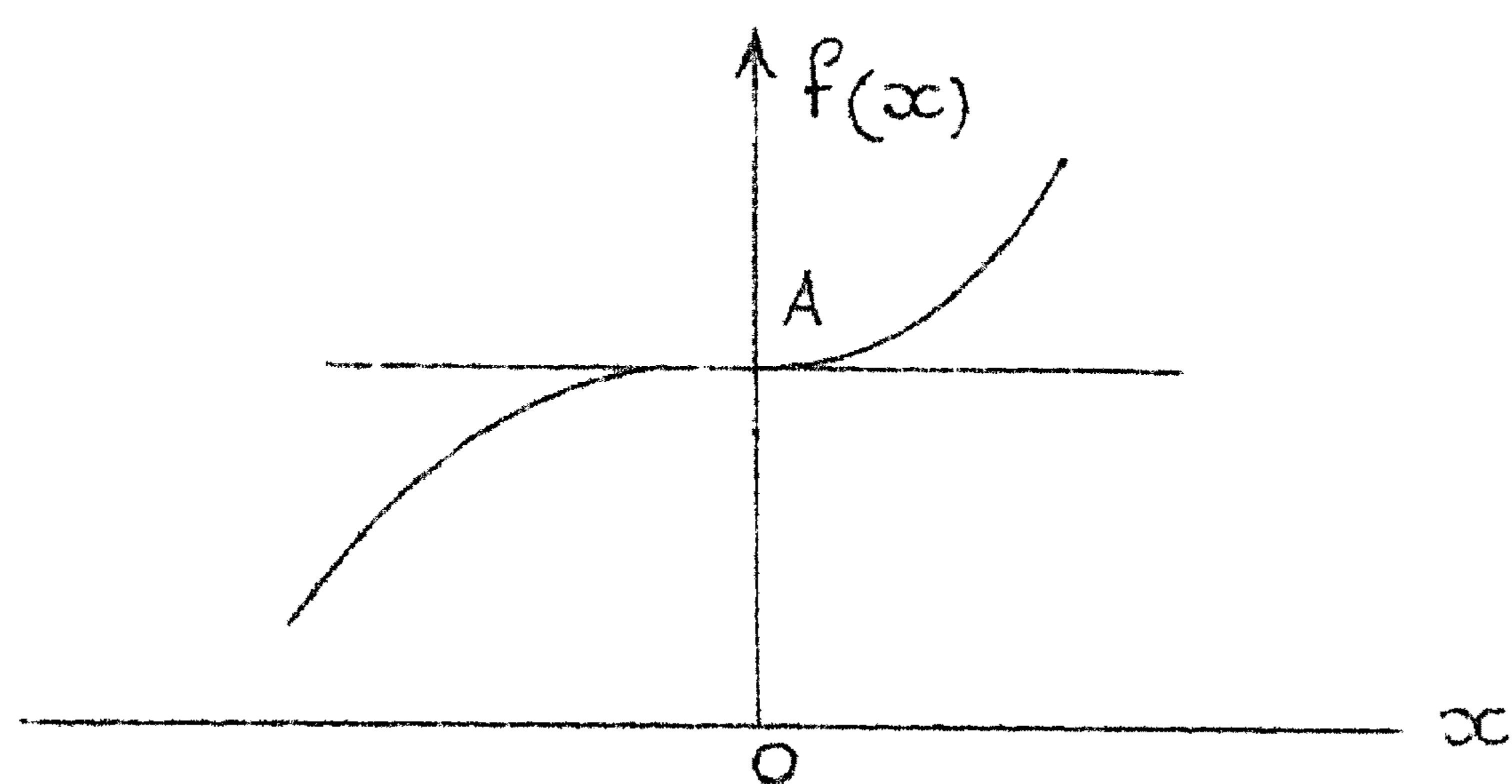


fig. 4.3

$\frac{\partial f}{\partial x} = 0$ en
 $f(0)$ is buigpunt

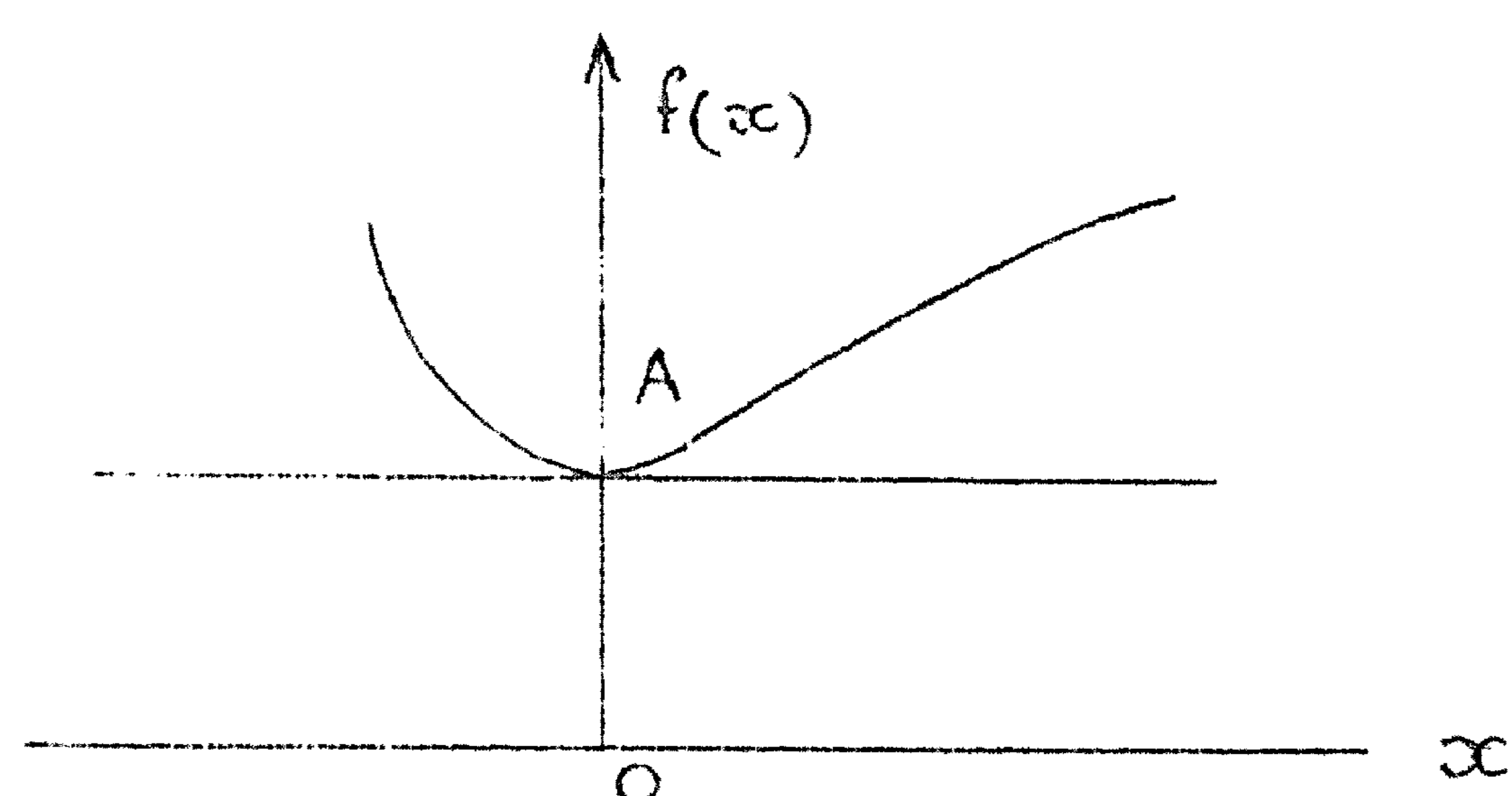


fig. 4.4

$\frac{\partial f}{\partial x} = 0$ en
minimum $f = f(0)$

4.2 Algemeen geldt: Een differentieerbare convexe functie is minimaal in 0, als alle partiële afgeleiden in 0 nul zijn; zie fig. 4.4 en 4.5.

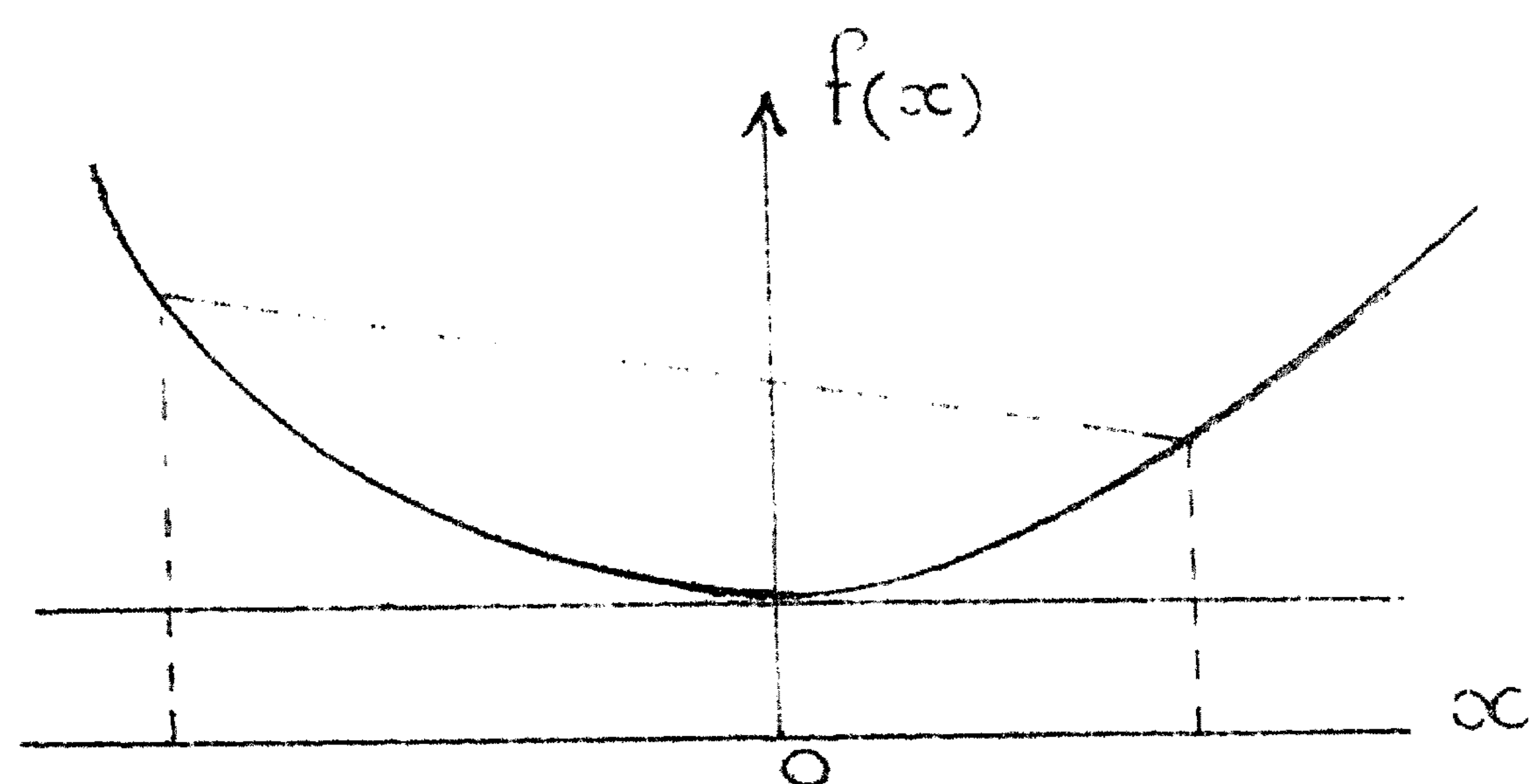


fig. 4.5

$f(x)$ is convex en
 $f(0)$ is minimum van $f(x)$

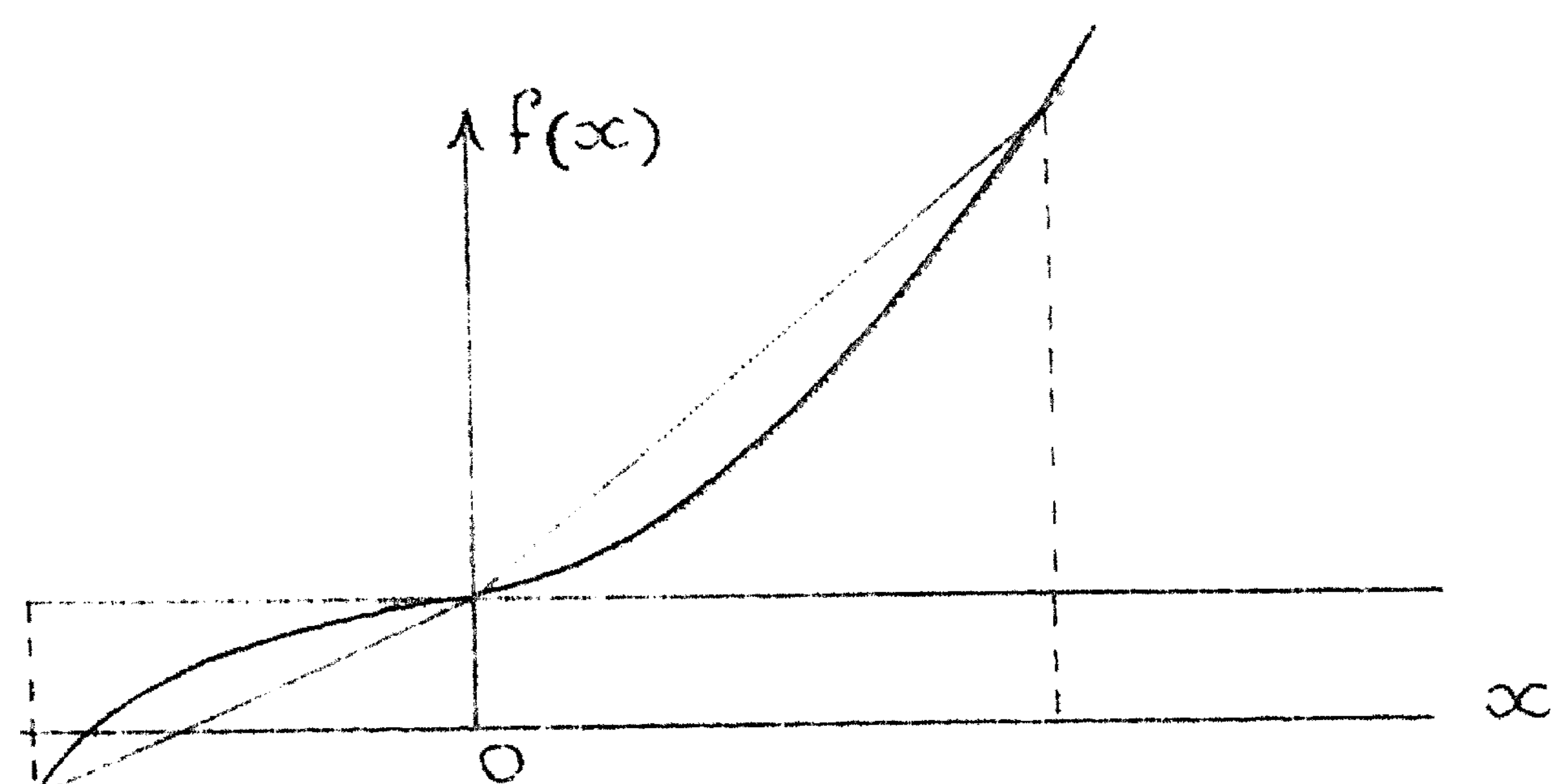


fig. 4.6

$f(x)$ is niet convex en
 $f(0)$ is niet het minimum van $f(x)$

4.3 Zijn echter de variabelen niet-negatief, dan behoeven de partiële afgeleiden in σ niet nul te zijn, als de differentieerbare convexe functie f minimaal is in σ ; zie fig. 4.6 en fig. 4.7.

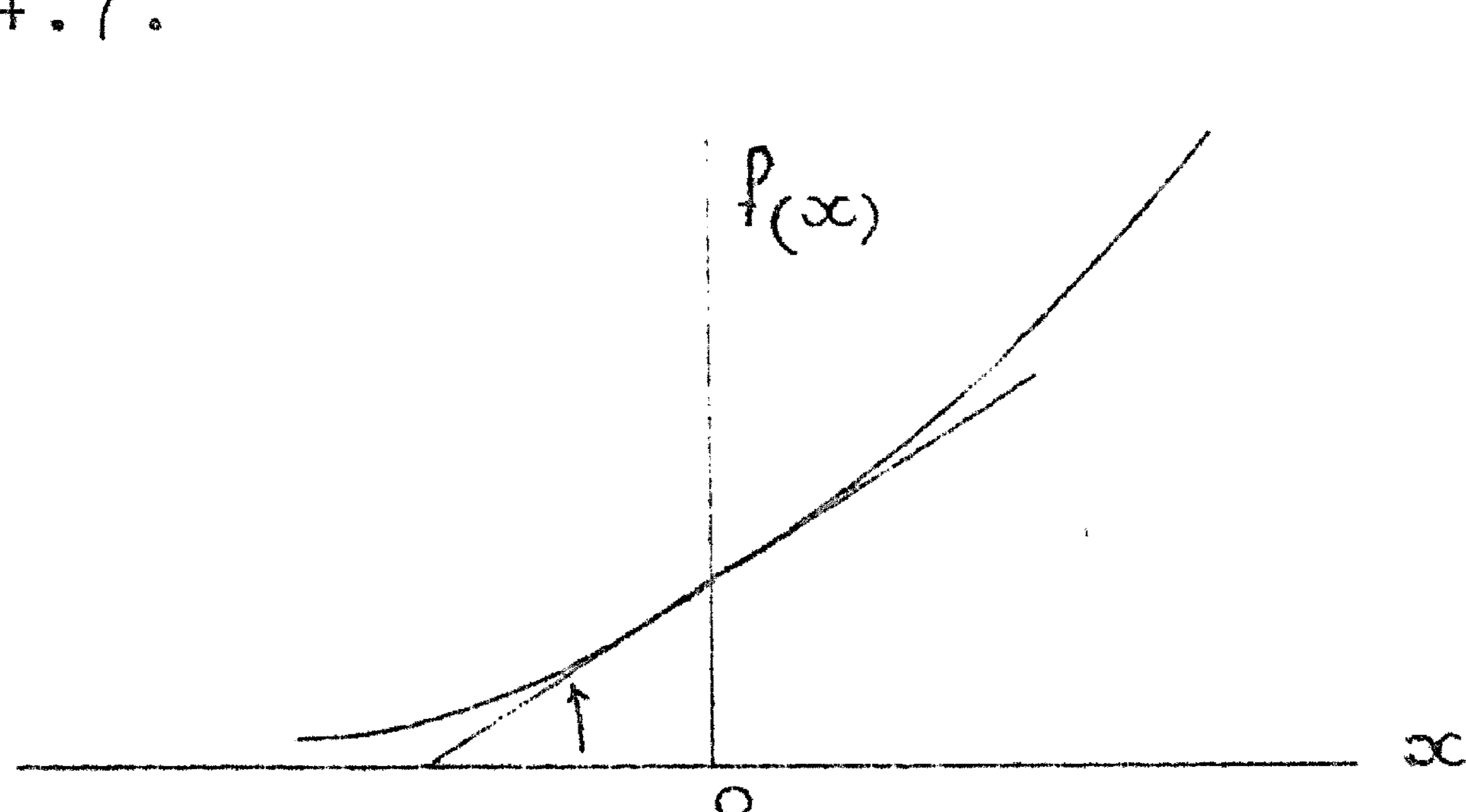


fig. 4.7

$\frac{\partial f}{\partial x} > 0; x \geq 0$ en
 $f(0)$ is minimum van $f(x)$

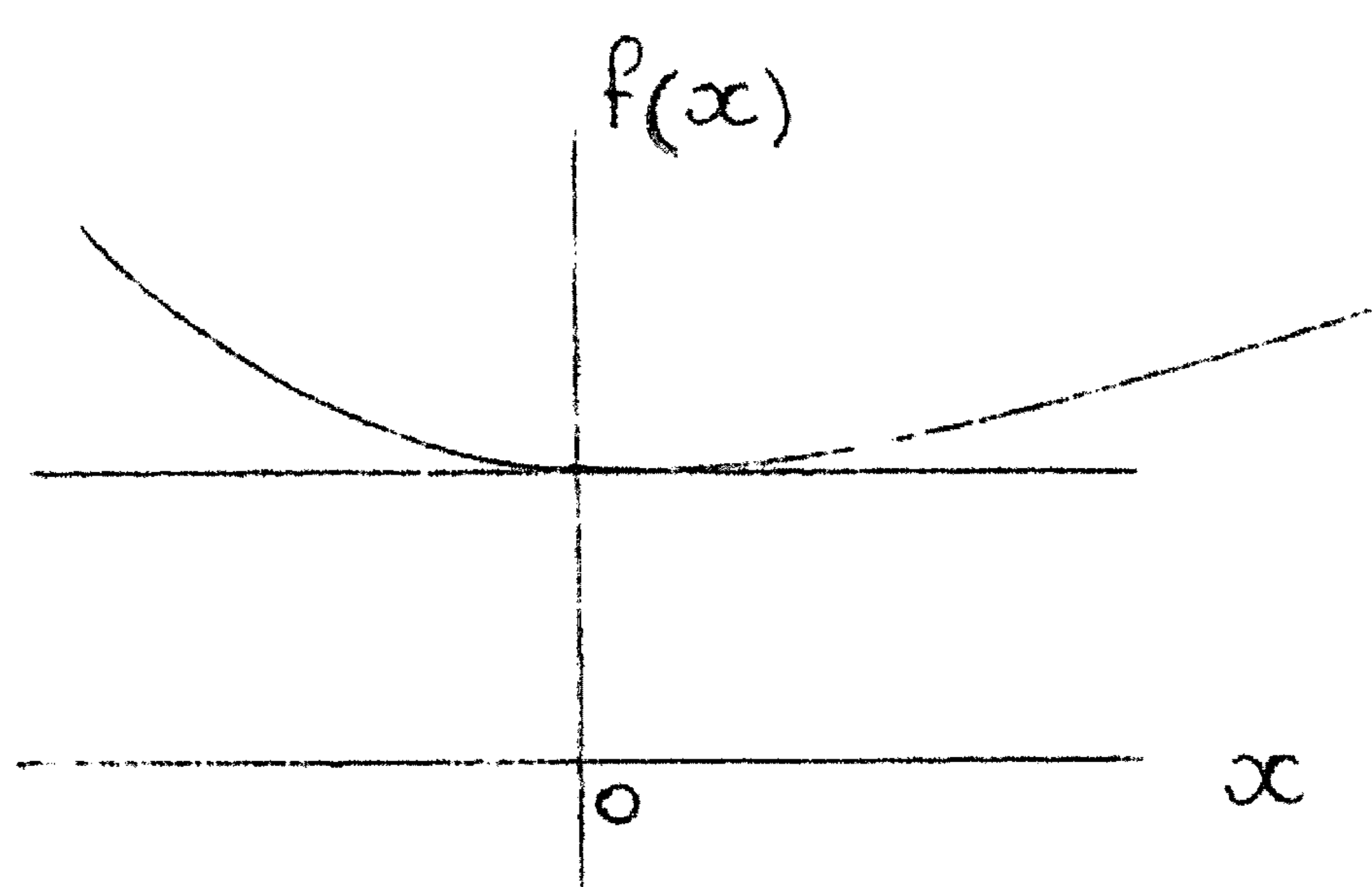


fig. 4.8

$\frac{\partial f}{\partial x} = 0, x \geq 0$ en
 $f(0)$ is minimum van $f(x)$

4.2 Notatie

Voor de beschrijving van het algoritme wordt het probleem als volgt genoteerd:

$$\text{Minimaliseer } f(X) = \sum_{i,j=0}^n \bar{c}_{ij} x_i x_j \quad x_0=1 \quad (4.17)$$

$$\text{onder de nevenvoorwaarden } \sum_{j=1}^n a_{ij} x_j = a_{i0} \quad i=1,2,\dots,m \quad (a_{i0} \geq 0) \quad (4.18)$$

$$x_j \geq 0 \quad j=1,2,\dots,n. \quad (4.19)$$

Evenals bij de simplexmethode nemen we aan, dat de laatste m kolommen van de matrix $A = (a_{ij})$ een éénheidsmatrix vormen. Mocht dit niet het geval zijn, dan voeren we verschilvariabelen in (zie hoofdstuk XII). Vergelijking (4.18) kan dus geschreven worden als:

$$\begin{aligned} x_{n-m+1} &= a_{10} - a_{11}x_1 - \dots - a_{1n-m}x_{n-m} \\ x_{n-m+2} &= a_{20} - a_{21}x_1 - \dots - a_{2n-m}x_{n-m} \\ &\vdots \\ x_n &= a_{m0} - a_{m1}x_1 - \dots - a_{mn-m}x_{n-m} \end{aligned} \quad (4.20)$$

Met behulp van de m -vergelijking (4.20) kunnen we de functie f (zie 4.17) schrijven als een functie van de variabelen x_1 t/m x_{n-m} . Als we bovendien de notatie als volgt wijzigen:

$$\begin{array}{llll} x_0, x_1, \dots, x_m & \text{noteren als} & z_0, z_1, \dots, z_m \\ x_{n-m+1}, \dots, x_n & " & " & x_1, x_2, \dots, x_m \\ -a_{11}, \dots, -a_{m1} & " & " & a_{11}, \dots, a_{m1} \\ - - - - - & & & - - - - - \\ -a_{1n-m}, \dots, -a_{mn-m} & " & " & a_{1n-m}, \dots, a_{mn-m} \end{array}$$

dan kan stelsel (4.17) t/m (4.19) geschreven worden als:

$$\text{Minimaliseer } f(X) = \sum_{i,j=0}^{n-m} c_{ij} z_i z_j^{1)} \quad z_0=1 \quad (4.21)$$

$$\text{onder de restricties } x_h = a_{h0} + \sum_{j=1}^{n-m} a_{hj} z_j \quad (4.22)$$

$$\begin{aligned} x_h, \quad a_{h0} &\geq 0 & h=1, 2, \dots, m \\ z_j &\geq 0. & j=1, 2, \dots, n-m \end{aligned} \quad (4.23)$$

Uit de voorbeelden 4.1, 4.2 en 4.3 bleek al, dat de functie f en de nevenvoorwaarden (4.18) en (4.19) aan het einde van iedere iteratie in de vorm (4.21) en (4.22) geschreven kunnen worden. Zie bijv. (4.5) en (4.14), (4.15), (4.16). Bovendien kunnen er onder de z -variabelen ook variabelen, die alle reële waarden kunnen aannemen - de z.g. vrije variabelen - voorkomen. Zie bijv. u_1 in voorbeeld 4.5.

1) Door de bovengenoemde substituties m.b.v. (4.20) in f uit te voeren kan men gemakkelijk verifiëren dat weer geldt $c_{ij} = c_{ji}$.

De x-variabelen noemt men ook wel de basisvariabelen en de z-variabelen de niet-basisvariabelen.

4.3. Beschrijving algoritme

De basisvergelijkingen zijn (4.21), (4.22) en (4.23). Voor $z_1=0, i=1,2,\dots,n-m$ heeft men een toegelaten basisoplossing $x_h=a_{h0}, h=1,2,\dots,m$, van (4.18) en $f=c_{00}$.

De afgeleide van f naar z_1 voor $z_1=z_2=\dots=z_{n-m}=0$ bepaalt men door f uit te schrijven en daarna naar z_1 te differentiëren.

$$\begin{aligned} f = \sum_{i,j=0}^{n-m} c_{ij} z_i z_j &= \sum_{i=0}^{n-m} (c_{i0} z_i + \dots + c_{ii} z_i^2 + \dots + c_{i,n-m} z_i z_{n-m}) \\ &= \begin{matrix} c_{00} & + & c_{01} z_1 & + \dots + & c_{0i} z_i & + \dots + & c_{0,n-m} z_{n-m} \\ \vdots & & \vdots & & \vdots & & \vdots \\ c_{i0} z_i & + & c_{ii} z_i^2 & + \dots + & c_{i,n-m} z_i z_{n-m} \\ \vdots & & \vdots & & \vdots & & \vdots \\ c_{n-m,0} z_{n-m} & + & c_{n-m,i} z_{n-m} z_i & + \dots + & c_{n-m,i} z_{n-m} z_i & + \dots + & c_{n-m,n-m} z_{n-m}^2 \end{matrix} \end{aligned} \quad (4.24)$$

$$\frac{\partial f}{\partial z_i} = \begin{matrix} c_{0i} \\ c_{ii} z_i \\ \vdots \\ c_{n-m,i} z_{n-m} \end{matrix} + c_{i0} + c_{ii} z_i + \dots + 2c_{ii} z_i + \dots + c_{i,n-m} z_{n-m} \quad (4.25)$$

Formule (4.25) kan, daar alle termen twee keer voorkomen - $c_{ij} = c_{ji}$ - ook geschreven worden als:

$$\frac{\partial f}{\partial z_i} = 2c_{i0} + 2 \sum_{j=1}^{n-m} c_{ij} z_j \quad (4.26)$$

De genoemde afgeleide is dus $2c_{i0}$. We definiëren nu:

$$u_i = \frac{1}{2} \frac{\partial f}{\partial z_i} = c_{i0} + \sum_{j=1}^{n-m} c_{ij} z_j \quad (4.27)$$

Stel nu $c_{r0} < 0$; dit betekent dat f afneemt door een kleine toename van z_r , als we de overige z-variabelen nul houden (zie figuur 4.9). Zolang dus $c_{r0} + c_{rr} z_r$ (is $\frac{1}{2} \frac{\partial f}{\partial z_r}$ voor $z_j=0, j \neq r$) negatief blijft, zal f door toename van z_r afnemen. Dit blijkt

gemakkelijk uit figuur 4.9. Zie verder ook fig. 4.1, 4.3 t/m 4.8.

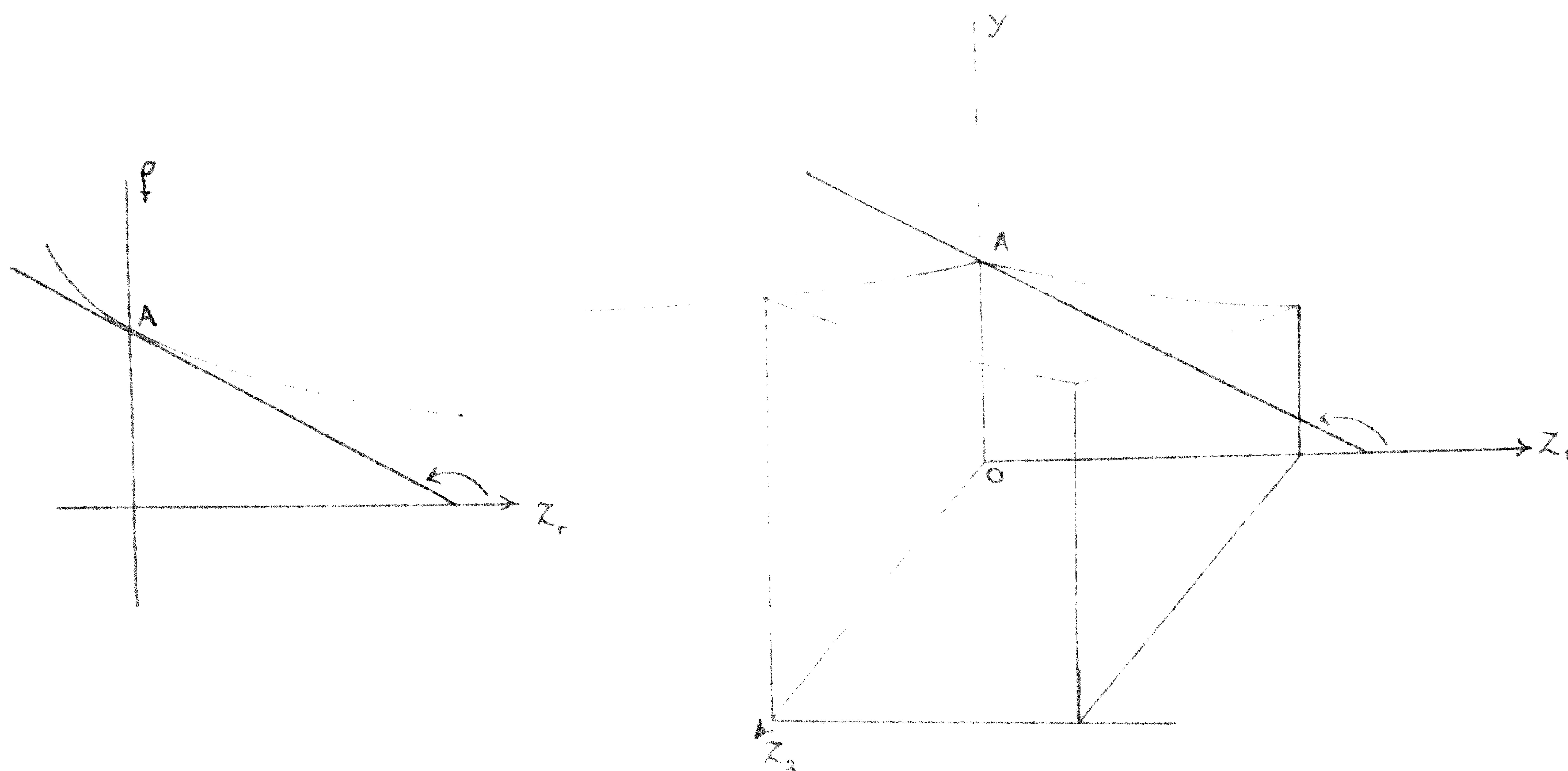


fig. 4.9

$$y = f(z_1, z_2), \quad z_r = z_1; \quad f \text{ neemt af daar } \frac{-\partial f}{\partial z_r} = c_{r0} < 0$$

$$\text{voor } z_1 = z_2 = \dots + z_{n-m} = 0$$

Stel nu dat er een negatieve c_{10} ($i \neq 0$), bijv. c_{r0} , bestaat. Mochten er nu meerdere $c_{10} < 0$ ($i \neq 0$) bestaan, dan ligt het voor de hand de meest negatieve te kiezen. Immers meetkundig - zie fig. (4.9) - is c_{10} de richtingscoëfficiënt van de raaklijn aan de doorsnijdingskromme van het oppervlak $y=f(z_1, z_2)$ en het z_1Oy -vlak. Hoe meer negatief c_{10} is, des te steiler loopt de genoemde raaklijn in A (fig. 4.9), en des te groter is de vermindering van f bij een kleine variatie van z_1 . In formule luidt deze selectieregel:

$$\text{Kies } \min_{c_{10} < 0} c_{10} = \max_{c_{10} < 0} |c_{10}|. \text{ Zijn er meerdere } c_{10} \text{ die hieraan}$$

voldoen dan neemt men hiervan de c_{10} met de kleinste i -index.

Er kunnen zich bij de variatie van z_r vanuit 0 naar $+\infty$ drie gevallen voordoen, in het vervolg aangegeven door resp. A, B en C. Voor de bespreking van deze gevallen hebben we de vergelijkingen (4.28) en (4.29)

$$u_r = c_{r0} + c_{rr}z_r \quad (4.28)$$

$$x_h = a_{h0} + a_{hr}z_r \quad h = 1, \dots, m \quad (4.29)$$

nodig. Deze ontstaan door in (4.22) en (4.27) $z_j=0$ voor $j \neq r$ te substitueren en voor i te kiezen $i=r$. De coëfficiënten a_{h0} en c_{rr} zijn altijd niet negatief. Voor a_{h0} volgt dit uit de veronderstellingen van het probleem (vgl. 4.18); tot $c_{rr} \geq 0$ concluderen we door op te merken dat een convexe functie in n -variabelen ook convex is in één variabele als we de overige variabelen nul stellen. In het

geval $f(x) = \sum_{i,j=0}^{n-m} c_{ij}z_i z_j$ moet dus $f = c_{00} + 2c_{r0}z_r + c_{rr}z_r^2$ convex zijn.

Hieruit volgt het gestelde.

A. $c_{r0} < 0$, $c_{rr} = 0$ en alle $a_{hr} \geq 0$.

Dit betekent dat bij variatie van z_r geen der vergelijkingen (4.28) en (4.29) nul kan worden.

Uit $f = c_{00} + 2c_{r0}z_r$ blijkt dat het minimum in dit geval $-\infty$ is.

B. Eén der basisvariabelen x_h is nul voordat u_r en de overige basisvariabelen nul worden.

De index p van de basisvariabele die nul wordt voordat $u_r < 0$ en de overige x -variabelen van teken veranderen, vindt men door het volgende betoog:

Uit (4.29) volgt, dat: $x_1=0$ voor $z_r = -\frac{a_{10}}{a_{1r}}$
 $x_m=0$ voor $z_r = -\frac{a_{m0}}{a_{mr}}$.

Aangezien z_r in positieve richting gevarieerd wordt, kunnen de x_i met $-\frac{a_{i0}}{a_{ir}} < 0$ nooit nul worden. Bovendien wordt de grootte van de variatie van z_r bepaald door de kleinste positieve waarde van $\frac{-a_{i0}}{a_{ir}}$; daar $a_{i0} \geq 0$ is behoeven we alleen maar naar de quotiënten met $a_{ir} < 0$ te kijken.

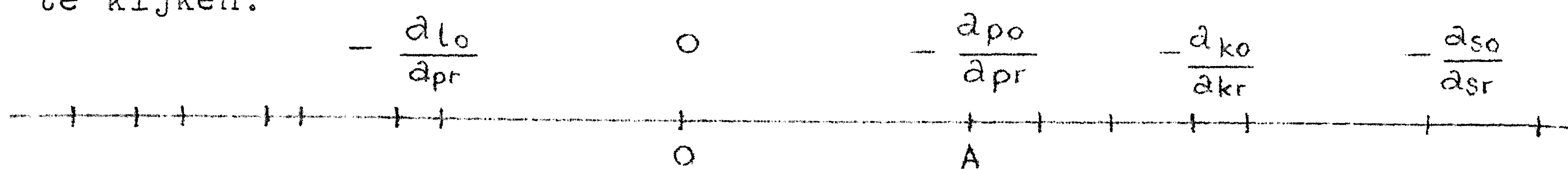


fig. 4.10

Het interval OA bepaalt de grootte van de variatie van z_r

Bovenstaand betoog kan samengevat worden door middel van formule (4.30).

$$\min_{a_{hr} < 0} \left(- \frac{a_{h0}}{a_{hr}} \right) = - \frac{a_{p0}}{a_{pr}} . \quad (4.30)$$

We voegen nu z_r aan de basisvariabelen en x_p aan de niet-basisvariabelen toe. Uit de p^{de} vergelijking van (4.22) lossen we z_r op.

$$z_r = - \frac{a_{p0}}{a_{pr}} + \frac{1}{a_{pr}} x_p - \sum_{j \neq r} \frac{a_{pj}}{a_{pr}} z_j . \quad (4.31)$$

Substitueren we voor z_r formule (4.31) in f , dan neemt f voor $z_1=z_2=\dots=z_{r-1}=z_{r+1}=\dots=z_{n-m}=0$, $x_p=0$ precies die waarde aan, die f bereikte vóór de substitutie van (4.31) voor $z_1=z_2=\dots=z_{r-1}=z_{r+1}=\dots=z_{n-m}=0$ en $z_r = - \frac{a_{p0}}{a_{pr}} .$

Als we (4.31) ook substitueren in de overige vergelijkingen van (4.22) en de variabelen herdopen als: $z'_1=z_1$, $z'_2=z_2$, ..., $z'_{r-1}=z_{r-1}$, $z'_r=x_p$, $z'_{r+1}=z_{r+1}$, ..., $z'_{n-m}=z_{n-m}$ en $x'_1=x_1$, ..., $x'_{p-1}=x_{p-1}$, $x'_p=z_r$, $x'_{p+1}=x_{p+1}$, ..., $x'_{n-m}=x_{n-m}$, dan zijn f en de nevenvoorwaarden weer in de vorm (4.21) en (4.22) gekomen.

De functie neemt echter voor $z'_i=0$ $i=1,2,\dots,n-m$ in het algemeen een kleinere waarde aan dan vóór de substitutie. Op deze nieuwe vorm van f en de restricties kunnen we weer eenzelfde iteratie toepassen (zie iteratie N) etc., totdat alle c_{i0} niet negatief zijn. In dit geval neemt f zijn minimumwaarde c_{00} aan voor $z_1=0$ en $x_h=a_{h0}$. Zie voor illustratie hiervan de voorbeelden 4.2 en 4.3.

De substituties van (4.31) in (4.21) en (4.22) verlopen aldus:

$$x_h = a_{h0} - \frac{a_{hr}}{a_{pr}} a_{p0} + \frac{a_{hr}}{a_{pr}} x_p + \sum_{j \neq r} (a_{hj} - a_{hr} \frac{a_{pj}}{a_{pr}}) z_j \quad h \neq p. \quad (4.32)$$

De (nieuwe) a_{ij} 's zijn dus:

$$a'_{hj} = a_{hj} - a_{hr} \frac{a_{pj}}{a_{pr}} \quad h \neq p, \quad j \neq r \quad (4.33)$$

$$a'_{pj} = - \frac{a_{pj}}{a_{pr}} \quad j \neq r \quad (4.34)$$

$$a'_{hr} = \frac{a_{hr}}{a_{pr}} \quad h \neq p \quad (4.35)$$

$$a'_{pr} = \frac{1}{a_{pr}} \quad (4.36)$$

Teneinde eenvoudige transformatieformules voor f te verkrijgen, schrijven we f weer als volgt:

$$f = (c_{00} + \dots + c_{0r}z_r + \dots + c_{0,n-m}z_{n-m}) + (c_{i0} + \dots + c_{ir}z_r + \dots + c_{i,n-m}z_{n-m})z_i + (c_{n-m,0} + \dots + c_{n-m,r}z_r + \dots + c_{n-m,n-m}z_{n-m})z_{n-m} \quad (4.37)$$

Eerst transformeren we nu z_i binnen de haken met (4.31).

$$f = \left\{ \left(c_{00} - \frac{a_{p0}}{a_{pr}} c_{0r} \right) + \dots + \frac{c_{0r}}{a_{pr}} x_p + \dots + \left(c_{0,n-m} - \frac{a_{p,n-m}}{a_{pr}} c_{0r} \right) z_{n-m} \right\} + \left\{ \left(c_{i0} - \frac{a_{p0}}{a_{pr}} c_{ir} \right) + \dots + \frac{c_{ir}}{a_{pr}} x_p + \dots + \left(c_{i,n-m} - \frac{a_{p,n-m}}{a_{pr}} c_{ir} \right) z_{n-m} \right\} z_i + \left\{ \left(c_{n-m,0} - \frac{a_{p0}}{a_{pr}} c_{n-m,r} \right) + \dots + \frac{c_{n-m,r}}{a_{pr}} x_p + \dots + \left(c_{n-m,n-m} - \frac{a_{p,n-m}}{a_{pr}} c_{n-m,r} \right) z_{n-m} \right\} z_{n-m} \quad (4.38)$$

Hieruit volgen de formules

$$c'_{ij} = c_{ij} - \frac{a_{pj}}{a_{pr}} c_{ir} \quad j \neq r \quad (4.39)$$

$$c'_{ir} = \frac{c_{ir}}{a_{pr}} \quad (4.40)$$

Na substitutie van deze waarden in (4.38) en vervanging van $z_1, \dots, z_{r-1}, x_p, z_{r+1}, \dots, z_{n-m}$, die binnen de haken staan, door $z'_1, \dots, z'_r, \dots, z'_{n-m}$, kunnen we (4.40) schrijven als:

$$f = (c'_{00} + c'_{01}z'_1 + \dots + c'_{0r}z'_r + \dots + c'_{0,n-m}z'_{n-m}) + (c'_{i0} + c'_{i1}z'_1 + \dots + c'_{ir}z'_r + \dots + c'_{i,n-m}z'_{n-m})z_i + (c'_{n-m,0} + c'_{n-m,1}z'_1 + \dots + c'_{n-m,r}z'_r + \dots + c'_{n-m,n-m}z'_{n-m})z_{n-m} \quad (4.41)$$

Door de haakjes weg te werken, en daarna de z'_i buiten haakjes te brengen wordt (4.41):

$$\begin{aligned}
 f = & (c'_{00} + c'_{10} z_1 + \dots + c'_{r0} z_r + \dots + c'_{n-m,0} z_{n-m}) + \\
 & (c'_{01} + c'_{11} z_1 + \dots + c'_{r1} z_r + \dots + c'_{n-m,1} z_{n-m}) z_1 + \\
 & (c'_{0,n-m} + c'_{1,n-m} z_1 + \dots + c'_{r,n-m} z_r + \dots + c'_{n-m,n-m} z_{n-m}) z_{n-m}
 \end{aligned} \quad (4.42)$$

De c'_{ij} kolommen van (4.41) zijn dus de rijen van (4.42) en omgekeerd. Uit (4.39) en (4.40) blijkt dat voor de c'_{ij} in het algemeen niet geldt: $c'_{ij} = c'_{ji}$.

Door substitutie van (4.31) in (4.42) wordt z_r weer vervangen door x_p . (4.42) wordt dan

$$\begin{aligned}
 f = & \left\{ (c'_{00} - \frac{a_{p0}}{a_{pr}} c'_{r0}) + \dots + \frac{c'_{r0}}{a_{pr}} x_p + \dots + (c'_{n-m,0} - \frac{a_{p,n-m}}{a_{pr}} c'_{r0}) z_{n-m} \right\} + \\
 & \left\{ (c'_{01} - \frac{a_{p0}}{a_{pr}} c'_{r1}) + \dots + \frac{c'_{r1}}{a_{pr}} x_p + \dots + (c'_{n-m,1} - \frac{a_{p,n-m}}{a_{pr}} c'_{r1}) z_{n-m} \right\} z_1 + \\
 & \left\{ (c'_{0,n-m} - \frac{a_{p0}}{a_{pr}} c'_{r,n-m}) + \dots + \frac{c'_{r,n-m}}{a_{pr}} x_p + \dots + (c'_{n-m,n-m} - \frac{a_{p,n-m}}{a_{pr}} c'_{r,n-m}) z_{n-m} \right\} z_{n-m}
 \end{aligned} \quad (4.43)$$

Hieruit volgen de formules:

$$c'_{ji} = c'_{ji} - \frac{a_{pj}}{a_{pr}} c'_{ri} \quad j \neq r \quad (4.44)$$

$$c'_{ri} = \frac{c'_{ri}}{a_{pr}} \quad (4.45)$$

Door vergelijking van (4.39) en (4.40) met (4.44) en (4.45) ziet men gemakkelijk in, dat men weer dezelfde transformatieformules (4.39) en (4.40) mag toepassen mits men de c_{ij} vervangt door de c'_{ji} ; de a_{ij} 's blijven dezelfde.

In een getallenvoorbeeld geeft de verwisseling van de indices i en j geen moeite, daar c'_{ji} dan automatisch als c'_{ij} gelezen wordt, mits men tenminste de indices op dezelfde manier bepaalt, als bij de nummering van de elementen van een matrix.

De coëfficiënten a_{ij} en c_{ij} kan men in tableaux opschrijven nl.:

A		
	a_{10}	$a_{1,n-m}$
	a_{m0}	$a_{m,n-m}$

C		
	c_{00}	$c_{0,n-m}$
	$c_{n-m,0}$	$c_{n-m,n-m}$

Het A-tableau wordt één keer getransformeerd met de formules (4.33) t/m (4.36).

Het C-tableau wordt getransformeerd met (4.39) en (4.40); daarna getransponeerd of gespiegeld ten opzichte van de hoofddiagonaal. De elementen worden hernummerd (c'_{ji} wordt gelezen als c'_{ij}) en daarna nog eens getransformeerd met (4.39) en (4.40). (dezelfde a_{ij} 's en de hernummerde c'_{ij} 's!) Door substitutie kan men bewijzen dat het twee keer getransformeerde C-tableau weer symmetrisch is.

C. De variabele u_r is nul voordat een der basisvariabelen nul wordt.

Uit formule (4.23) volgt, analoog aan het betoog van (4.30), dat in dit geval geldt:

$$\min_{a_{hr} < 0} \left(-\frac{a_{h0}}{a_{hr}}, -\frac{c_{r0}}{c_{rr}} \right) = -\frac{c_{r0}}{c_{rr}}.$$

Bij verdere toename van z_r zal f niet meer afnemen (een convexe functie kan geen buigpunten hebben; zie fig. 4.9).

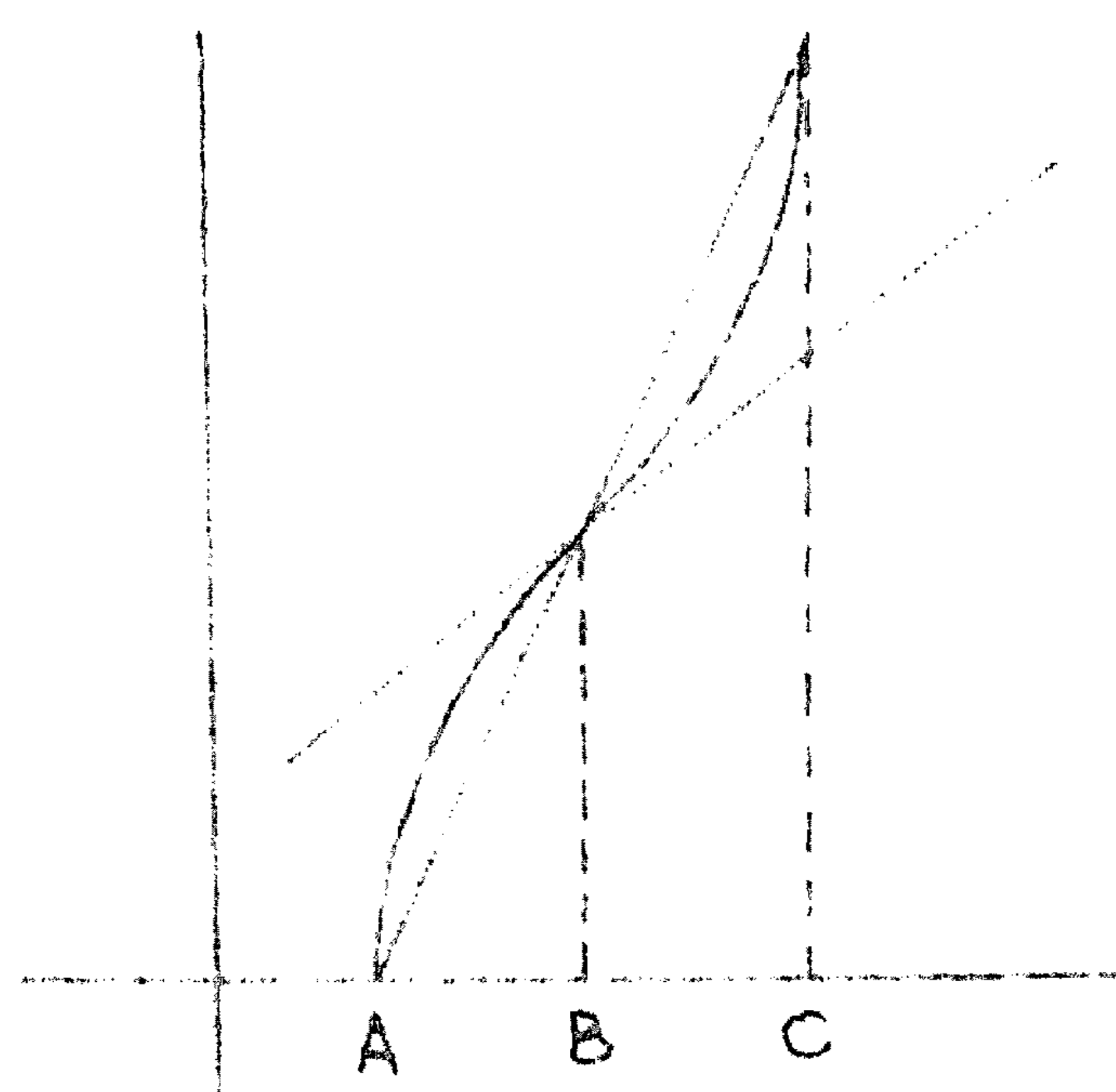


fig. 4.11

B is een buigpunt. Tussen A en B is f concaaf, tussen B en C is f convex.

In dit geval wordt de variabele u_r , zie (4.27), aan de niet-basisvariabelen en z_r aan de basisvariabelen toegevoegd.

Als men de tableaux gebruikt, dan worden de coëfficiënten $c_{r0}, \dots, c_{r,n-m}$ op de eerste regel van het A-tableau geplaatst en als $a_{10}, a_{11}, \dots, a_{1,n-m}$ gelezen. Van de oorspronkelijke a_{ij} 's wordt de index i met één verhoogd. Op grond van de argumentatie onder B blijkt, dat het uitgebreide A- en het C-tableau met behulp van (4.33) t/m (4.36), (4.39) en (4.40) getransformeerd worden.

Verder kan men bewijzen dat na toevoeging van de vrije variabele u_r aan de niet-basisvariabelen (z_r wordt dan basisvariabele) voor de twee keer getransformeerde c_{ij} geldt: $c_{ir}=0$ voor $i \neq r$ (zie fig. 4.12). Wordt bij de direct daarop volgende iteratie weer een vrije variabele aan de niet-basisvariabelen toegevoegd, dan geldt hetzelfde voor de zojuist ingevoerde vrije variabele, maar ook voor de bij de vorige stap ingevoerde vrije variabele. Dit blijft gelden, totdat er een niet-negatieve variabele aan de z -variabelen wordt toegevoegd (zie fig. 4.13). Alle nullen van het C-tableau kunnen nu veranderen in een reëel getal, dat niet gelijk aan nul is.

C		z_1	z_2	z_3
	150	6	0	7
z_1	7	8	0	-7
z_2	0	0	0	0
z_3	5	-11	0	7

fig. 4.12

z_2 is een zojuist ingevoerde vrije variabele
($n-m=3$)

C		z_1	z_2	z_3	z_4	z_5
	2	0	7	0	8	9
z_1	0	0	0	0	0	0
z_2	-7	0	9	0	12	-6
z_3	0	0	0	0	0	0
z_4	9	0	4	0	8	3
z_5	12	0	2	0	-4	6

fig. 4.13

z_1 en z_3 zijn twee achtereenvolgend ingevoerde vrije variabelen ($n-m=5$)

Aan het eind van de eerste iteratie kan de verzameling der z -variabelen een vrije variabele bevatten. In dit geval is de verzameling der basisvariabelen met één toegenomen. Hieruit blijkt dat bij voortgezette iteratie het aantal basisvariabelen variabel, en het aantal niet-basisvariabelen constant, nl. $n-m$, is.

Uit de onder B en C genoemde transformaties volgt verder nog:

1. Als $a_{h0} \geq 0$, dan geldt ook $a'_{h0} \geq 0$. Stel nl. dat $x_p = a_{p0} + a_{pr} z_r$ de eerste basisvariabele is, die nul wordt bij toenemende z_r . Dus voor $z_r = -\frac{a_{p0}}{a_{pr}}$ is $x_p = 0$. Substitueert men deze waarde in $x_h = a_{h0} + a_{hr} z_r$ ($h \neq p$), dan volgt hieruit $a_{h0} - \frac{a_{p0}}{a_{pr}} a_{hr} \geq 0$ of $a'_{h0} \geq 0$ (vgl. (4.33)).
2. Door toepassing van de transformatieformules (4.39) en (4.40) op c_{00} (twee keer!), kan men bewijzen dat in geval B voor $a_{p0} > 0$, geldt $\bar{c}_{00} < c_{00}$. ($\bar{c}_{00}$ is de twee keer getransformeerde c_{00}). Evenzo kan men hetzelfde voor geval C bewijzen.
3. Bovenstaande iteratie kan ook toegepast worden als f lineair is. Men kan dan bewijzen, als men de $c_{10}, c_{20}, \dots, c_{n-m,0}$ op de onderste regel van het A-tableau plaatst en c_{i0} leest als $a_{n-m+1,i}$, dat het zo aangevulde tableau met (4.33) t/m (4.36) getransformeerd kan worden. De $a_{n-m+1,i}$ komen dan overeen met de $c_i - z_i$ van het simplextableau.

Iteratie N

Als basisvergelijkingen dienen de vergelijkingen (4.21) en (4.22), met dien verstande, dat het aantal basisvariabelen nu groter dan m kan zijn en er één of meerdere vrije z -variabelen kunnen voorkomen. Deze worden door de onder C genoemde transformaties ingevoerd. In het geval de onder iteratie 1 genoemde z_r een vrije variabele is, kan f altijd, mits $c_{r0} \neq 0$, onafhankelijk van het teken van de afgeleide naar z_r voor $z_1=z_2=\dots=z_{n-m}=0$ vermindert worden door variatie van z_r . Immers, als $c_{r0} < 0$, dan volgt dit uit de bespreking bij fig. 4.8 en als $c_{r0} > 0$, dan kan men f verminderen door variatie van z_r vanuit 0 naar $-\infty$, hierbij weer $z_i=0$ latend voor $i \neq r$. Zie hiervoor fig. 4.14.

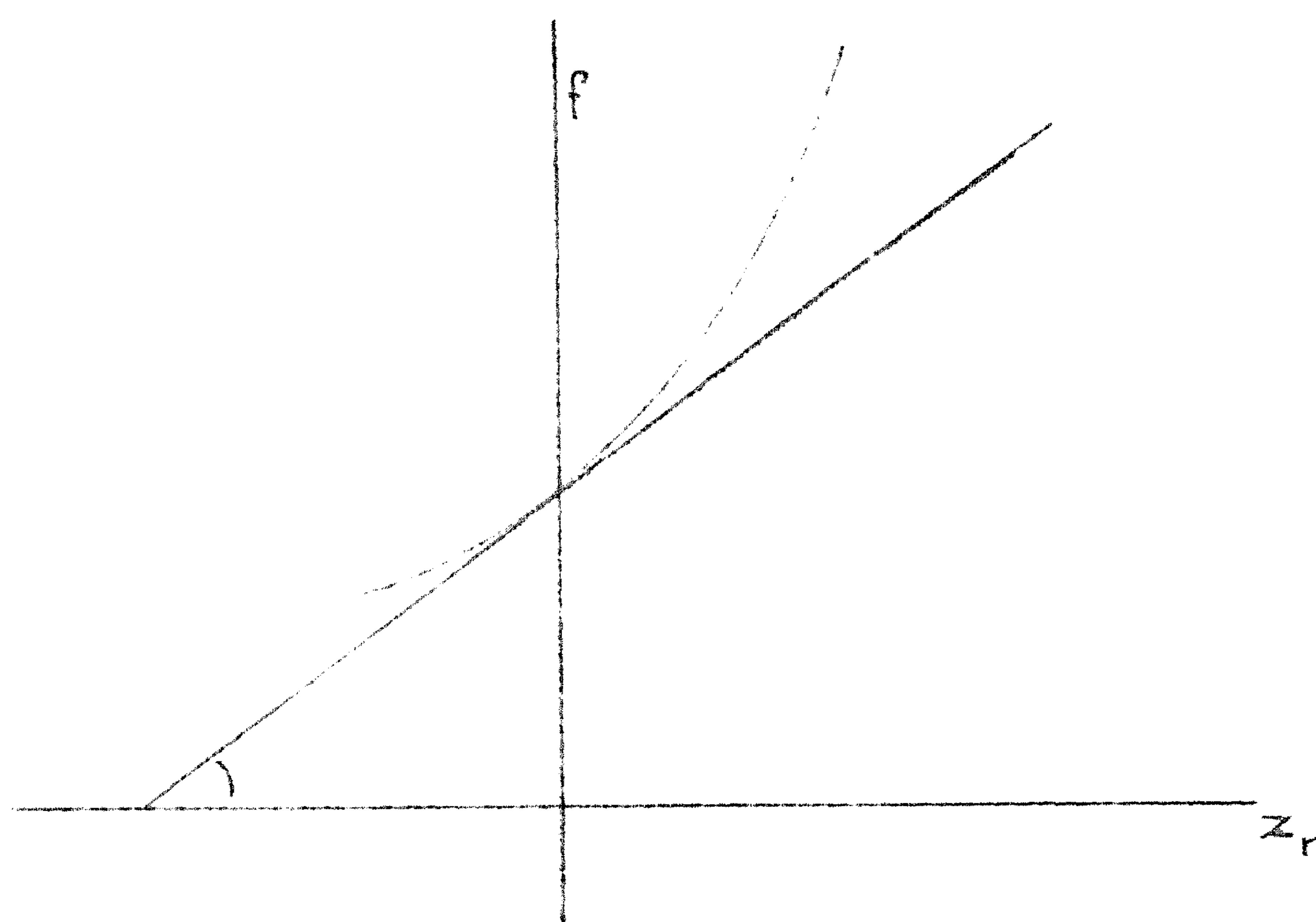


fig. 4.14

$$\frac{\partial f}{\partial z_r} = c_{r0} > 0 \text{ voor } z_i=0; f \text{ neemt af}$$

Hieruit blijkt, dat zolang de verzameling der niet-basisvariabelen een vrije variabele z_i met $c_{i0} \neq 0$ bevat, het de voorkeur verdient deze uit deze variabelen te verwijderen. Immers de functie kan in dit geval door variatie van de vrije variabele z_r , $z_i=0$ voor $i \neq r$, altijd vermindert worden. Dit verwijderen van de vrije variabele z_r uit de niet-basisvariabelen kan niet direct, nadat deze aan de z -variabelen is toegevoegd, plaatsvinden want na de toevoeging geldt: $c_{ir}=0$ $i \neq r$ (zie iteratie 1 onder C).

Evenals bij de eerste iteratie, kunnen we bij de variatie van z_r weer dezelfde gevallen A, B en C onderscheiden, die, onge-

acht de aard van de variabele z_r , weer dezelfde transformatieformules opleveren voor de a_{ij} en de c_{ij} . Voor het geval dat z_r een niet negatieve of een vrije variabele met $c_{r0} < 0$ is, blijft de onder iteratie 1 gehouden redenering gelijk.

Voor z_r , een vrije variabele met $c_{r0} > 0$, dienen de volgende kleine toevoegingen bij het onder iteratie 1 onder A resp. B en C vermelde, aangebracht te worden.

Stel dus $c_{r0} > 0$.

A. $c_{rr} = 0$ en alle $a_{hr} \leq 0$.

Uit de formules (4.28) en (4.29) volgt dat z_r ($z_i = 0$ voor $i \neq r$) willekeurig kan afnemen vanuit nul, zonder dat u_r of één der x_h van teken verandert. Het minimum van $f = c_{00} + 2c_{r0}z_r$ ($c_{r0} > 0$ en $z_r \leq 0$) is dus $-\infty$.

B. Eén der basisvariabelen wordt nul voordat de overige x_h en u_r van teken veranderen.

Stel dat de basisvariabele x_p nul wordt, voordat de overige basisvariabelen en u_r van teken veranderen, bij een negatieve variatie van z_r vanuit nul met $z_i = 0$, $i \neq r$. Uit vergelijking (4.29) blijkt, dat x_p alleen maar nul kan worden voor een niet positieve waarde van z_r als $a_{pr} > 0$. De index p van x_p kan dus gevonden worden met behulp van (zie beredenering van (4.30)):

$$\min_{a_{hr} > 0} \left(-\frac{a_{hr}}{a_{hr}}, -\frac{c_{r0}}{c_{rr}} \right) = -\frac{a_{pr}}{a_{pr}}.$$

De transformatieformules etc. zijn weer volkomen gelijk als bij iteratie 1 punt B.

Als er een vrije variabele uit de niet-basisvariabelen verwijderd wordt, dan kan deze zonder meer geschrapt worden uit de basisvariabelen. Het heeft nl. geen zin in dit geval verder op het teken te letten, aangezien dit bij variatie van een niet-basisvariabele geen belemmering vormt; dit dus in tegenstelling tot het teken van de variabele $x_h = a_{h0} + \sum a_{hj}z_j$ ($x_h \geq 0$!).

C. De variabele u_r wordt nul voordat de basisvariabelen x_h van teken veranderen.

In dit geval is (zie beredenering (4.30)):

$$\min_{a_{hr} > 0} \left(-\frac{a_{h0}}{a_{hr}}, -\frac{c_{r0}}{c_{rr}} \right) = -\frac{c_{r0}}{c_{rr}}. \quad (4.46)$$

Verder gelden dezelfde opmerkingen als onder B.

Uit het bovenstaande blijkt, dat het aantal basisvariabelen kan variëren tussen m en n , en dat het aantal niet-basisvariabelen constant blijft (nl. $n-m$).

Het is na het besprokene onder de iteraties 1 en N gemakkelijk in te zien, dat f niet meer verminderd kan worden als de c_{i0} , die met een niet negatieve variabele z_i corresponderen, niet negatief, en de c_{i0} die met een vrije variabele z_i corresponderen, nul zijn.

Men zegt dat de functie in de standaardvorm (s.v) is, als $c_{i0}=0$ voor z_i een vrije variabele. De functie f kan dus alleen zijn minimum aannemen in een s.v. Uit de beschrijving van het algoritme blijkt (zie iteratie N), dat we f zo spoedig mogelijk in een s.v. trachten te transformeren, d.w.z. we eerst de vrije variabelen met $c_{i0} \neq 0$ uit de niet-basisvariabelen verwijderen.

4.3. Convergentie

Uit de definitie van de standaardvorm volgt, dat het aantal s.v. (standaardvormen) eindig is. Een s.v. kan $n-m$, $n-m-1$, ..., 1, of 0 niet negatieve z -variabelen bevatten. Het aantal s.v. is dus hoogstens:

$$\binom{n}{n-m} + \binom{n}{n-m-1} + \dots + \binom{n}{1} + \binom{n}{0}.$$

Het algoritme zal nu in een eindig aantal stappen eindigen, als men kan bewijzen:

- dat men na k iteraties met een s.v. bereikt, die al eerder is aangenomen,
- dat men, met behulp van het algoritme, f in een eindig aantal stappen in een s.v. kan transformeren.

Voor het bewijs wordt verwezen naar de literatuur, bijv. BEALE E.M.L. - "On minimizing a convex function subject to linear inequalities" - Journal of the Royal Stat. Soc. (Series B) 17 (1955) pp. 173-184.

4.4. Schematisch overzicht van het algorithm

$$\begin{array}{c|c|c}
 A & & r \\
 \hline
 & a_{ij} - \frac{a_{pi} a_{rj}}{a_{pr}} & \frac{a_{ir}}{a_{pr}} \\
 \hline
 p & -\frac{a_{pj}}{a_{pr}} & \frac{1}{a_{pr}} \\
 \hline
 & & \frac{a_{ir}}{a_{pr}}
 \end{array} \quad (4.47)$$

$$\begin{array}{c|c|c}
 c & & r \\
 \hline
 & c_{ij} - \frac{a_{pi} a_{rj}}{a_{pr}} c_{ir} & \frac{c_{ir}}{a_{pr}} \\
 \hline
 & & c_{ij} - \frac{a_{pi} a_{rj}}{a_{pr}} c_{ir}
 \end{array} \quad (4.48)$$

1. Kies $\max_{c_{i0} \neq 0} |c_{i0}| = |c_{r0}|$ ¹⁾ (i varieert over de indices van de

vrije z-variabelen). Schrijf de coëfficiënten $c_{r0}, c_{r1}, \dots, c_{r,n-m}$ op de eerste regel van het A-tableau en lees de coëfficiënten als $a_{10}, a_{12}, \dots, a_{1n-m}$.

De i-indices van de oorspronkelijke a_{ij} 's worden met één vermeerderd.

Bepaal nu, als:

a) $c_{r0} > 0$ (eigenlijk $a_{10} > 0$) $\min_{a_{hr} > 0} \left(-\frac{a_{h0}}{a_{hr}} \right) = -\frac{a_{p0}}{a_{pr}}$ ¹⁾ (4.49)

b) $c_{r0} < 0$ (eigenlijk $a_{10} < 0$) $\min_{a_{hr} < 0} \left(-\frac{a_{h0}}{a_{hr}}, -\frac{a_{10}}{a_{1r}} \right) = -\frac{a_{p0}}{a_{pr}}$ ¹⁾ (4.50)

Als $p \neq 1$, dan wordt de eerste regel van het A-tableau geschrapt vóór de transformatie. (Is $p=1$ dan wordt de eerste regel niet geschrapt.) Hierna worden de beide tableaux met (4.47) en (4.48) getransformeerd.

Deze bewerkingen worden herhaald, totdat alle vrije variabelen uit de niet-basis variabelen verdwenen zijn, of de c_{i0} , die met een vrije variabele corresponderen nul zijn (m.a.w. f is in de s.v.).

2. Kies $\max_{c_{i0} < 0} |c_{i0}| = |c_{r0}|$ ¹⁾ (i varieert over de indices van de niet-negatieve z-variabelen.)

 1) Mochten er meerdere r- of p-indices zijn die hieraan voldoen, dan kan men bijv. de kleinste r- resp. p-index nemen die voldoet. De index p kan men, als $p=1$ ook toegelaten is, beter groter dan 1 nemen, opdat het A-tableau nietodeloos vergroot wordt. (Zie voorbeeld paragraaf 4.6.)

Plaats de coëfficiënten $c_{r0}, c_{r1}, \dots, c_{r, n-m}$ op de eerste regel van het A-tableau en herdoop de c_{ij} 's en a_{ij} 's (zie 1).

Bepaal nu:

$$\min_{a_{hr} < 0} \left(-\frac{a_{h0}}{a_{hr}}, -\frac{a_{10}}{a_{1r}} \right) = -\frac{a_{p0}}{a_{pr}} \quad 1) \quad (4.51)$$

In het geval $p \neq 1$ wordt de eerste regel van het A-tableau vóór de transformatie geschrapt (geldt $p=1$ dan wordt de eerste regel niet geschrapt). Hierna worden beide tableaux met (4.47) en (4.48) getransformeerd. De bewerkingen onder 2 worden herhaald, totdat er een c_{i0} , die met een vrije variabele correspondeert, ongelijk nul wordt. Vindt dit plaats dan wordt de regel onder 1 weer toegepast etc.

3. Is geen der onder 1 of 2 genoemde transformaties meer mogelijk, dan heeft f zijn minimum ($=c_{00}$) bereikt voor $x_h = a_{h0}$ en $z_j = 0$.

4.5. Uitbreiding

Het volgens het algoritme berekende locale minimum is, aangezien f convex is, gelijk aan het absolute minimum volgens stelling 2.2.

Als f echter een willekeurige kwadratische functie is, dan is het zelfs niet zeker dat de zo berekende waarde een lokaal minimum voorstelt. Geldt bijv. $c_{ij} = 0$ voor $i \neq j$, $c_{ii} < 0$ (dit kan daar f niet convex is), $\frac{\partial f}{\partial z_i} > 0$ voor z_i een niet negatieve variabele en $\frac{\partial f}{\partial z_i} = 0$ voor z_i een vrije variabele, dan zou volgens regel 3 van 4.4. ¹ f zijn minimum aannemen van $z_i = 0$ ($i = 1, 2, \dots, n-m$). Uit $f = c_{00} + c_{ii} z_i^2$ blijkt echter, dat f verminderd kan worden door variatie van z_i , en $z_i = 0$ voor $i \neq i$.

Door enkele aanvullingen op de regels van 4.4. kan het algoritme echter ook gebruikt worden voor de berekening van een lokaal minimum of een stationnair punt (een stationnair punt is een punt, dat geen (locaal) minimum is en waarin $\frac{\partial f}{\partial z_i} \geq 0$, als z_i een niet-negatieve variabele, en $\frac{\partial f}{\partial z_i} = 0$ als z_i een vrije variabele is).

- 1) Mochten er meerdere r - of p -indices zijn die hieraan voldoen, dan kan men bijv. de kleinste r - resp. p -index nemen die voldoet. De index p kan men, als $p=1$ ook toegelaten is, beter groter dan 1 nemen, opdat het A-tableau nietodeloos vergroot wordt. (Zie voorbeeld 3.6.)

4.6. Voorbeeld

Een auto-assemblagebedrijf produceert $10x_1$ autobussen en $10x_2$ trucks per maand. De aanvoer van grondstoffen en de capaciteit van het bedrijf zijn zodanig, dat er niet meer dan $\lambda > 0$ autobussen en trucks per maand afgeleverd kunnen worden. De winst per autobus (truck) is lineair afhankelijk van x_1 en wordt gegeven door $v_i = a_i - 10b_i x_i$ ($i = 1, 2, \dots$, v_i , a_i en b_i worden bijv. uitgedrukt in geldeenheden van 10.000 gld.).

Gevraagd wordt hoeveel autobussen en trucks er per maand afgeleverd moeten worden om een maximale winst te verkrijgen.

$$a_1=40, \quad a_2=20, \quad b_1=b_2=1, \quad \lambda=20.$$

Oplossing

In formule luidt dit probleem:

$$\text{Maximaliseer: } f = 10x_1 v_1 + 10x_2 v_2 = 10x_1(40 - 10x_1) + 10(20 - 10x_2)x_2 \quad (4.52)$$

$$\text{onder de nevenvoorwaarden: } 10x_1 + 10x_2 \leq 20.$$

$$x_1, x_2 \geq 0$$

of

$$\text{Minimaliseer: } f = -4x_1 - 2x_2 + x_1^2 + x_2^2 \quad (4.53)$$

$$\text{onder de nevenvoorwaarden: } x_1 + x_2 + x_3 = 2$$

$$x_1, x_2, x_3 \geq 0.$$

Opmerking

x_3 is een verschilvariabele om van $x_1 + x_2 \leq 2$ een gelijkheid te maken;

$$f = -\frac{1}{100} f'.$$

De A- en C-tableaux luiden respectievelijk:

A		x_1	x_2
x_3	2	-1	-1

C		x_1	x_2
	0	-2	-1
x_1	-2	1	0
x_2	-1	0	1

iteratie I

A_1		x_1	x_2	C_1		x_1	x_2
x_3	2	-1	-1		0	-2	-1
				x_1	-2	1	0
				x_2	-1	0	1

Volgens regel 2: $\max_{c_{10} < 0} |c_{10}| = |c_{01}| = 2$

$$\min_{a_{hr} < 0} \left(-\frac{a_{h0}}{a_{hr}}, -\frac{a_{10}}{a_{1r}} \right) = -\frac{a_{10}}{a_{11}} = -\frac{a_{20}}{a_{21}} = 2.$$

Om geen vrije variabelen aan de z-variabele toe te voegen kiezen we hier $p=2$. Vóór de transformatie wordt de eerste regel van het A-tableau geschrapt.

C_1		x_3	x_2
	-4	2	1
x_1	0	-1	-1
x_2	-1	0	1

C_1		x_1	x_2
	-4	0	-1
x_3	2	-1	0
x_2	1	-1	1

iteratie II

A_2		x_3	x_2	C_2		x_3	x_2
U_2	-1	1	2		-4	0	-1
x_1	2	-1	-1	x_3	0	1	1
				x_2	-1	1	2

Volgens regel 2: $\max_{c_{10} < 0} |c_{10}| = |c_{02}| = 1.$

$$\min_{a_{hr}} \left(-\frac{a_{h0}}{a_{hr}}, -\frac{a_{10}}{a_{1r}} \right) = -\frac{a_{10}}{a_{12}} = \frac{1}{2}.$$

De eerste regel van het A-tableau wordt dus meegetransformeerd.

C_2'	x_3	u_2
	$-4\frac{1}{2}$	$\frac{1}{2}$
x_3	$\frac{1}{2}$	$\frac{1}{2}$
x_2	0	1

C_2''	x_3	x_2
	$-4\frac{1}{2}$	0
x_3	$\frac{1}{2}$	0
u_2	$-\frac{1}{2}$	$\frac{1}{2}$

iteratie III

A_3	x_3	u_2	C_3	x_3	u_2
x_2	$\frac{1}{2}$	$-\frac{1}{2}$		$-4\frac{1}{2}$	0
x_1	$\frac{1}{2}$	$-\frac{1}{2}$	x_3	$\frac{1}{2}$	0
			u_2	0	$\frac{1}{2}$

Aangezien aan de voorwaarden van regel 3 voldaan is, wordt het minimum bereikt voor $x_1=1\frac{1}{2}$ en $x_2=\frac{1}{2}$. De winst is dus maximaal bij een produktie van 15 autobussen en 5 trucks per maand.

Opmerkingen

- 1) De vrije variabelen worden in de praktijk van de berekening genummerd naar de volgorde waarin ze in het probleem verschijnen.
- 2) De berekening van het voorbeeld kan ook aanschouwelijk voorgesteld worden door op te merken, dat de krommen $-4x_1 - 2x_2 + x_1^2 + x_2^2 = \text{const.}$ ($c > -5$) cirkels voorstellen met straal $\sqrt{c+5}$ - $(x_1-2)^2 + (x_2-1)^2 = c+5$ - en middelpunt (2,1).

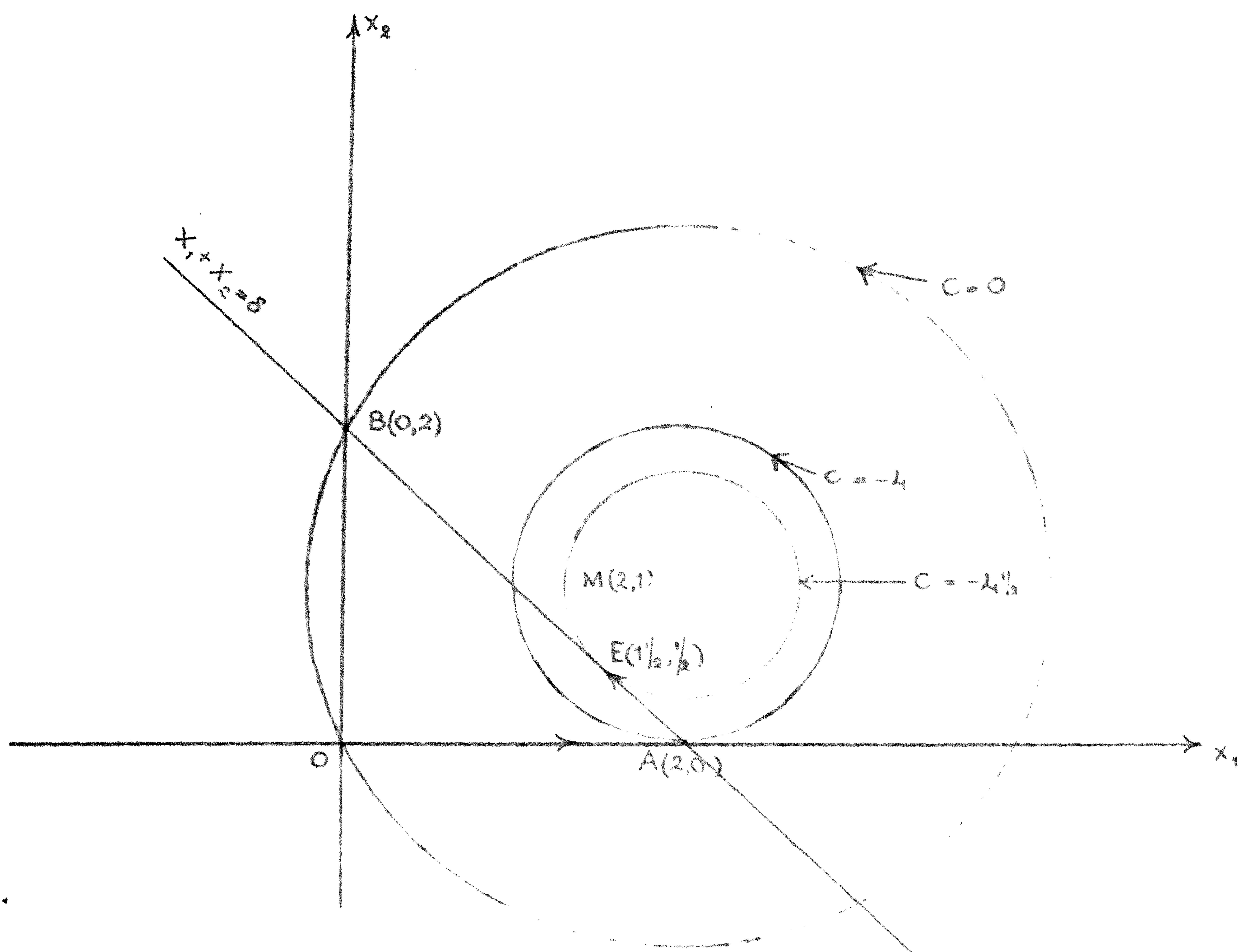


fig. 4.13

De iso-winstkrommen zijn cirkels. Het optimum van het probleem wordt bereikt via OAE. Het toegelaten gebied wordt begrensd door $\triangle OAB$.

Bij de eerste iteratie wordt x_1 ($x_2=0$!) gevarieerd van 0 tot 2. f neemt dan af van 0 tot -4. Meetkundig betekent dit, dat we - langs de x_1 -as - bewegen van 0 naar A.

Gedurende de tweede iteratie wordt $x_3=0$ gehouden m.a.w. x_1 en x_2 moeten voldoen aan $x_1+x_2=8$. Uit de berekening van iteratie II blijkt dat x_2 gevarieerd kan worden tot $\frac{1}{2}$. Dus x_1 is - de beweging vindt plaats langs de lijn AE - $1\frac{1}{2}$.

5. Algorithme van WOLFE

Het algorithme van WOLFE levert door middel van de simplex-methode een oplossing van probleem (1.4), (1.5) en (1.6). Het

algorithme kent twee vormen nl. de korte en de lange vorm. De korte vorm levert zonder meer een oplossing van (1.4), (1.5) en (1.6) op; de lange vorm - onder de restricties (1.5) en (1.6) - levert voor $f(X) = \lambda \sum_{j=1}^n p_j x_j + \frac{1}{2} \sum_{i,j=1}^n c_{ij} x_i x_j$ een oplossing, die een functie is van λ . Het nadeel van de korte vorm is, dat geëist wordt dat $\bar{C}$ positief definitief is, $\bar{C}$ of alle lineaire termen ontbreken (P is dan een nulvector) en C semi-positief definitief is. Bij de lange vorm is het voldoende als C semi-positief definitief is.

In dit hoofdstuk zal alleen een versie van de verkorte vorm besproken worden. Voor de berekening wordt gebruik gemaakt van de volgende stelling:

Stelling 5.1. Een vector $\bar{X} = (\bar{x}_1, \dots, \bar{x}_n) \geq 0$, die aan $\sum_{j=1}^n a_{ij} \bar{x}_j = a_{i0}$ ($i=1, 2, \dots, m$) voldoet, is een optimale oplossing van (1.4) t/m (1.6). Als er een n -vector $V = (v_1, \dots, v_n) \geq 0$ en een m -vector $U = (\lambda_1, \dots, \lambda_m)$ bestaan, zodat:

$$\begin{array}{ccccccc} v_1 \bar{x}_1 & + & \dots & + & v_n \bar{x}_n & & (5.1) = 0 \\ c_{11} \bar{x}_1 & + & \dots & + & c_{1n} \bar{x}_n & - v_1 & a_{11} \lambda_1 + \dots + a_{m1} \lambda_m + p_1 = 0 \\ \vdots & & & & \vdots & \ddots & \vdots \\ c_{1n} \bar{x}_1 & + & \dots & + & c_{nn} \bar{x}_n & - v_n & a_{1n} \lambda_1 + \dots + a_{mn} \lambda_m + p_n = 0 \end{array}$$

Toelichting voor $m=2$ en $n=3$.

In dit geval luiden f en de restricties (zie 1.7):

$$f = 0 + \frac{1}{2} p_1 x_1 + \frac{1}{2} p_2 x_2 + \frac{1}{2} p_3 x_3 \quad (5.2)$$

$$\begin{aligned} & \frac{1}{2} p_1 x_1 + \frac{1}{2} c_{11} x_1^2 + \frac{1}{2} c_{12} x_1 x_2 + \frac{1}{2} c_{13} x_1 x_3 \\ & p_2 x_2 + c_{21} x_1 x_2 + \frac{1}{2} c_{22} x_2^2 + \frac{1}{2} c_{23} x_2 x_3 \\ & p_3 x_3 + \frac{1}{2} c_{31} x_1 x_3 + \frac{1}{2} c_{32} x_2 x_3 + \frac{1}{2} c_{33} x_3^2 \end{aligned}$$

$$a_{11} x_1 + a_{12} x_2 + a_{13} x_3 - a_{10} = 0 \quad (5.3)$$

$$a_{21} x_1 + a_{22} x_2 + a_{23} x_3 - a_{20} = 0$$

$$x_1, x_2, x_3 \geq 0. \quad (5.4)$$

Als we de voorwaarde (5.4) buiten beschouwing laten, dan kan het optimum van f onder de restricties (5.3) met behulp van de multiplicatorenmethode van LAGRANGE berekend worden (zie voorb. 4.2 en hoofdstuk II blz. 39). De LAGRANGE-functie F luidt dan:

$$F = f + \lambda_1(a_{11}x_1 + \dots + a_{13}x_3 - a_{10}) + \lambda_2(a_{21}x_1 + \dots - a_{20}) \quad (5.5)$$

De optimale oplossing $\bar{X} = (\bar{x}_1, \bar{x}_2, \bar{x}_3)$ en de multiplicatoren λ_1 en λ_2 moeten dan voldoen aan:

$$\frac{\partial F}{\partial x_1} = c_{11}\bar{x}_1 + c_{12}\bar{x}_2 + c_{13}\bar{x}_3 + a_{11}\lambda_1 + a_{21}\lambda_2 + p_1 = 0 \quad (5.6)$$

$$\frac{\partial F}{\partial x_2} = c_{21}\bar{x}_1 + c_{22}\bar{x}_2 + c_{23}\bar{x}_3 + a_{12}\lambda_1 + a_{22}\lambda_2 + p_2 = 0$$

$$\frac{\partial F}{\partial x_3} = c_{31}\bar{x}_1 + c_{32}\bar{x}_2 + c_{33}\bar{x}_3 + a_{13}\lambda_1 + a_{23}\lambda_2 + p_3 = 0$$

$$\frac{\partial F}{\partial \lambda_1} = a_{11}\bar{x}_1 + a_{12}\bar{x}_2 + a_{13}\bar{x}_3 - a_{10} = 0 \quad (5.7)$$

$$\frac{\partial F}{\partial \lambda_2} = a_{21}\bar{x}_1 + a_{22}\bar{x}_2 + a_{23}\bar{x}_3 - a_{20} = 0$$

Als het uit (5.6) en (5.7) berekende minimum $\bar{X} = (\bar{x}_1, \bar{x}_2, \bar{x}_3)$ voldoet aan $X \geq 0$, dan is het probleem (5.2), (5.3), (5.4) opgelost. Dit zal i.h.a. echter niet het geval zijn, zoals blijkt uit de voorbeelden 4.1, 4.2 en 4.3. Uit de opmerkingen aan het einde van paragraaf 4.1 blijkt, dat de partiële afgeleiden - als $\bar{x}$ ook moet voldoen aan 4.4 - in $\bar{X}$ moeten voldoen aan:

$$\frac{\partial F}{\partial x_1} \geq 0, \quad \frac{\partial F}{\partial x_2} \geq 0, \quad \frac{\partial F}{\partial x_3} \geq 0 \quad (5.8)$$

of anders geschreven:

$$\frac{\partial F}{\partial x_1} - v_1 = 0, \quad \frac{\partial F}{\partial x_2} - v_2 = 0, \quad \frac{\partial F}{\partial x_3} - v_3 = 0 \quad (5.9)$$

$$V = (v_1, v_2, v_3) \geq 0.$$

Bovendien moet gelden $v_1\bar{x}_1 = v_2\bar{x}_2 = v_3\bar{x}_3 = 0$, want, als voor het minimum $\bar{X}$ geldt $x_1=0$, dan voldoet de partiële afgeleide $\frac{\partial F}{\partial x_1}$ in $\bar{X}$ aan $\frac{\partial F}{\partial x_1} \geq 0$. Als $\bar{x}_1 \neq 0$, dan geldt $\frac{\partial F}{\partial x_1} = 0$ (zie ook de opmerkingen aan het einde van paragraaf 3.1). Analoge redeneringen gelden voor

$v_2 \bar{x}_2 = 0$ en $v_3 \bar{x}_3 = 0$. Aangezien V en $\bar{X}$ niet negatief zijn, is $v_1 \bar{x}_1 = v_2 \bar{x}_2 = v_3 \bar{x}_3 = 0$ equivalent met:

$$v_1 \bar{x}_1 + v_2 \bar{x}_2 + v_3 \bar{x}_3 = 0 \quad \text{of} \quad V \cdot \bar{X} = 0. \quad (5.10)$$

Samenvattend zien we, dat een optimale oplossing $\bar{X} = (\bar{x}_1, \bar{x}_2, \bar{x}_3)$ van (5.2), (5.3) en (5.4) moet voldoen aan:

$$\begin{array}{rcl} a_{11}\bar{x}_1 + a_{12}\bar{x}_2 + a_{13}\bar{x}_3 & & -a_{10} = 0 \\ a_{21}\bar{x}_1 + a_{22}\bar{x}_2 + a_{23}\bar{x}_3 & & -a_{20} = 0 \\ c_{11}\bar{x}_1 + c_{12}\bar{x}_2 + c_{13}\bar{x}_3 - v_1 & + v_{11}\lambda_1 + a_{21}\lambda_2 + a_{31}\lambda_3 + p_1 & = 0 \\ c_{21}\bar{x}_1 + c_{22}\bar{x}_2 + c_{23}\bar{x}_3 - v_2 & + a_{21}\lambda_1 & + p_2 = 0 \\ c_{31}\bar{x}_1 + c_{32}\bar{x}_2 + c_{33}\bar{x}_3 - v_3 & & + p_3 = 0 \end{array} \quad (5.11)$$

$$\bar{X} = (\bar{x}_1, \bar{x}_2, \bar{x}_3) \geq 0, \quad (5.12)$$

$$v_1 \bar{x}_1 + v_2 \bar{x}_2 + v_3 \bar{x}_3 = 0. \quad (5.13)$$

Het bewijs van stelling (5.1) zal hier niet gegeven worden.

De voorwaarden (5.11), (5.12) en (5.13) maken deel uit van een stelsel, dat men de voorwaarden van KUHN en TUCKER noemt. KUHN en TUCKER bewijzen, dat de voorwaarden (5.11), (5.12) en (5.13) niet alleen voldoende, zoals uit bovenstaande "redenering" volgt, maar ook noodzakelijk zijn opdat $\bar{X}$ de kleinste waarde van f oplevert (mits f convex is).

Stelsel (5.11), (5.12) en (5.13) kan in de volgende verzwakte vorm geschreven worden:

$$\begin{array}{rcl} a_{11}x_1 + \dots + a_{1n}x_n & & +w_1 = a_{10} \\ \vdots & & \vdots \\ a_{m1}x_1 + \dots + a_{mn}x_n & & +w_m = a_{m0} \\ c_{11}x_1 + \dots + c_{1n}x_n - v_1 & + a_{11}\lambda_1 + \dots + a_{m1}\lambda_m + z_{11} - z_{21} & = -p_1 \\ \vdots & & \vdots \\ c_{n1}x_1 + \dots + c_{nn}x_n & -v_n + a_{1n}\lambda_1 + \dots + a_{mn}\lambda_m + z_{1n} - z_{2n} & = -p_n \end{array} \quad (5.14)$$

$$\begin{array}{l} (x_1, \dots, x_n) \geq 0, \quad (v_1, \dots, v_n) \geq 0, \quad (z_{11}, \dots, z_{1n}) \geq 0, \quad (z_{21}, \dots, z_{2n}) \geq 0 \\ v_1 x_1 + \dots + v_n x_n = 0. \end{array}$$

Het oplossen van (5.14) gaat in twee etappes. In (5.14) worden $U = (\lambda_1, \dots, \lambda_m)$ en $V = (v_1, \dots, v_n)$ nul gesteld, en daarna lossen we het lineaire programmeringsprobleem (5.15), (5.16) op.

$$\text{Minimaliseer: } f_1 = \sum_{j=1}^m w_j \quad (5.15)$$

onder de nevenvoorwaarden:

$$\begin{array}{rcl} a_{11}x_1 + \dots + a_{1n}x_n & + w_1 & = a_{10} \\ \vdots & \vdots & \\ a_{m1}x_1 + \dots + a_{mn}x_n & + w_m & = a_{m0} \\ c_{11}x_1 + \dots + c_{1n}x_n + z_{11} & - z_{21} & = -p_1 \\ \vdots & \vdots & \\ c_{n1}x_1 + \dots + c_{nn}x_n + z_{1n} & - z_{2n} & = -p_n \end{array} \quad (5.16)$$

$$X = (x_1, \dots, x_n) \geq 0, \quad Z_1 = (z_{11}, \dots, z_{1n}) \geq 0, \quad Z_2 = (z_{21}, \dots, z_{2n}) \geq 0, \\ W = (w_1, \dots, w_m) \geq 0.$$

Voordat dit probleem opgelost wordt, vermenigvuldigen we de vergelijkingen die een negatieve $-p_1$ hebben met -1 . Door de keuze van de Z_1 en Z_2 zijn we altijd in staat uit de coëfficiënten van de z_{ij} en w_j een basis, die uit eenheidsvectoren bestaat, te kiezen.

Voor het voorbeeld in par. 4.5 luiden deze vergelijkingen:

$$\begin{array}{rcl} \text{Minimaliseer } f_1 = w & \text{onder de} & (5.17) \\ \text{restricties: } x_1 + x_2 + x_3 & + w & = 2 \\ 2x_1 & + z_{11} & - z_{21} = 4 \\ 2x_2 & + z_{12} & - z_{22} = 2 \\ & + z_{13} & - z_{23} = 0 \\ x_1, x_2, x_3, z_{11}, z_{12}, z_{13}, z_{21}, z_{22}, z_{23}, w & \geq 0. \end{array}$$

Het simplextableau van (5.17) luidt:

Tabel 5.I
Aanvangstableau van 5.17 ¹⁾

C	B	X ₀	X ₁	X ₂	X ₃	X ₈	X ₉	X ₁₀	X ₁₁	X ₁₂	X ₁₃	X ₁₄
1	x ₁₄	2	1	1	1	0	0	0	0	0	0	1
0	x ₈	4	2	0	0	1	0	0	-1	0	0	0
0	x ₉	2	0	2	0	0	1	0	0	-1	0	0
0	x ₁₀	0	0	0	0	0	0	1	0	0	-1	0
$z_j - c_j$		1	1	1	1	0	0	0	0	0	0	0

Door x₁₄ uit de basis te verwijderen en x₁ erin op te nemen, blijkt probleem (5.17) opgelost te zijn. Zie voor transformatieformules e.d. hoofdstuk XII.

Tabel 5.II
Eindtableau van (5.17)

C	B	X ₀	X ₁	X ₂	X ₃	X ₈	X ₉	X ₁₀	X ₁₁	X ₁₂	X ₁₃	X ₁₄
0	x ₁	2	1	1	1	0	0	0	0	0	0	1
0	x ₈	0	0	-2	-2	1	0	0	-1	0	0	-2
0	x ₉	2	0	2	0	0	1	0	0	-1	0	0
0	x ₁₀	0	0	0	0	0	0	1	0	0	-1	0
$z_j - c_j$		0	0	0	0	0	0	0	0	0	0	-1

De oplossing van (5.17) luidt dus:

$$x_1=2, x_2=x_3=0, z_{11}=z_{12}=z_{13}=1, z_{21}=z_{22}=z_{23}=0.$$

De oplossing van (5.16) levert dus altijd een toegelaten basisoplossing²⁾ met $U=V=0$ en $V.X=0$ ($v_1x_1+\dots+v_nx_n=0$) op van:

-
- 1) De nummering is niet doorlopend gekozen, omdat dit tableau deel uitmaakt van een groter.
 - 2) Het is altijd mogelijk een oplossing van (5.16) te vinden die niet meer dan n van de z_{ij} variabelen bevat. (Pas, nadat er een oplossing met $k > n$ z_{ij} variabelen bepaald is, nog k-n keer de simplexmethode toe.)¹⁾ De zo gevonden oplossing van (5.16) levert een toegelaten basisoplossing met $U=V=0$ en $V.X=0$ op van (5.18) en (5.19).

$$\begin{aligned}
 &\text{Minimaliseer } f_2 = z_{11} + \dots + z_{1n} + z_{21} + \dots + z_{2n} \quad \text{onder de} \quad (5.18) \\
 &\text{restricties: } \begin{aligned} &a_{11}x_1 + \dots + a_{1n}x_n &&= a_{10} \\ &\vdots &&\vdots \\ &a_{n1}x_1 + \dots + a_{nn}x_n &&= a_{n0} \end{aligned} \quad (5.19) \\
 &\begin{aligned} &c_{11}x_1 + \dots + c_{1n}x_n + a_{11}\lambda_1 + \dots + a_{n1}\lambda_1 + z_{11} &&-z_{21} &&= -p_1 \\ &\vdots &&\vdots &&\vdots &&\vdots \\ &c_{n1}x_1 + \dots + c_{nn}x_n + a_{1n}\lambda_1 + \dots + a_{nn}\lambda_n &&+z_{1n} &&-z_{2n} &&= -p_n \end{aligned} \\
 &X, V, z_1, z_2 \geq 0 \quad V.X = 0
 \end{aligned}$$

met dien verstande, dat alle z_{ij} , die nul zijn (n stuks) in de oplossing van (5.16) geschrapt worden in (5.18) en (5.19) en bij de oplossing van (5.18) en (5.19) niet meer meegerekend worden.

Bovendien moet nu (in verband met de voorwaarde $V.X=0$) de volgende keuzeregel toegepast worden:

Keuzeregel: Als x_k in de basis is, dan wordt v_k niet toegelaten in de basis en omgekeerd ($V.X=0$ moet steeds gelden!).

Voor het voorbeeld uit par. 4.5 zien de vergelijkingen (5.18) en (5.19) er uit als:

$$\begin{aligned}
 &\text{Minimaliseer } f_2 = z_{11} + z_{12} + z_{13} \quad \text{onder de} \quad (5.20) \\
 &\text{restricties: } \begin{aligned} &x_1 + x_2 + x_3 &&= 2 \\ &2x_1 + \lambda_1 + z_{11} &&= 4 \\ &\quad + 2x_2 + \lambda_2 + z_{12} &&= 2 \\ &\quad \quad + \lambda_3 + z_{13} &&= 0 \end{aligned} \\
 &x_1, x_2, x_3, z_{11}, z_{12}, z_{13} \geq 0.
 \end{aligned}$$

Het simplextableau van (5.15) en (5.16) is een deelttableau van (5.20). (Aangenomen wordt dat alle $-p_i$ niet negatief zijn.)

f_1	0	---	0	0	---	0	0	---	0	0	---	---	---	0	1	---	1
f_2	0	---	0	0	---	0	0	---	0						0	---	0
	x_1	---	x_n	v_1	---	v_n	u_1	---	u_m	z_n	---	z_{n1}	z_{12}	z_{n2}	w_1	---	w_m
X_0	X_1		X_n	X_{n+1}		X_{2n}	X_{2n+1}		X_{3n+m}	X_{3n+m+1}		X_{4n+m}	X_{4n+m+1}		X_{4n+2m}		
a_{1n}	a_{11}		a_{1n}	0		0	0		0	0		0	0	0	1		0
a_{mn}	a_{m1}		a_{mn}	0		0	0		0	0		0	0	0		1	0
$-p_1$	c_1		c_n	-1		0	a_{11}		a_{m1}	1		-1	0	0		0	0
$-p_n$	c_n		c_{nn}	0		-1	a_{1n}		a_{mn}	0		1	0	-1	0		0

$\underbrace{\hspace{10em}}_n$
 $\underbrace{\hspace{10em}}_n$
 $\underbrace{\hspace{10em}}_m$
 $\underbrace{\hspace{10em}}_n$
 $\underbrace{\hspace{10em}}_n$
 $\underbrace{\hspace{10em}}_m$

Het deelttableau van (5.15) en (5.16) ontstaat door de vectoren X_{n+1} tot en met X_{2n+m} en de coëfficiënten van f_2 te schrappen. Het minimum van f_1 is nul en wordt zeker bereikt (zie hoofdstuk XII). Het tableau van (5.18) en (5.19) ontstaat (aangenomen wordt dat men de herziene simplexmethode - zie hoofdstuk XII - toepast, m.a.w. dat niet steeds het grondtableau meegetransformeerd wordt) door uit de kolommen X_{2n+m+1} t/m X_{4n+m} die kolommen te schrappen, waarvan de bijbehorende variabele z_{ij} nul is in de oplossing van (5.15) en (5.16) en door de kolommen X_{4n+m+1} t/m X_{4n+2m} te schrappen. Simplextableau (5.20) ziet er dus voor voorbeeld par. 4.5 uit als:

	x_1	x_2	x_3	v_1	v_2	v_3	u	z_{11}	z_{21}	z_{31}	z_{12}	z_{22}	z_{32}	w
f_1	0	0	0	0	0	0	0	0	0	0	0	0	0	1
f_2	0	0	0	0	0	0	0	1	1	1	0	0	0	0
X_0	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	X_{11}	X_{12}	X_{13}	X_{14}
2	1	1	1	0	0	0	0	0	0	0	0	0	0	1
4	2	0	0	-1	0	0	1	1	0	0	-1	0	0	0
2	0	2	0	0	-1	0	1	0	1	0	0	-1	0	0
0	0	0	0	0	0	-1	1	0	0	1	0	0	-1	0

Het tableau van (5.17) ontstaat door de kolommen X_4 t/m X_7 én de rij f_2 te schrappen uit (5.21); van (5.20) door de kolommen X_{11} t/m X_{14} en de rij f_1 te schrappen.

De vrije variabelen u_1 behoeven, nadat ze in de basis zijn opgenomen, niet meer uit de basis verwijderd te worden (zie hoofdstuk XII blz. 19 e.v., de keuze van θ).

Men kan nu bewijzen, dat het minimum nul van f_2 in een eindig aantal stappen bereikt wordt. Dit volgt niet zonder meer uit hoofdstuk XII, want er is aan de simplexmethode een nieuwe keuzeregel toegevoegd.

Uit het bovenstaande volgt, dat men na het vinden van de optimale oplossing van (5.15), (5.16) de simplexmethode met het deeltableau van (5.20) kan voortzetten als men:

- a) de vectoren $A_{n+1}, \dots, A_{2n+m}$ aan het deeltableau toevoegt;
- b) uit de vectoren $A_{2n+m+1}, \dots, A_{4n+m}$ die vectoren, waarvan de bijbehorende z_{ij} nul is in de optimale oplossing van (5.15), (5.16) schrapt uit het deeltableau;
- c) de vectoren $A_{4n+m+1}, \dots, A_{4n+2m}$ verwijdert uit het deeltableau;
- d) in plaats van de coëfficiënten c_1 van f_1 nu de coëfficiënten c_1 van f_2 gebruikt.

De uiteindelijke oplossing verkrijgt men door uit de optimale oplossing van (5.18), (5.19) alle niet x-componenten te schrappen.

Opmerking

- 1) Gezien het grote aantal nullen in het tableau (5.20) (zie (5.21)) verdient het aanbeveling de herziene simplexmethode te gebruiken (zie hoofdstuk XII, par. 11).
- 2) Bij de werkelijke afleiding van stelling 5.1 en het convergentiebewijs is essentieel gebruik gemaakt van het (semi-) positief definit zijn van de matrix C (ook buiten het door de nevenvoorwaarden $\sum_{j=1}^n a_{ij}x_j = a_{i0}$, $i=1,2,\dots,m$, $x \geq 0$ begrensde gebied moet $f(x)$ convex zijn voor het convergentiebewijs). Het algoritme levert dus bij een willekeurige kwadratische functie in het algemeen géén zinvolle oplossing (bijv. een lokaal minimum of een stationnair punt, zie voor def. par. 4.5).

Voorbeeld 5.1

Het voorbeeld in 4.5 wordt met behulp van het algoritme van Wolfe hieronder nog eens uitgerekend. Bij de "simplexberekeningen" wordt de herziene simplexmethode toegepast (zie hoofdstuk XII, par. 11). Tableau (5.20) ziet er na substitutie van de getalwaarden als volgt uit:

	x_1	x_2	x_3	v_1	v_2	v_3	u	z_{11}	z_{21}	z_{31}	z_{12}	z_{22}	z_{32}	w
f_1	0	0	0	0	0	0	0	0	0	0	0	0	0	1
f_2	0	0	0	0	0	0	0	1	1	1	0	0	0	0
x_0	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}	x_{12}	x_{13}	x_{14}
2	1	1	1	0	0	0	0	0	0	0	0	0	0	1
4	2	0	0	-1	0	0	1	1	0	0	-1	0	0	0
2	0	2	0	0	-1	0	1	0	1	0	0	-1	0	0
0	0	0	0	0	0	-1	1	0	0	1	0	0	-1	0

(5.21)

Phase 1. Oplossing van het probleem (5.15) en (5.16)

Bij de minimalisering van (5.15) worden de vectoren x_4 , x_5 , x_6 en x_7 niet meegerekend. Bij de berekening is de herziene simplexmethode (zie hoofdstuk XII) toegepast.

I	C_0	B	x_0	B^{-1}				x_1	
	1	x_{14}	2	1	0	0	0	1	←
	0	x_8	4	0	1	0	0	2	$C_0 B^{-1} = (1, 0, 0, 0)$
	0	x_9	2	0	0	1	0	0	$D = (1, 1, 1, 0, 0, 0, 0, 0, 0)$
	0	x_{10}	0	0	0	0	1	0	↑

Phase 2. Oplossing van probleem (5.18) en (5.19)

De vectoren x_{11} , x_{12} , x_{13} en x_{14} worden geschrapt, de vectoren x_4 , x_5 , x_6 en x_7 worden van nu af meegeteld, f_1 wordt vervangen door f_2 en de nieuwe keuzeregels wordt, zonodig, toegepast.

II

C_0	B	X_0	B^{-1}				X_7
0	X_1	2	1	0	0	0	0
1	X_8	0	-2	1	0	0	1
1	X_9	2	0	0	1	0	1
1	X_{10}	0	0	0	0	1	1

$\leftarrow C_0 B^{-1} = (-2, 1, 1, 1)$
 $D = (0, 0, -2, -1, -1, -1, 3, 0, 0, 0)$
 $\uparrow$

III

C_0	B	X_0	B^{-1}				X_2
0	X_1	2	1	0	0	0	1
0	X_7	0	-2	1	0	0	-2
1	X_9	2	2	-1	1	0	4
0	X_{10}	0	2	-1	0	1	2

$C_0 B^{-1} = (4, -2, 1, 1)$
 $D = (0, 6, 4, 2, -1, -1, 0, -3, 0, 0)$
 $\leftarrow \uparrow$

IV

C_0	B	X_0	B^{-1}				X_6
0	X_1	2	0	$\frac{1}{2}$	0	$-\frac{1}{2}$	$\frac{1}{2}$
0	X_7	0	0	0	0	1	-1
1	X_9	2	-2	1	1	-2	2
0	X_2	0	1	$-\frac{1}{2}$	0	$\frac{1}{2}$	$-\frac{1}{2}$

$C_0 B^{-1} = (2, 1, 1, -2)$
 $\leftarrow D = (0, 0, -2, -1, -1, 2, 0, 0, 0, -3)$
 $\uparrow$

V

C_0	B	X_0	B^{-1}				X
0	X_1	<u>$1\frac{1}{2}$</u>	$\frac{1}{2}$	0	$-\frac{1}{2}$	$\frac{1}{2}$	
0	X_7	1	1	1	1	-1	
0	X_6	1	-1	1	1	-2	
0	X_2	<u>$\frac{1}{2}$</u>	$\frac{1}{2}$	0	$\frac{1}{2}$	$-\frac{1}{2}$	

$C_0 B^{-1} = (0, 0, 0, 0)$
 $D = (0, 0, 0, 0, 0, 0, 0, 0, 0, 0)$

Aangezien alle $z_j - c_j = 0$ is het minimum bereikt voor $x_1 = 1\frac{1}{2}$ en $x_2 = \frac{1}{2}$.

6. Vergelijking van de algorithmen

Voor het geval dat de objectieve functie kwadratisch en convex is en de restricties lineair zijn, kan men voor de berekening een keuze doen uit verscheidene algorithmen (zie par. 6 t/m 8). De meeste algorithmen zijn eindig, d.w.z. dat het optimum in een eindig aantal iteraties bereikt wordt. Op zichzelf is dit niet zo belangrijk, mits er maar een convergentie-criterium bestaat, en de convergentie niet al te langzaam is.

Als we ons nu beperken tot het A.W.¹⁾ en het A.B., dan blijkt, dat een programma voor het lineaire programmeringsmodel heel gemakkelijk aangepast kan worden aan het A.W.. Het SHARE programma voor de I.B.M. 704 vereist ongeveer 10 veranderingen in de instructies. Het A.B. vereist daarentegen veel meer veranderingen.

Daartegenover staat, dat het simplextableau (zie 5.20) van het A.W. voor m restricties met n variabelen een $(m+n)$ bij $(m+3n)$ tableau omvat (de "deelttableaux" van (5.20) ontstaan door $n+m$ kolommen van tableau (5.20) te schrappen; zie (5.15), (5.16) en (5.13), (5.19)). Dit kan bij grote m en n in veel gevallen een te zware belasting zijn van de geheugencapaciteit van de computer; tevens zal het aantal afrondingsfouten sterk toenemen.

De tableaux van het A.B. zijn veel kleiner. Bovendien is het algoritme eenvoudiger en aanschouwelijker "georganiseerd". Voor het geval dat $f(X)$ niet convex (echter wel kwadratisch!) is, kan het algoritme van BEALE nog toegepast worden. Men vindt dan echter op zijn hoogst een lokaal optimum (dat toevallig het absolute optimum kan zijn!). De methode van WOLFE faalt in dit geval. Soms is het gewenst - als de lineaire termen van $f(x_1, \dots, x_n)$ een eerstegraadsfunctie van de parameter λ zijn - de oplossing van (1.4) en (1.5) als een functie van λ op te schrijven. In dit geval is de lange vorm van het A.W. superieur aan het A.B.. Tot slot kunnen we opmerken dat het A.B. waarschijnlijk de voorkeur verdient als f weinig - t.o.v. het aantal variabelen - kwadratische termen bevat; verder leent het A.B. zich gemakkelijk voor uitbreidingen.

7. Literatuur

De meest voor de hand liggende oplossing van het algemene

1) A.W. en A.B. zullen gebruikt worden als de afkorting van resp. algoritme van WOLFE en algoritme van BEALE.

programmeringsmodel (1.1) is het optimaliseren van een kwadratische approximatie van $f(x_1, \dots, x_n)$ onder lineaire approximaties van de $g_k(x_1, \dots, x_n)$ en $x_i \geq 0$. In deze zin kunnen o.a. de A.W. en A.B. als een benaderende oplossingsmethode van (1.1) beschouwd worden. Er zijn echter ook wel algorithmen die een betere oplossing van (1.1) geven.

ZOUTENDIJK's [7] methode van de toelaatbare richtingen is een van de meest algemene. Veel programmeringsalgorithmen - ook lineaire programmeringsmethoden - kunnen nl. als bijzondere gevallen hiervan beschouwd worden. Deze methode is zelfs van toepassing (onder zekere voorwaarden van differentieerbaarheid) als de functies $f(x_1, \dots, x_n)$ en $g_k(x_1, \dots, x_n)$ niet convex zijn. Men kan dan echter niet altijd zeggen of f in het zo gevonden punt absoluut minimaal is. Voor praktijkproblemen is dit niet zo bezwaarlijk, daar men vaak door de aard van het probleem wel weet, of het zo gevonden punt een lokaal of absoluut minimum kan zijn.

HILDRETH [4] reduceert probleem (1.4), (1.5) en (1.6) tot het optimaliseren van een convexe kwadratische functie met niet-negatieve variabelen. De bepaling van het optimum geschiedt - analoog aan het A.B. - door coördinaatsgewijze optimalisering (Univariate method). Bewezen kan worden, dat deze procedure convergeert (het optimum zal i.h.a. niet in een eindig aantal iteraties bereikt worden).

D'ESOPPO [2] bespreekt van bovenstaande methode een generalisatie; de functie $f(x_1, \dots, x_n)$ moet convex en de restricties moeten rechtehoekig zijn (d.w.z. alleen nevenvoorwaarden van de vorm $\alpha_i \leq x_i \leq \beta_i$ zijn toegestaan). In [4] wordt bewezen dat een stelsel lineaire restricties tot een stelsel rechthoekige nevenvoorwaarden gereduceerd kan worden. Het algoritme van D'Esopo is dus, mits $f(x_1, \dots, x_n)$ convex en $g_k(x_1, \dots, x_n)$ lineair zijn, toepasbaar op (1.1). Een andere convexe programmeringsmethode is die van FRANK en WOLFE [3] (hierin zijn $f(x_1, \dots, x_n)$ en de $g_k(x_1, \dots, x_n)$ convex). Deze methode is een generalisatie van hun kwadratisch programmeringsalgoritme, dat in hetzelfde artikel wordt besproken. Laatstgenoemde methode convergeert echter vaak zeer langzaam.

8. Verwijzingen

- [1] BEALE E.M.L. - "On minimizing a convex function subject to linear inequalities". Journal of the Royal Stat. Soc. (series B) 17 (1955) pp. 173-184.
- [2] D'ESOPPO A.A. - "A convex programming procedure". Nav. Log. Quat. 6 (1959) pp. 33-42.
- [3] FRANK M. en WOLFE P. - "An algorithm for quadratic programming". Nav. Res. Log. Quat. 3 (1956) pp. 95-110.
- [4] HILDRETH C.G. - "A quadratic programming procedure". Nav. Res. Log. Quat. 4 (1957) pp. 79-85.
- [5] MARKOWITZ H. - "The optimization of a quadratic function subject to linear constraints". Nav. Res. Log. Quat. 3 (1956) pp. 111-133.
- [6] WOLFE P. - "The simplex method for quadratic programming". Econometrika 27 (1959) pp. 382-398.
- [7] ZOUTENDIJK G. - "Maximizing a function in a convex region. Journal of the Royal Stat. Soc. (series B) 21 (1959) pp. 338-355.