

STICHTING
MATHEMATISCH CENTRUM
2e BOERHAAVESTRAAT 49
AMSTERDAM

AFDELING MATHEMATISCHE STATISTIEK

Rapport S 265 (C13)

Leergang Besliskunde

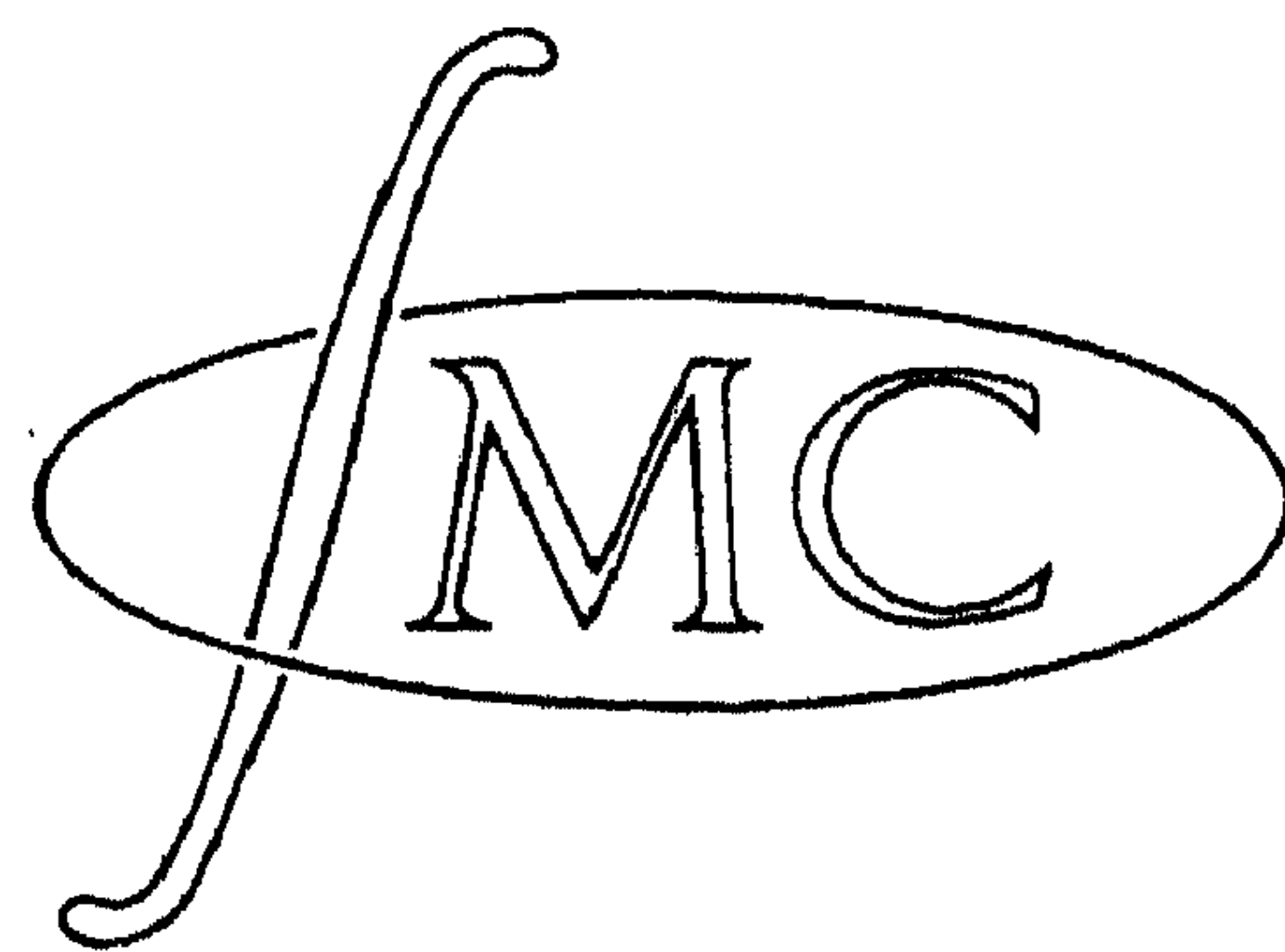
Hoofdstuk XIII

Niet-lineaire Programmering

door

J. Kriens

(Gedeeltelijk een bewerking van hoofdstuk XIII
uit de eerste druk door S.C. van Westrhenen)



2e druk
maart 1963

The Mathematical Centre at Amsterdam, founded the 11th of February 1946, is a non-profit institution aiming at the promotion of pure mathematics and its applications, and is sponsored by the Netherlands Government through the Netherlands Organization for Pure Research (Z.W.O.) and the Central National Council for Applied Scientific Research in the Netherlands (T.N.O.), by the Municipality of Amsterdam and by several industries.

1. Indeling van de wiskundige programmering

Onder wiskundige programmering verstaan wij in dit hoofdstuk het bepalen van extrema van expliciet gegeven reële functies van een eindig aantal variabelen, welke eveneens alleen reële waarden aan kunnen nemen. De te optimaliseren functie kan dus steeds in de vorm

$$y = f(x_1, \dots, x_n) \quad (1.1)$$

geschreven worden, of, afgekort als

$$y = f(X), \quad (1.2)$$

waarin X de n -dimensionale vector $(x_1 x_2 \dots x_n)$ voorstelt.

De wiskundige programmering kan als volgt worden ingedeeld.

- A. Wiskundige programmering zonder bijvoorwaarden: de variabelen, als functie waarvan het extreem bepaald moet worden, kunnen alle, onafhankelijk van elkaar, de gehele as van reële getallen doorlopen.
- B. Wiskundige programmering met bijvoorwaarden: de waarden, welke de variabelen aan kunnen nemen zijn aan bepaalde beperkingen onderworpen.

ad A. Het bepalen van extremen bij een belangrijke groep van problemen, die hieronder valt, is reeds eerder in deze cursus behandeld, n.l. in hoofdstuk II. De stellingen 4.2 en 6.7 uit dat hoofdstuk geven voldoende voorwaarden voor een relatief extreem voor functies van één, resp. twee variabelen, welke tweemaal differentiëerbaar zijn. Analoge stellingen kunnen afgeleid worden voor tweemaal differentiëerbare functies van meer variabelen.¹⁾

Deze methoden kunnen niet zondermeer toegepast worden op functies, die wel continu zijn, maar niet overal differentiëerbaar. Voor een functie $f(x_1)$ van één variabele, welke bij geschikt gekozen intervallen voor x_1 , in alle intervallen een lineaire functie is van x_1 (vgl. fig. 1.1), kan aangetoond worden dat het optimaliseren gelijk

1) Zie o.a. Th. SAATY, Mathematical Methods of Operations Research, Mc. Graw Hill, 1959, blz. 128-129.

waardig is aan het oplossen van een discreet programmeringsprobleem (een omschrijving van deze problemen volgt verderop in deze paragraaf). Hetzelfde kan bewezen worden voor functies $f(x_1, \dots, x_n)$, welke te beschrijven zijn in de vorm

$$f(x_1, \dots, x_n) = \sum_{j=1}^n \varphi_j(x_j) \quad (1.3)$$

waarbij alle functies φ_j de zojuist genoemde eigenschap bezitten.

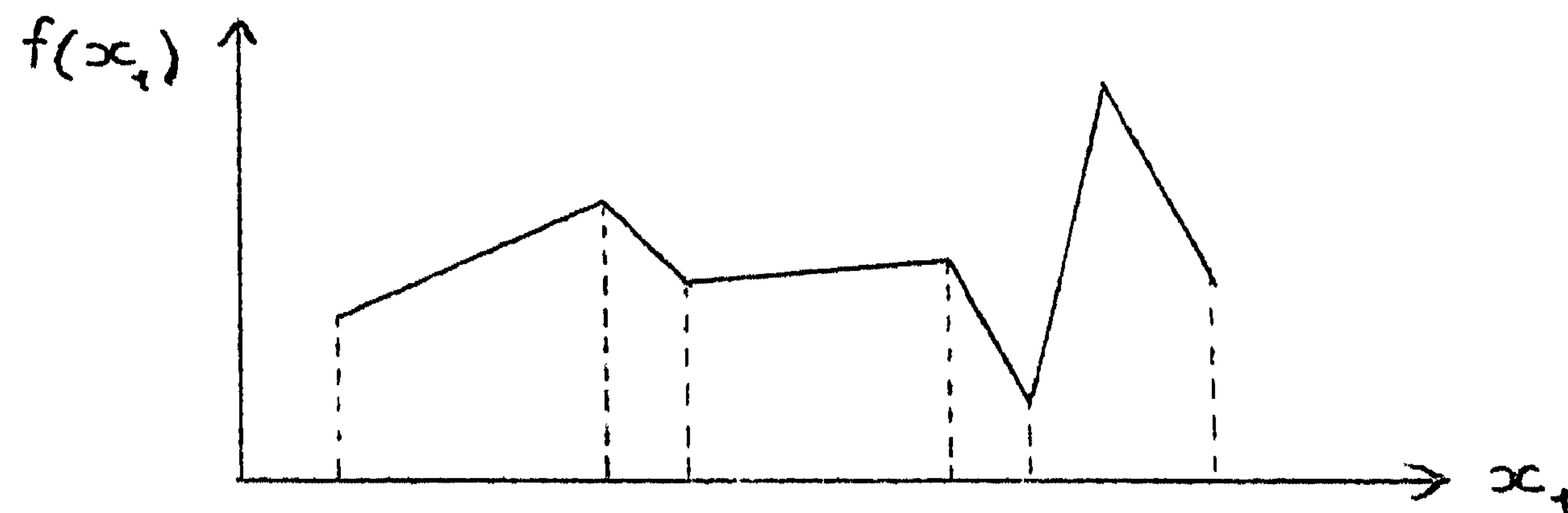


fig. 1.1

Een functie, lineair op bepaalde intervallen.

Het bepalen van extremen van functies, die niet voor alle reële waarden van de variabelen gedefiniëerd zijn, rangschikken wij onder B door te handelen, alsof de variabelen alleen waarden aan mogen nemen in het definitiegebied van de functie.

ad B Noemen wij de verzameling van waarden, die de vector X kan doorlopen, het toegelaten gebied, dan kan deel B verder gesplitst worden door onderscheidt te maken tussen

B 1: het toegelaten gebied is samenhangend,

B 2: het toegelaten gebied is niet samenhangend.

Hierbij noemen wij het toegelaten gebied G samenhangend, wanneer het mogelijk is, vanuit ieder punt uit het gebied te wandelen naar ieder ander punt uit het gebied langs een weg, welke geheel in dit gebied ligt. Wanneer dit niet mogelijk is, heet het gebied niet samenhangend. In fig. 1.2 is G wel samenhangend als het alleen uit de punten van A, of alleen uit de punten van B bestaat, maar niet, wanneer het uit de punten

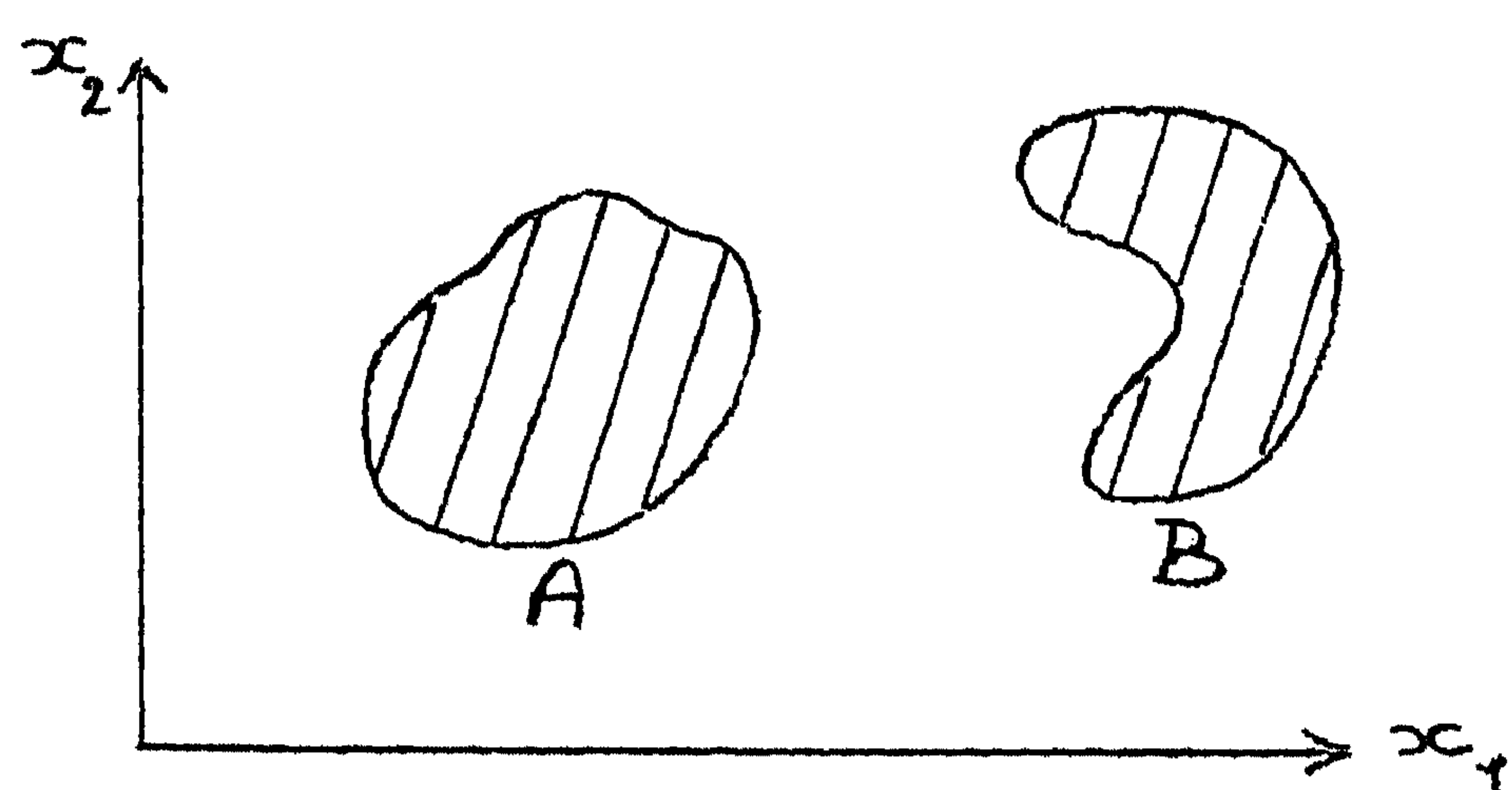


fig. 1.2

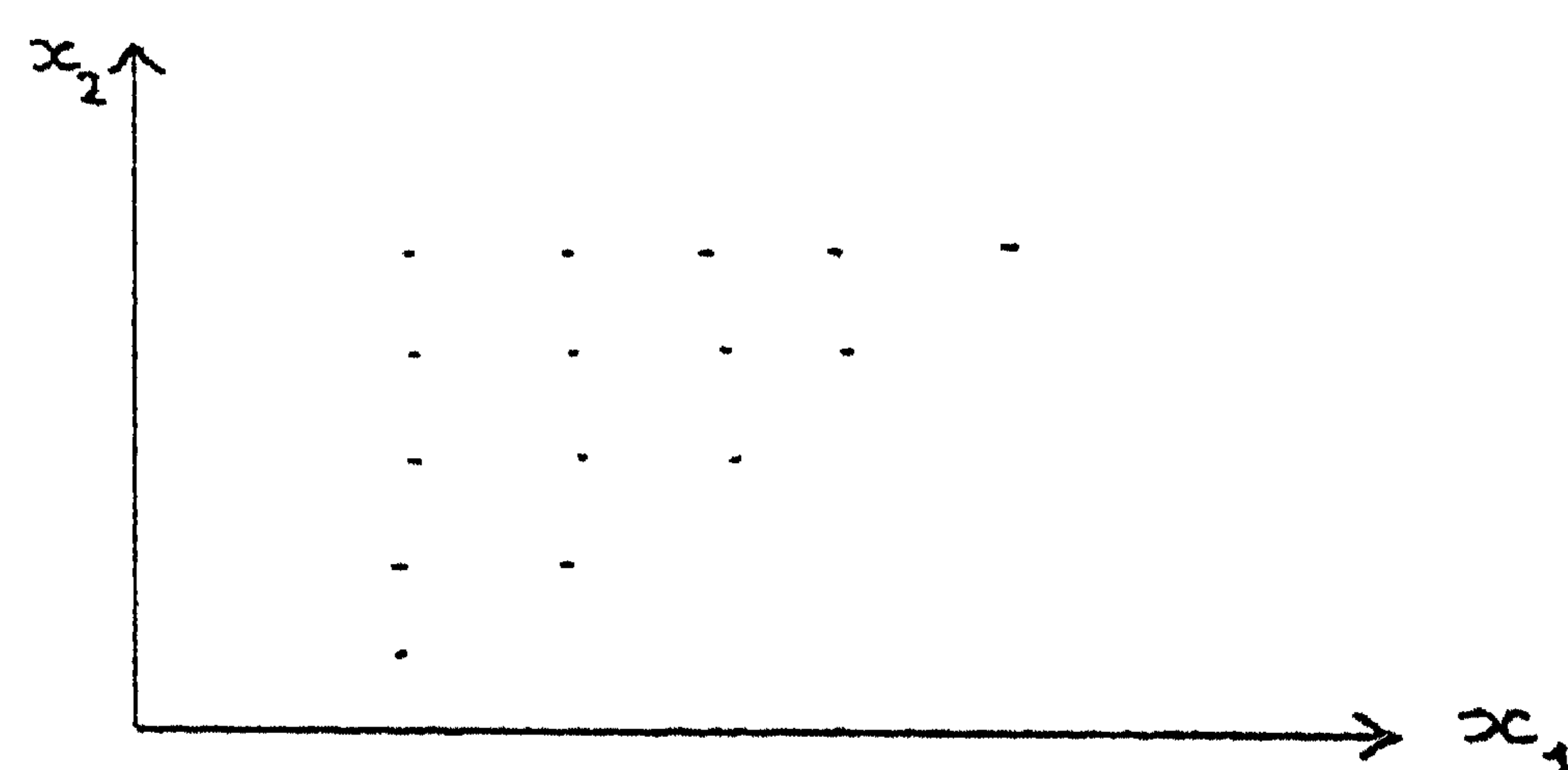


fig. 1.3

Samenhangende en niet samenhangende gebieden.

van A en B tezamen bestaat. De punten, aangegeven in fig.1.3 vormen uiteraard een niet samenhangend gebied.

Tot B1 behoort de lineaire programmering; de functie $f(X)$ is hierbij lineair en gedefiniëerd voor alle reële vectoren X . Het toegelaten gebied wordt vastgelegd door voorwaarden van de vorm $\sum_{j=1}^n a_{ij}x_j \leq a_{i0}$, terwijl bovendien geëist wordt dat alle, of een deel van de variabelen x_j slechts niet negatieve waarden kunnen aannemen. De figuren 1.1, 1.2 en 3.1 uit hoofdstuk XII illustreren, dat het toegelaten gebied bij deze problemen samenhangend is.

Veel algemener dan het lineaire programmeringsprobleem is: optimaliseer $f(X)$ onder voorwaarden van de vorm

$$g_i(x_1, \dots, x_n) = g_i(X) \leq a_{i0} \quad (1.4)$$

waarin $f(X)$ en $g_i(X)$ differentiëerbaar zijn, gedefiniëerd voor alle reële vectoren X . Bij de meeste algoritmen om deze problemen op te lossen stelt men naast de differentiëerbaarheid nog andere eisen aan de functies f en g_i , waardoor het probleem vereenvoudigd wordt. Men eist dan dat de functies g_i behalve differentiëerbaar ook convex zijn en dat f in het minimaliseerprobleem convex is, resp. in het maximaliseerprobleem concaaf (zie voor deze begrippen § 2). Gewoonlijk spreekt men dan van convexe programmering. Is $f(X)$ kwadratisch en bovendien convex (resp. concaaf) en wordt gevraagd f te minimaliseren (resp. te maximaliseren) onder lineaire bijvoorwaarden, dan wordt de term kwadratische programmering gebruikt.

Een voorbeeld van B2 is het geval, waarin de variabelen x_j hetzij 0, hetzij geheel positief moeten zijn. Soms kunnen deze problemen ten gevolge van de vorm van de andere bijvoorwaarden met de onder B1 vallende methoden worden opgelost; vergelijk b.v. § 13 uit hoofdstuk XII. Wanneer dat niet zo is, spreekt men veelal over discrete programmering. Het speciale geval, waarin het extreem van een functie van één of twee variabelen bepaald moet worden, terwijl de variabele(n) 0 of geheel positief moeten zijn, doch verder niet aan voorwaarden zijn onderworpen is reeds kort besproken in § 7.2 van hoofdstuk II en toegepast in o.a. de §§ 5 en 6 van hoofdstuk XI.

De term gemengde programmering heeft betrekking op situaties, waarin sommige variabelen alleen discrete waarden aan

kunnen nemen en andere een samenhangend gebied kunnen doorlopen.

Opmerking 1.1

Naast de bovengenoemde begrippen komt men verder nog de woorden stochastische programmering en dynamische programmering tegen.

Onder stochastische programmering verstaat men meestal programmering, waarbij één of meer van de coëfficiënten van de te optimaliseren functie en / of één of meer coëfficiënten in de bijvoorwaarden stochastische grootheden zijn. Daarnaast bestaan er situaties, waarin de variabelen x_j zelf stochastische grootheden zijn (vgl. o.a. voorbeeld 1.4 uit hoofdstuk VII). Al dergelijke problemen kunnen herleid worden tot gevallen van de eerder in deze paragraaf genoemde typen, b.v. door in plaats van $\underline{f}$ de verwachting $\mathcal{E} \underline{f}$ te optimaliseren, of door een voorwaarde met stochastische coëfficiënten $\underline{g}_i(x_1, \dots, x_n) \leq \underline{a}_{i0}$ te vervangen door een voorwaarde van de vorm $P [\underline{g}(x_1, \dots, x_n) \leq \underline{a}_{i0}] \leq \infty$, in welke laatste geval het probleem veelal in een convex programmeringsprobleem overgaat. Dit is gebeurd in het genoemde voorbeeld uit hoofdstuk VIII. Fraaier zou het uiteraard zijn, wanneer men de verdelingsfunctie van de optimale waarde van $\underline{f}$ kon bepalen. Dit leidt echter tot grote wiskundige moeilijkheden en er bestaat tot nog toe slechts weinig literatuur over.

Het woord dynamische programmering wordt hoofdzakelijk gebruikt in de betekenis, die R. BELLMAN er aan heeft gegeven. Het is dan een oplossingstechniek, waarmee wiskundige programmeringsproblemen kunnen worden opgelost en staat als zodanig dus naast b.v. de simplexmethoden. Tal van programmeringsproblemen kunnen hiermee efficiënter worden opgelost dan met technieken, gebaseerd op simplexmethoden; vergelijk o.a. het eerste voorbeeld uit § 2 van hoofdstuk XIV, in welk hoofdstuk nader op de dynamische programmering wordt ingegaan.

In dit hoofdstuk beperken wij ons verder tot de convexe programmering. Paragraaf 2 bevat de definities van convexe en concave functies en aanwijzingen om na te gaan of functies hieraan voldoen. De belangrijkste stelling van de convexe programmering, de stelling van KUHN en TUCKER wordt in § 3 besproken. Vervolgens bevat § 4 enkele voorbeelden van kwadratische

programmeringsproblemen, terwijl in § 5 het algoritme van WOLFE wordt behandeld om deze problemen op te lossen en in § 6 dat van BEALE.

2. Enige 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.

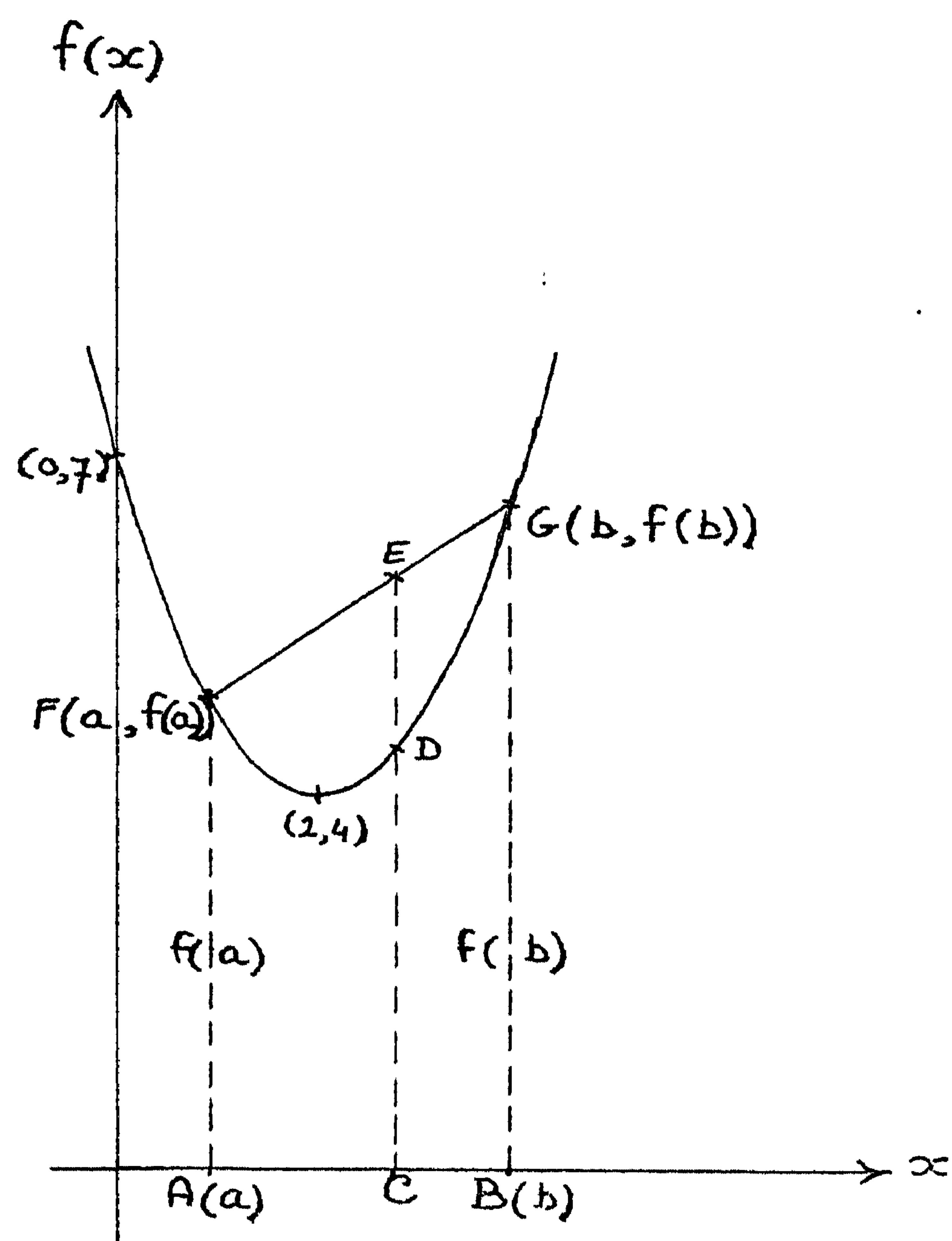


fig. 2.1

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

$f(x)$ is convex

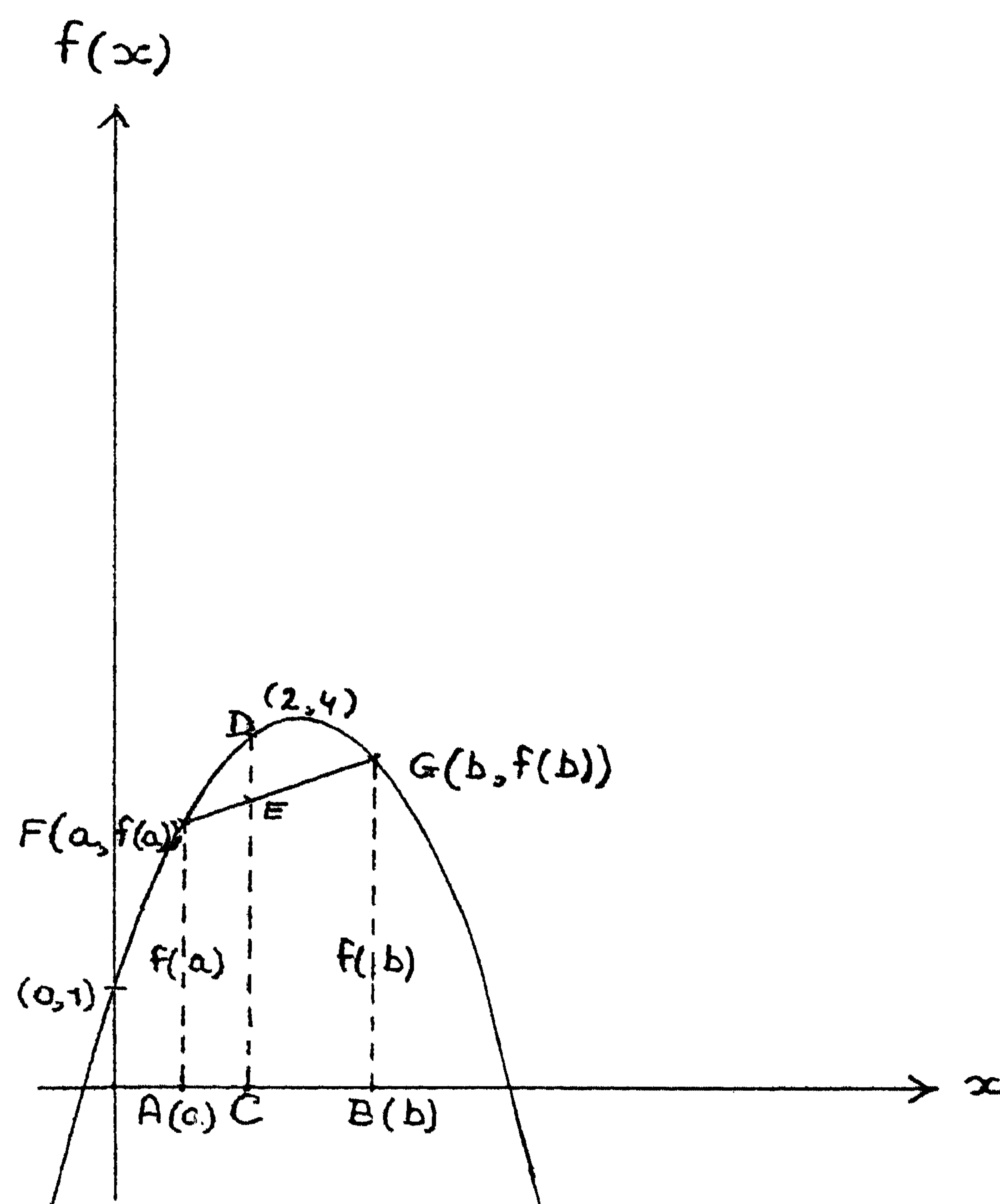


fig. 2.2

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

$f(x)$ is concaaf

Ieder punt E van de koorde FG ($E \neq F$ of G) heeft tot coördinaten $(\overline{1-\lambda} a + \lambda b, \overline{1-\lambda} f(a) + \lambda f(b))$, waarin $0 < \lambda < 1$.¹⁾ Aangezien in fig. 2.1 voor E geldt $EC > DC$ is dus

$$(1-\lambda) f(a) + \lambda f(b) > f(\overline{1-\lambda} a + \lambda b).$$

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

Dit is in overeenstemming met de algebraïsche definitie van convexiteit, te weten: een functie $f(x)$ is convex, wanneer voor alle a en b en alle λ , $0 < \lambda < 1$, geldt

$$(1-\lambda) f(a) + \lambda f(b) \geq f(\overline{1-\lambda} a + \lambda b). \quad (2.1)$$

Een functie $f(x)$ is concaaf, wanneer voor alle a en b en alle λ , $0 < \lambda < 1$, geldt

$$(1-\lambda) f(a) + \lambda f(b) \leq f(\overline{1-\lambda} a + \lambda b) \quad (2.2).$$

Voor functies van meer variabelen kan men convexiteit en concaviteit definiëren met behulp van formules analoog aan (2.1) en (2.2). De algemene definitie luidt dan:

Definitie 2.1. Een functie $f(X)$, gedefiniëerd voor alle reële waarden van $X = (x_1 x_2 \dots x_n)$ is
convex als $f(\overline{1-\lambda} A + \lambda B) \leq (1-\lambda) f(A) + \lambda f(B)$ (2.3),
concaaf als $f(\overline{1-\lambda} A + \lambda B) \geq (1-\lambda) f(A) + \lambda f(B)$ (2.4)
 geldt voor iedere $A = (a_1 a_2 \dots a_n)$, $B = (b_1 b_2 \dots b_n)$,
 $A \neq B$ en iedere λ , $0 < \lambda < 1$.
 De functie $f(X)$ is strikt convex (strikt concaaf)
 als onder bovenstaande voorwaarden alleen het
 ongelijkteken geldt.

In figuur 2.3 is nog een convexe functie getekend van twee variabelen, $y = f(x_1, x_2)$. Het punt E van de koorde FG ligt boven het punt D op het oppervlak $y = f(x_1, x_2)$.

Uit bovenstaande definitie kunnen de volgende eigenschappen direct door substitutie worden afgeleid:

- a) een lineaire functie is convex en concaaf, maar noch strikt convex, noch strikt concaaf,
- b) de som van een aantal convexe (concave) functie s is weer convex (concaaf),
- c) de som van een aantal convexe (concave) functies is strikt convex (concaaf), wanneer tenminste één van de functies strikt convex (concaaf) is,
- d) als $f(X)$ (strikt) convex is, dan is $-f(X)$ (strikt) concaaf en andersom.

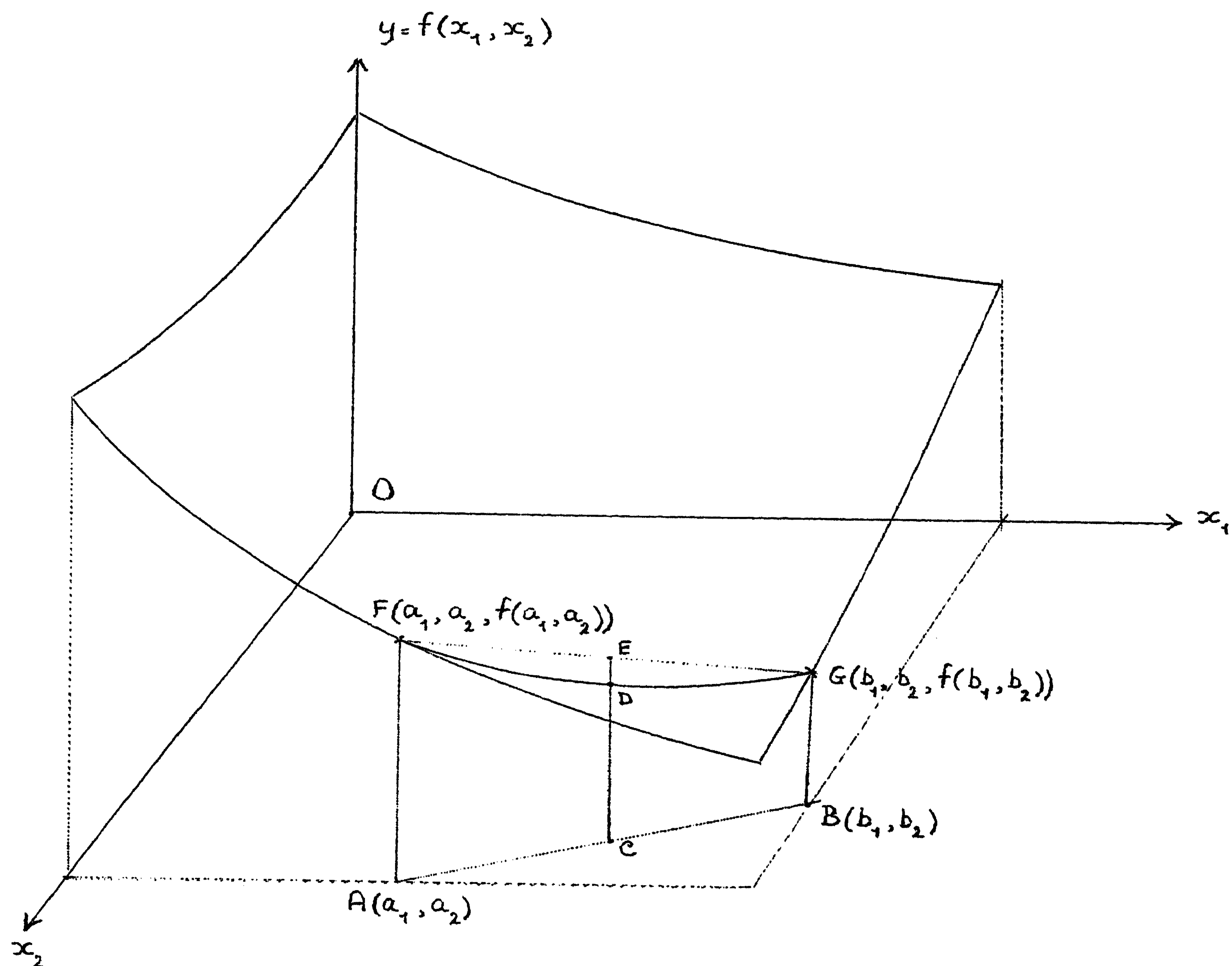


fig. 2.3

Een convexe functie van twee variabelen.

Zoals reeds in § 1 is opgemerkt wordt in de meeste algoritmen voor kwadratische en convexe programmering geëist dat de betreffende functies hetzij convex, hetzij concaaf zijn. Het is dus van belang te beschikken over criteria, waarmee men kan nagaan of een functie al dan niet convex is, of concaaf. De formules (2.3) en (2.4) zijn hiervoor in de praktijk niet erg geschikt. Hieronder wordt daarom een ander criterium nader besproken, doch ter inleiding dienen eerst enige nieuwe begrippen ingevoerd te worden.

Definitie 2.2. Een symmetrische $n \times n$ -matrix $\mathcal{C}$ is positief definit als voor iedere $X \neq 0$ geldt:
 $X \cdot \mathcal{C} \cdot X > 0$,
 respectievelijk positief semi-definit als voor iedere $X \neq 0$ geldt $X \cdot \mathcal{C} \cdot X \geq 0$.

Definitie 2.3. Een symmetrische matrix $\mathcal{C}$ is negatief definitief, resp. negatief semi-definitief, wanneer $-\mathcal{C}$ positief definitief is, respectievelijk positief semi-definitief.

Definitie 2.4. Een symmetrische $n \times n$ -matrix $\mathcal{C}$ is indefinitief als er minstens één paar vectoren X en Y bestaat, zodanig dat $X \cdot \mathcal{C} \cdot X > 0$ en $Y \cdot \mathcal{C} \cdot Y < 0$.

Voorbeeld 2.1.

Ga van de matrices

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

na, aan welke van de in de definities 2.3 t/m 2.4 gegeven eigenschappen zij voldoen.

$\mathcal{A}_1$ is positief definitief, 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 positief semi-definitief, 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 voor $X \neq 0$, b.v. voor $X = (0 \ 2 \ -2)$.

$\mathcal{A}_3$ is negatief semi-definitief 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 b.v. $X = (3 \ -3)$.

$\mathcal{A}_4$ is indefinitief, 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 b.v. $X = (0 \ 0 \ 4)$, $X \mathcal{A}_4 X > 0$ voor b.v. $X = (2 \ -3 \ 0)$ en $X \mathcal{A}_4 X = 0$ voor b.v. $X = (3 \ 0 \ -3)$.

De aard van een symmetrische matrix kan dus onderzocht worden door uit te gaan van bovenstaande definities. In de praktijk

gaat men vaak anders te werk en maakt men gebruik van een bepaalde eigenschap van positief definitie matrices. Deze eigenschap wordt in een voetnoot vermeld, daar hiervoor het begrip determinant nodig is, dat in deze cursus niet is behandeld ²⁾.

Wij kunnen nu overgaan op de behandeling van de eigenschap, waarmee het convex zijn van een functie kan worden onderzocht. Daartoe onderstellen wij dat de functie $f(x_1, \dots, x_n)$ tweemaal partiëel differentiëerbaar is naar alle variabelen. De afgeleiden

$$f''_{x_i x_j} = \frac{\partial^2 f}{\partial x_i \partial x_j}$$

kunnen dus alle bepaald worden en vervolgens opgeschreven in de vorm van een matrix als volgt

$$\mathcal{C}(x_1, \dots, x_n) = \begin{pmatrix} f''_{x_1 x_1} & f''_{x_1 x_2} & \dots & f''_{x_1 x_n} \\ f''_{x_2 x_1} & f''_{x_2 x_2} & \dots & f''_{x_2 x_n} \\ \vdots & \vdots & \ddots & \vdots \\ f''_{x_n x_1} & f''_{x_n x_2} & \dots & f''_{x_n x_n} \end{pmatrix} \quad (2.5).$$

Wanneer $f(X)$ een functie van de tweede graad is, zijn de tweede partiële afgeleiden constanten en dan is $\mathcal{C}$ dus een gewone matrix. In het algemeen zal dit echter niet zo zijn en dan is ieder element van $\mathcal{C}$ zelf weer een functie van $x_1, \dots, x_n$. Wel geldt onder zeer algemene voorwaarden dat $f''_{x_i x_j}$ gelijk is aan $f''_{x_j x_i}$ (vgl. § 6 uit hoofdstuk II), waardoor de matrix (2.5) dus wel symmetrisch is.

-
- 2) Laat een symmetrische $n \times n$ matrix gegeven zijn en laat $\mathcal{C}_k$ de matrix voorstellen, welke uit $\mathcal{C}$ ontstaat door alle elementen weg te laten, die niet zowel in de eerste k rijen als in de eerste k kolommen voorkomen. Er zijn dus n matrices $\mathcal{C}_k$; $k=1, \dots, n$. De determinant van $\mathcal{C}_k$ geven wij aan met $|\mathcal{C}_k|$. De stelling luidt dan: een symmetrische matrix $\mathcal{C}$ is dan en slechts dan positief definit, wanneer alle determinanten $|\mathcal{C}_k|$ positief zijn. Met behulp van deze stelling kan men de aard van symmetrische matrices dus onderzoeken door de betreffende determinanten uit te rekenen. Een voorbeeld staat in voetnoot 3) op blz. 11. Een uitvoeriger behandeling van deze eigenschap kan men o.a. vinden in F.B. HILDEBRAND, Methods of Applied Mathematics, Prentice-Hall, Inc. (1952), hfdst. I, §18.

De stelling luidt nu:

Stelling 2.1. Een tweemaal differentiëerbare functie $f(x_1, \dots, x_n)$ is dan en slechts dan convex, wanneer de matrix $\mathcal{C}(x_1, \dots, x_n)$ van de tweede partiële afgeleiden voor alle waarden van $x_1, \dots, x_n$ positief semi-definiet is. De functie is strikt convex, wanneer deze matrix positief definiet is voor alle waarden van $x_1, \dots, x_n$.

Voorbeeld 2.2.

De functie $e^{2x_1 + \frac{1}{3}x_2 + x_3}$ is convex.

De matrix van de tweede partiële afgeleiden is

$$\mathcal{C}(x_1, x_2, x_3) = \begin{vmatrix} 4 e^{2x_1 + \frac{1}{3}x_2 + x_3} & \frac{2}{3} e^{2x_1 + \frac{1}{3}x_2 + x_3} & 2 e^{2x_1 + \frac{1}{3}x_2 + x_3} \\ \frac{2}{3} e^{2x_1 + \frac{1}{3}x_2 + x_3} & \frac{1}{9} e^{2x_1 + \frac{1}{3}x_2 + x_3} & \frac{1}{3} e^{2x_1 + \frac{1}{3}x_2 + x_3} \\ 2 e^{2x_1 + \frac{1}{3}x_2 + x_3} & \frac{1}{3} e^{2x_1 + \frac{1}{3}x_2 + x_3} & e^{2x_1 + \frac{1}{3}x_2 + x_3} \end{vmatrix}$$

$$= e^{2x_1 + \frac{1}{3}x_2 + x_3} \begin{vmatrix} 4 & \frac{2}{3} & 2 \\ \frac{2}{3} & \frac{1}{9} & \frac{1}{3} \\ 2 & \frac{1}{3} & 1 \end{vmatrix}.$$

Deze matrix is positief semi-definiet voor alle waarden van x_1, x_2 en x_3 , omdat

$$(y_1 \ y_2 \ y_3) \cdot e^{2x_1 + \frac{1}{3}x_2 + x_3} \begin{pmatrix} 4 & \frac{2}{3} & 2 \\ \frac{2}{3} & \frac{1}{9} & \frac{1}{3} \\ 2 & \frac{1}{3} & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} =$$

$$= e^{2x_1 + \frac{1}{3}x_2 + x_3} \cdot (2y_1 + \frac{1}{3}y_2 + y_3)^2 \geq 0 \text{ voor alle waarden van } x_1, x_2, x_3 \text{ en } y_1, y_2, y_3.$$

Voorbeeld 2.3.

De functie $x_1^3 + 2x_1^2x_2 - x_2^3 + x_2$ is niet convex.

Immers de matrix van de tweede afgeleiden is

$$\mathcal{C}(x_1, x_2) = \begin{pmatrix} 6x_1 + 4x_2 & 4x_1 \\ 4x_1 & -6x_2 \end{pmatrix} = 2 \begin{pmatrix} 3x_1 + 2x_2 & 2x_1 \\ 2x_1 & -3x_2 \end{pmatrix}.$$

Bekijken wij nu de vorm

$$(y_1 \ y_2) \begin{pmatrix} 3x_1 + 2x_2 & 2x_1 \\ 2x_1 & -3x_2 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = (3x_1 + 2x_2) \cdot y_1^2 + 4x_1 \cdot y_1 y_2 - 3x_2 \cdot y_2^2,$$

dan is eenvoudig in te zien, dat er waarden van x_1 en x_2 bestaan, waarbij wij waarden van y_1 en y_2 aan kunnen geven, z.d.d.

deze vorm negatief wordt, b.v. $x_1 = +1$, $x_2 = -1$, $y_1 = -2$ en $y_2 = 1$.³⁾

Wij merkten reeds op dat de afgeleiden $f'_{x_i x_j}$ bij een functie van de tweede graad constanten zijn. Nu kan het tweede graads-deel van zo'n functie altijd geschreven worden in de vorm

$\frac{1}{2} \sum_{j=1}^n \sum_{k=1}^n c_{jk} x_j x_k$, of afgekort als $\frac{1}{2} X \mathcal{C} X$, waarin $\mathcal{C}$ de symmetrische matrix

$$\begin{pmatrix} c_{11} & c_{12} & \dots & c_{1n} \\ c_{21} & c_{22} & \dots & c_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ c_{n1} & c_{n2} & \dots & c_{nn} \end{pmatrix} \quad (2.6)$$

voorstelt. De matrix (2.6) is dan ook de matrix van de tweede partiële afgeleiden. Voor dit bijzondere geval krijgt stelling 2.1 de volgende eenvoudige vorm.

Stelling 2.2. De functie $f(X) = \frac{1}{2} X \mathcal{C} X$, waarin $\mathcal{C}$ een symmetrische $n \times n$ matrix is, is dan en slechts dan convex, wanneer $\mathcal{C}$ positief semi-definiet is. De functie is strikt convex, wanneer $\mathcal{C}$ positief definiet is.

Het bewijs van stelling 2.1 zullen wij niet geven. Stelling 2.2. kan natuurlijk uit stelling 2.1 afgeleid worden, maar ook rechtstreeks als volgt.

Bewijs van stelling 2.2.

Stel $A \neq B$ en $0 < \lambda < 1$, dan geldt:

$$\begin{aligned} f((1-\lambda)A + \lambda B) &= ((1-\lambda)A + \lambda B) \mathcal{C} ((1-\lambda)A + \lambda B) = (1-\lambda)^2 A \mathcal{C} A + \lambda(1-\lambda) A \mathcal{C} B + \lambda(1-\lambda) B \mathcal{C} A + \lambda^2 B \mathcal{C} B \\ &= (1-\lambda) A \mathcal{C} A + \lambda B \mathcal{C} B \\ \frac{f((1-\lambda)A + \lambda B) - ((1-\lambda)f(A) + \lambda f(B))}{f((1-\lambda)A + \lambda B) - ((1-\lambda)f(A) + \lambda f(B))} &= \frac{(1-\lambda) A \mathcal{C} A + \lambda B \mathcal{C} B - ((1-\lambda) A \mathcal{C} A + \lambda B \mathcal{C} B)}{f((1-\lambda)A + \lambda B) - ((1-\lambda)f(A) + \lambda f(B))} \\ &= \frac{-\lambda(1-\lambda) A \mathcal{C} A + \lambda(1-\lambda) A \mathcal{C} B + \lambda(1-\lambda) B \mathcal{C} A - \lambda(1-\lambda) B \mathcal{C} B}{f((1-\lambda)A + \lambda B) - ((1-\lambda)f(A) + \lambda f(B))} \\ &= -\lambda(1-\lambda) A \mathcal{C} (A-B) + \lambda(1-\lambda) B \mathcal{C} (A-B) = \\ &= -\lambda(1-\lambda) (A-B) \mathcal{C} (A-B). \end{aligned}$$

3) Zouden wij de stelling uit voetnoot ²⁾ toepassen, dan zou $\mathcal{C}(x_1, x_2)$ voor iedere greep (x_1, x_2) positief definiet zijn, wanneer

$$3x_1 + 2x_2 \text{ en } \begin{vmatrix} 3x_1 + 2x_2 & 2x_1 \\ 2x_1 & -3x_2 \end{vmatrix} = -4x_1^2 - 9x_1 x_2 - 6x_2^2 \text{ voor alle}$$

waarden van x_1 en x_2 positief waren. Men ziet direct in, dat dit niet het geval is.

Wegens $0 < \lambda < 1$ is $\lambda(1-\lambda) > 0$. Wanneer nu gegeven is dat $\mathcal{C}$ positief definit is, dan is dus voor iedere A en B

$$f(\overline{1-\lambda}A + \lambda B) - (1-\lambda)f(A) - \lambda f(B) = -\lambda(1-\lambda)(A-B)\mathcal{C}(A-B) < 0$$

en dus is $f(X)$ strikt convex. Is omgekeerd $f(X)$ strikt convex, dan is voor iedere A en B,

$$f(\overline{1-\lambda}A + \lambda B) < (1-\lambda)f(A) + \lambda f(B)$$

en dus is voor iedere $Y \neq 0$, $Y\mathcal{C}Y > 0$, daar iedere vector Y geschreven kan worden als het verschil van twee vectoren A en B.

Voor het geval $\mathcal{C}$ positief semi-definiet is en $f(X)$ convex, volgt de stelling uit bovenstaande redenering, wanneer men onder de ongelijktekens een = teken plaatst.

Opmerking 2.1. Hoewel in stelling 2.2. $f(X)$ geen lineaire termen bezit, geldt de stelling toch voor alle kwadratische functies. Immers volgens eigenschap a), genoemd na definitie 2.1 is een lineaire functie convex, terwijl volgens eigenschap b) de som van twee convexe functies weer convex is. Tenslotte leert eigenschap c) ons, dat de kwadratische functie strikt convex is, wanneer het tweedegraadsdeel strikt convex is.

Wij besluiten het onderzoek naar het al dan niet convex zijn van een functie met een voorbeeld, dat ook grafisch eenvoudig te verifiëren is.

Voorbeeld 2.4:

1) Neem $\mathcal{C} = \begin{pmatrix} 1 \\ 4 \end{pmatrix}$, $f(X) = x \cdot \frac{1}{4} \cdot x = \frac{1}{4}x^2$. $\mathcal{C}$ is positief definit, terwijl uit fig. 2.4 blijkt dat $f(X)$ convex is. Hetzelfde geldt voor:

2) $\mathcal{C} = \begin{pmatrix} \frac{1}{4} & 0 \\ 0 & \frac{1}{4} \end{pmatrix}$, $f(X) = \frac{1}{4}x_1^2 + \frac{1}{4}x_2^2$; zie fig. 2.5.

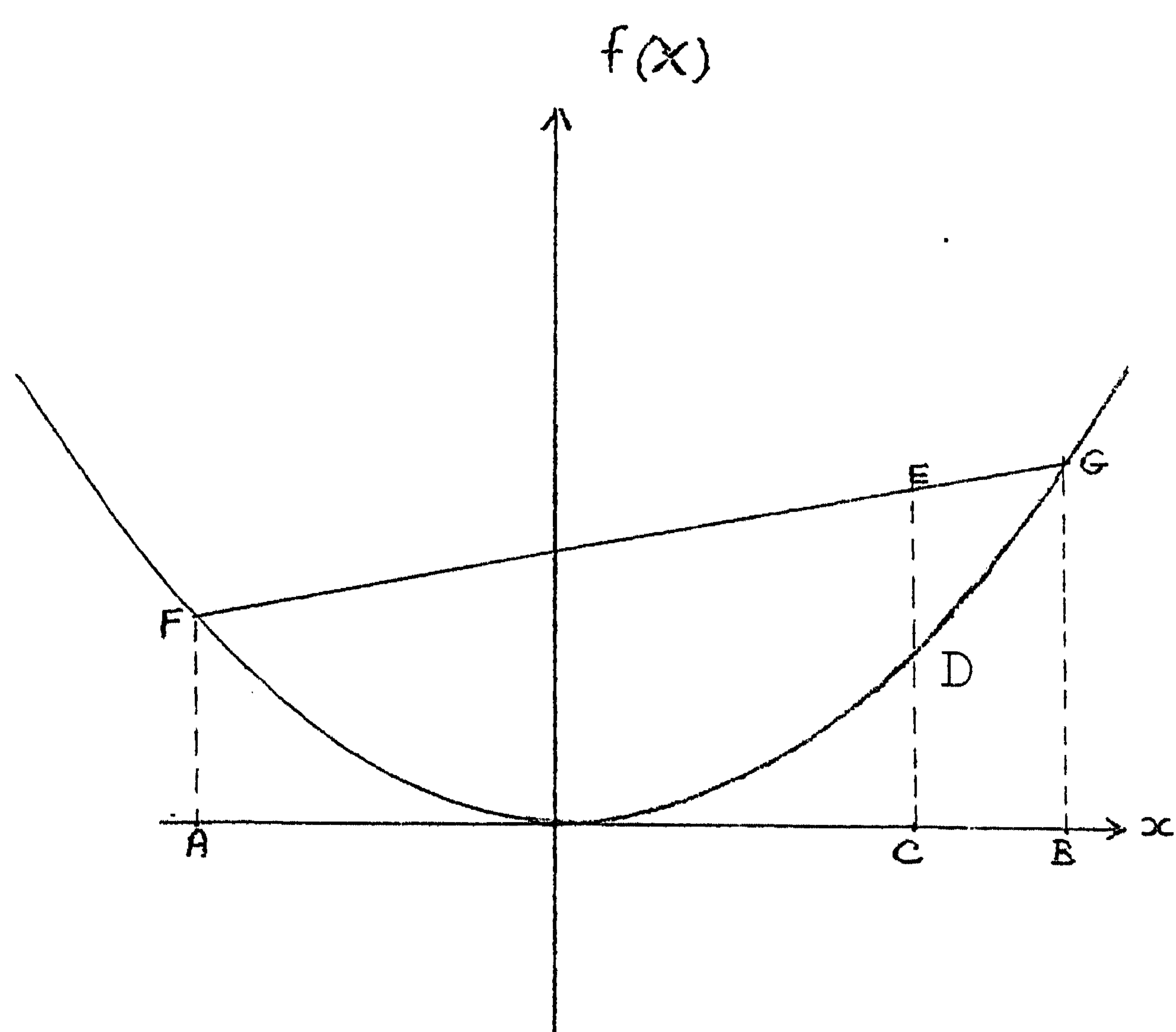


fig. 2.4

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

$f(X)$ is convex

De letters A, B, C, D, E, F en G hebben dezelfde betekenis als in fig. 2.1 t/m 2.3.

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

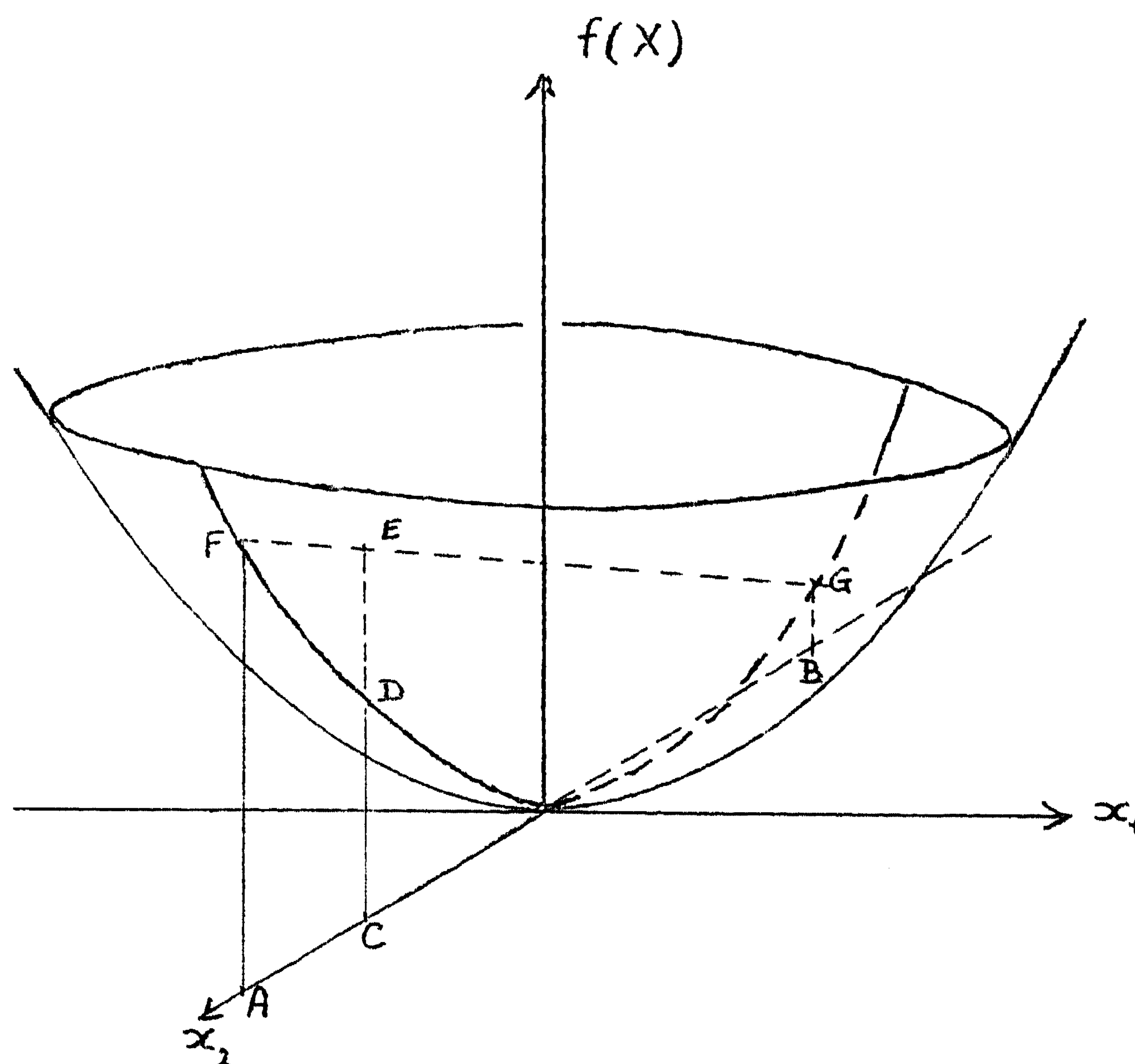


fig. 2.5

$$f(X) = \frac{1}{4} x_1^2 + \frac{1}{4} x_2^2$$

$f(X)$ is convex

Naast de eis, dat de te optimaliseren functie convex resp. concaaf is, neemt men in de meeste algoritmen aan, dat de functies $g_i(X)$ convex zijn. Dit doet men om gebruik te kunnen maken van de volgende stelling.

Stelling 2.3. De vectoren X , welke voldoen aan de voorwaarden $g_i(X) \leq a_{i0}$ ($i=1, \dots, m$) en $X \geq 0$, waarin de functies $g_i(X)$ convex zijn, vormen een convex gebied.

Wij herinneren er aan, dat een gebied convex is, wanneer als de vectoren A en B in dit gebied liggen, ook alle vectoren $\lambda A + (1-\lambda)B$ met $0 < \lambda < 1$ er toe behoren (zie § 3 van hoofdstuk XII). Een convex gebied is dus altijd samenhangend; het omgekeerde behoeft uiteraard niet waar te zijn (vgl. fig 1.2).

Bewijs van stelling 2.3.

Het is duidelijk, dat de vectoren $X \geq 0$ een convex gebied vormen. Evenzo vormen de vectoren X , welke voldoen aan

$g_i(X) \leq a_{i0}$ een convex gebied. Immers, stel dat voor A en B geldt $g_i(A) \leq a_{i0}$ resp. $g_i(B) \leq a_{i0}$, dan is voor $0 < \lambda < 1$

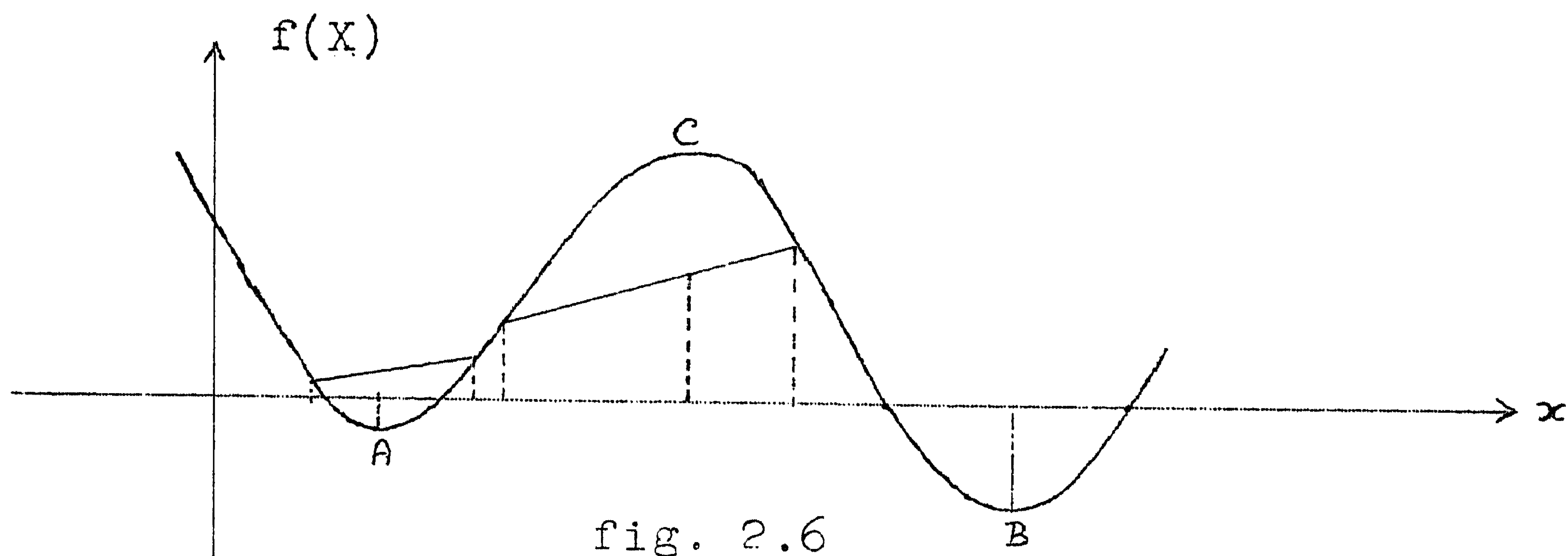
$$g_i((1-\lambda)A + \lambda B) \leq (1-\lambda)g_i(A) + \lambda g_i(B) \leq (1-\lambda)a_{i0} + \lambda a_{i0} = a_{i0},$$

zodat ook de vector $(1-\lambda)A + \lambda B$ voldoet aan de voorwaarde $g_i(X) \leq a_{i0}$. Door alle voorwaarden worden dus convexe gebieden vastgelegd. Het enige wat nog aangetoond moet worden, is dat ook de vectoren, die aan alle voorwaarden tegelijkertijd voldoen een convex gebied vormen. Stel daarom, dat A en B aan alle voorwaarden voldoen. De voorwaarden vormen alle convexe gebieden en dus ligt voor $0 < \lambda < 1$ ook de vector $(1-\lambda)A + \lambda B$ in al die gebieden en voldoet derhalve aan alle voorwaarden. Met andere woorden: ook het door alle voorwaarden bepaalde, toegelaten gebied is convex.

Waarom beperken de algoritmen zich nu meestal tot programmeringsproblemen, waarin hetzij de te minimaliseren functie convex is, hetzij de te maximaliseren functie concaaf, terwijl de functies $g_i(X)$, die het toegelaten gebied bepalen steeds convex verondersteld worden?

Wij bekijken eerst de functie $f(X)$, geschetst in fig. 2.6 en zien, dat een lokaal minimum niet steeds een absoluut minimum behoeft te zijn. Zo is A een lokaal minimum, maar geen absoluut minimum want het andere lokale minimum B ligt lager en is in dit geval wel het absolute minimum. Evenals de simplexmethode voor lineaire programmering, tasten de meeste algoritmen voor kwadratische en convexe programmering steeds de omgeving af van de vector X, die men in een bepaalde stap heeft bereikt, terwijl de berekeningen gestopt worden, wanneer blijkt, dat men in deze omgeving geen betere waarden meer kan vinden. Men vindt in principe dus hoogstens een lokaal extreem en dan is het dus van belang, te weten of dit nu absoluut is, of niet. De volgende stelling leert ons, wanneer een lokaal extreem in ieder geval een absoluut extreem is.

Stelling 2.4 Een lokaal minimum (maximum) van een convexe (concaaf) functie $f(X)$, gedefiniëerd op een convexe verzameling, is een absoluut minimum (maximum) van deze functie.



$f(X)$ is soms convex, soms concaaf.

Bewijs:

Stel dat de convexe functie $f(X)$ voor $X=A$ een lokaal minimum bereikt gelijk aan m , dus $f(A)=m$. Laten wij verder onderstellen, dat $f(X)$ voor $X=B$ een waarde $< m$ bezit.

Daar de toegelaten verzameling convex is, behoren alle punten $(1-\lambda)A + \lambda B$ ($0 < \lambda < 1$) tot deze verzameling. Uit $f(A)=m$ en $f(B)<m$ volgt

$$f(\overline{(1-\lambda)A + \lambda B}) \leq (1-\lambda)f(A) + \lambda f(B) < (1-\lambda)m + \lambda m = m,$$

zodat dus voor alle λ in het interval $0-1$ geldt $f(\overline{(1-\lambda)A + \lambda B}) < m$. Nu ligt de vector $(1-\lambda)A + \lambda B$ voor voldoende kleine λ willekeurig dicht bij A en ook daar zal $f(\overline{(1-\lambda)A + \lambda B})$ dus kleiner zijn dan m . Maar dit is in strijd met de onderstelling, dat f voor $X=A$ een lokaal minimum bezit en dus kan $f(X)$ voor $X=B$ geen kleinere waarde bezitten dan m .

Voor concave functies wordt de stelling geheel analoog bewezen.

Passen wij de algoritmen uit de kwadratische en convexe programmering toe op problemen, waarin aan de in stelling 2.4 gestelde eisen is voldaan, dan vinden wij dus absolute extremen. Is de functie $f(X)$ niet convex, dan leveren sommige algoritmen een relatief minimum op, andere leiden echter in het geheel niet tot een zinvol resultaat. Het analoog geldt voor de bepaling van maxima van niet concave functies.

3. De stelling van KUHN en TUCKER

De stelling, welke de basis vormt van de convexe programmering is de zogenaamde stelling van KUHN en TUCKER, welke in 1950 geplubliceerd werd in de Proceedings van het Tweede Berkeley-symposium ¹⁾. Deze stelling geeft nodige en voldoende voorwaarden opdat een vector $X = \bar{X}$ het optimum is van een convex programmeringsprobleem. De voorwaarden zijn enerzijds theoretisch interessant en leiden anderzijds ook tot numerieke oplossingsmethoden.

Ter inleiding herhalen wij hier de in hoofdstuk II, § 6 besproken methode van LAGRANGE, waarmee de extremen van differentiëerbare functies bepaald kunnen worden als de variabelen moeten voldoen aan één of meer vergelijkingen.

Laten gegeven zijn de differentiëerbare functies $f(x_1, \dots, x_n)$ en $g_i(x_1, \dots, x_n)$, $i=1, \dots, m$ en stel dat gevraagd wordt de extremen van f te bepalen als functie van $x_1, \dots, x_n$, welke variabelen moeten voldoen aan de vergelijkingen

$$g_i(x_1, \dots, x_n) = a_{i0} \quad (i=1, \dots, m) \quad (3.1).$$

Hiertoe vormen wij eerst de LAGRANGE functie

$$F(x_1, \dots, x_n; y_1, \dots, y_m) = f(x_1, \dots, x_n) + \sum_{i=1}^m y_i [g_i(x_1, \dots, x_n) - a_{i0}] \quad (3.2)$$

Men kan nu aantonen, dat de waarden van $x_1, \dots, x_n$ in een voorwaardelijk extreem van f moeten voldoen aan de vergelijkingen, die men krijgt door de eerste partiële afgeleiden van F naar de variabelen $x_1, \dots, x_n$ en $y_1, \dots, y_m$ gelijk aan nul te stellen. Deze voorwaarde is nodig, maar niet voldoende; voor voorwaarden, die nodig en voldoende zijn verwijzen wij naar de literatuur ²⁾.

1) H.W. KUHN and A.W. TUCKER, Non-Linear Programming, Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability, 1950, 481-492.

2) B. v. C. CARATHÉODORY, Variationsrechnung und Partielle Differentialgleichungen Erster Ordnung, B.G. Teubner, 1935, hoofdstuk XI.

De bij een extreem behorende waarden $x_1 = \bar{x}_1, \dots, x_n = \bar{x}_n$ kunnen dus gevonden worden door $n+m$ vergelijkingen op te lossen. De y_i worden de multiplicatoren van LAGRANGE genoemd.

Deze methode kan niet meer ongewijzigd toegepast worden, wanneer de voorwaarden ongelijkheden bevatten en een aantal, of alle variabelen alleen niet negatieve waarden aan mogen nemen.

Wij gaan weer uit van het algemene convexe programmeringsprobleem:

$$\text{minimaliseer } f(X) = f(x_1, \dots, x_n) \quad (3.3)$$

$$\text{onder de voorwaarden } g_i(X) = g_i(x_1, \dots, x_n) \leq a_{i0} \quad (3.4)$$

$$\text{en } X = (x_1, \dots, x_n) \geq 0 \quad (3.5),$$

waarin de functies f en g_i convexe, differentiëerbare functies zijn. Door de laatste onderstelling is aan de voorwaarden van stelling 2.4 voldaan, zodat een lokaal extreem tevens een absoluut extreem is.

De stelling van KUHN en TUCKER zegt dan het volgende.

Stelling 3.1. Een vector $X = \bar{X}$ is dan en slechts dan een oplossing van het probleem (3.3) t/m (3.5), wanneer er een niet negatieve vector $Y = \bar{Y} = (\bar{y}_1, \dots, \bar{y}_m)$ bestaat, zodanig dat

$$\frac{\partial f}{\partial x_j}(\bar{X}) + \sum_{i=1}^m \bar{y}_i \frac{\partial g_i}{\partial x_j}(\bar{X}) \geq 0 \quad (j=1, \dots, n) \quad (3.6)$$

$$\sum_{j=1}^n \left[\frac{\partial f}{\partial x_j}(\bar{X}) + \sum_{i=1}^m \bar{y}_i \frac{\partial g_i}{\partial x_j}(\bar{X}) \right] \bar{x}_j = 0 \quad (3.7)$$

$$\sum_{i=1}^m \left[g_i(\bar{X}) - a_{i0} \right] \bar{y}_i = 0 \quad (3.8).$$

Wij zullen deze stelling hier niet bewijzen. Wel moeten wij opmerken, dat de in de stelling genoemde voorwaarden strikt genomen niet volledig zijn. Er moet namelijk nog een regelmatigheidseis gesteld worden aan het toegelaten gebied van vectoren X , wanneer er zich onder de functies $g_i(X)$ niet lineaire functies bevinden. Aan deze eis is b.v. voldaan als gegeven is, dat er bij iedere niet lineaire bijvoorwaarde een vector X bestaat, die in het toegelaten gebied ligt en waarvoor geldt $g_i(X) < a_{i0}$. Wij eisen dus dat er geen enkele niet lineaire bijvoorwaarde is, waarvan alleen de rand $g_i(X) = a_{i0}$ (gedeeltelijk) tot het toegelaten gebied behoort. Daar hieraan bij praktische problemen altijd wel zal zijn voldaan, laten wij dit punt verder buiten beschouwing.

Toelichting.

De voorwaarden (3.6) t/m (3.8) kunnen, uitgaande van de functie

$$F(x_1, \dots, x_n; y_1, \dots, y_m)$$

uit (3.2), als volgt nader worden verduidelijkt. In de eerste plaats merken wij op, dat men de methode van LAGRANGE zo kan interpreteren, dat men door $F(x_1, \dots, x_n; y_1, \dots, y_m)$ te minimaliseren als functie van $x_1, \dots, x_n$ en te maximaliseren als functie van $y_1, \dots, y_m$, een oplossing $X=\bar{X}$ en $Y=\bar{Y}$ vindt, welke de eigenschap bezit, dat $X=\bar{X}$ het probleem oplost, $f(x_1, \dots, x_n)$ te minimaliseren als functie van $x_1, \dots, x_n$ onder de voorwaarden (3.1). Hiertoe dienen o.a. de afgeleiden van F naar $x_1, \dots, x_n$ nul gesteld te worden en de optimale waarden voldoen dus aan

$$\frac{\partial f}{\partial x_j}(\bar{X}) + \sum_{i=1}^m \bar{y}_i \frac{\partial g_i}{\partial x_j}(\bar{X}) = 0 \quad (3.9).$$

Afgezien van het = teken is deze voorwaarde identiek met (3.6). Ten gevolge van de voorwaarden $x_j \geq 0$ in probleem (3.3) t/m (3.5) heeft daar in het optimum niet aan (3.9) te zijn voldaan. Ter verduidelijking is in fig. 3.1 een convexe functie van één variabele getekend. Moet deze als functie van x

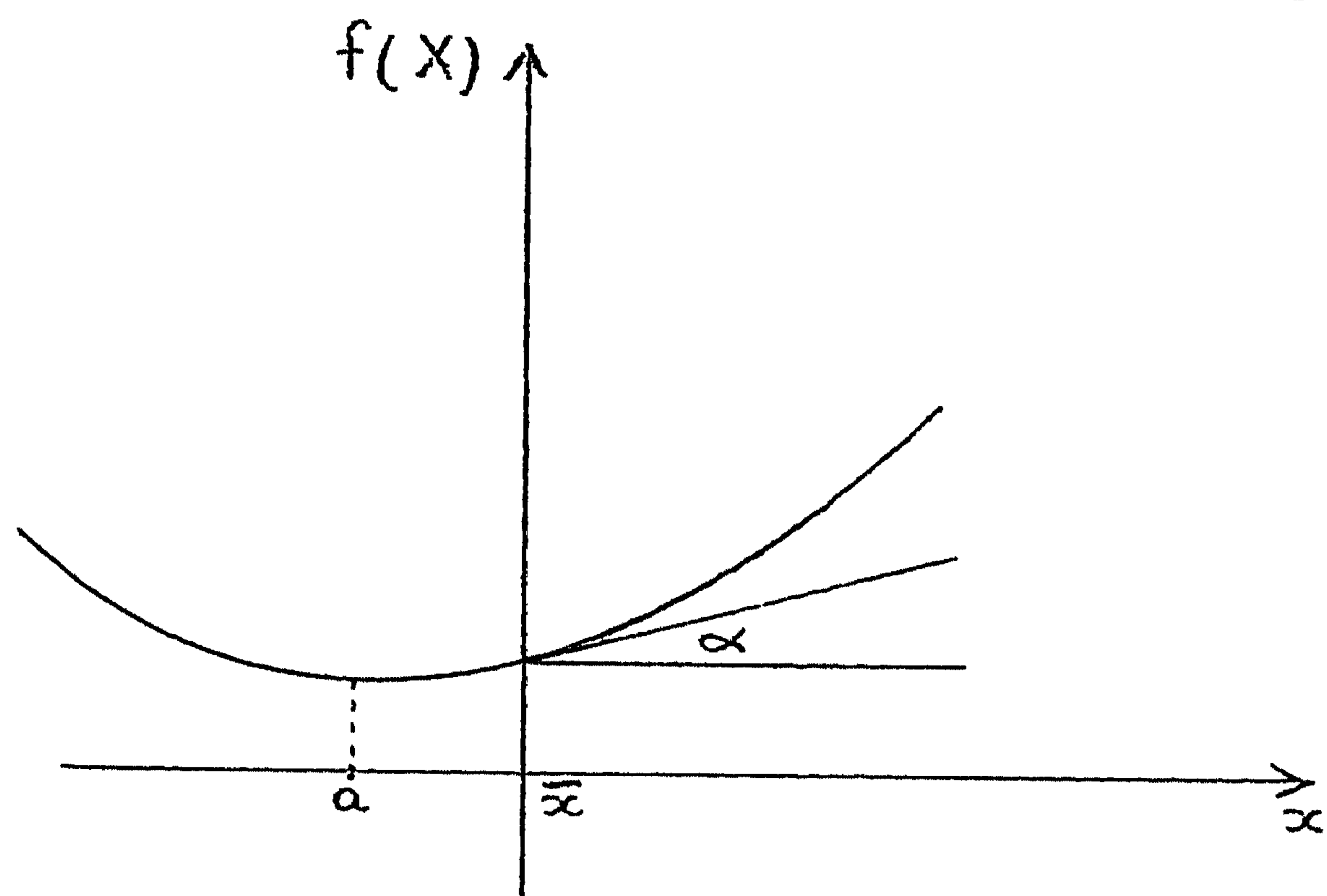


fig. 3.1

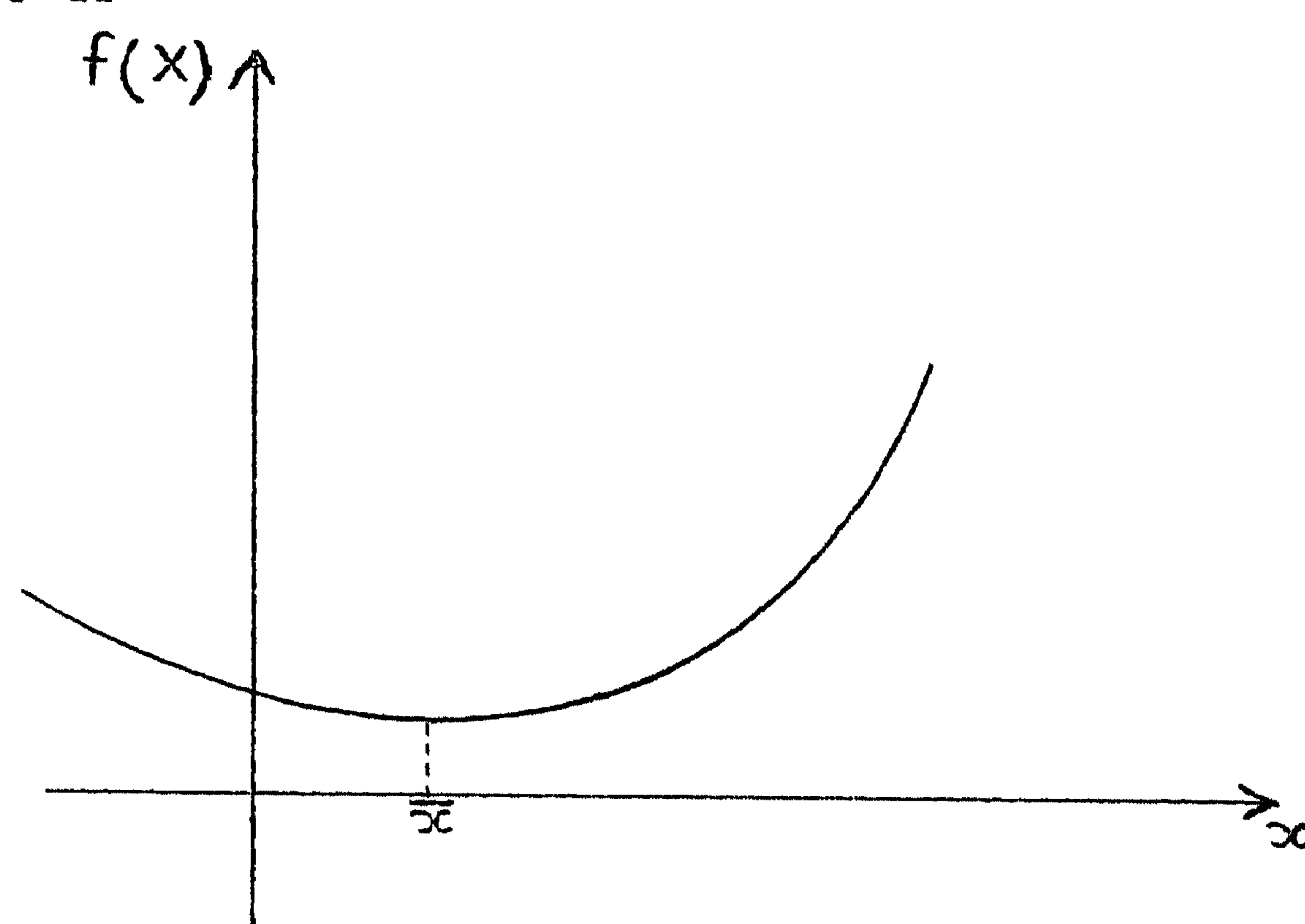


fig. 3.2

Convexe functies, welke geminimaliseerd moeten worden.

worden geminimaliseerd, terwijl x alle waarden mag aannemen, dan wordt dit minimum bereikt voor $x=a$ en is in dat punt $\frac{df}{dx}(a) = 0$. Kan x echter alleen niet negatieve waarden aannemen, dan vinden wij dit voorwaardelijke minimum voor $x=0$. De afgeleide van f naar x is in dit punt positief; $\text{tg } \alpha = \frac{df}{dx}(0) > 0$. Hetzelfde effect doet zich bij de partiële afgeleiden van functies van meer variabelen voor,

wanneer men deze minimaliseert onder de voorwaarde $X \geq 0$. Vandaar dat dan (3.9) vervangen moet worden door de zwakkere voorwaarde (3.6).

Bij de convexe functie, getekend in fig. 3.2 wordt het minimum bereikt voor een positieve waarde van x en dan maakt het dus geen verschil of de voorwaarde $x \geq 0$ aanwezig is of niet; de afgeleide in het minimum blijft ook bij het voorwaardelijke minimum nul.

Uit deze redenering volgt, dat bij een voorwaardelijk minimum van $f(X)$ in het punt $x=\bar{x}$ altijd is voldaan aan

$$\frac{df}{dx}(\bar{x}) \cdot \bar{x} = 0 \quad (3.10).$$

Ook deze voorwaarde handhaaft zich bij functies van meer variabelen. In het minimum van F als functie van $x_1, \dots, x_n$ geldt dus voor iedere partiële afgeleide van $F(x_1, \dots, x_n; y_1, \dots, y_m)$ naar $x_1, \dots, x_n$

$$\frac{\partial F}{\partial x_j}(\bar{X}) \cdot \bar{x}_j = 0,$$

of

$$\left[\frac{\partial f}{\partial x_j}(\bar{X}) + \sum_{i=1}^m \bar{y}_i \frac{\partial g_i}{\partial x_j}(\bar{X}) \right] \cdot \bar{x}_j = 0 \quad (3.11).$$

Maar dan moet in het voorwaardelijke minimum van $f(X)$ ook de som over j van de in (3.11) gegeven produkten nul zijn en dat is juist voorwaarde (3.7).

Een dergelijke interpretatie kan ook aan voorwaarde (3.8) gegeven worden. In het onvoorwaardelijke maximum van $F(x_1, \dots, x_n; y_1, \dots, y_m)$ als functie van de y_i moet voor iedere partiële afgeleide naar resp. $y_1, \dots, y_m$ gelden

$$g_i(\bar{X}) - a_{i0} = 0 \quad (i=1, \dots, m) \quad (3.12).$$

Eist men echter $Y \geq 0$, dan leidt eenzelfde redenering als hierboven is gegeven voor het minimum van een convexe functie van één variabele, maar nu voor het maximum van een concave functie van één variabele tot de voorwaarden

$$\begin{aligned} g_i(\bar{X}) - a_{i0} &\leq 0, \\ \left[g_i(\bar{X}) - a_{i0} \right] \cdot \bar{y}_i &= 0 \end{aligned} \quad (3.13)$$

en

$$\sum_{i=1}^m \left[g_i(\bar{X}) - a_{i0} \right] \bar{y}_i = 0.$$

Uit (3.13) volgt weer, dat voor iedere i minstens één van de factoren $g_i(\bar{X}) - a_{i0}$ en $\bar{y}_i$ gelijk aan nul moet zijn; ligt dus $\bar{X}$ niet op de rand van $g_i(X) \leq a_{i0}$, dan is $\bar{y}_i = 0$ en is $\bar{y}_i > 0$, dan ligt $\bar{X}$ op de rand van het gebied $g_i(X) \leq a_{i0}$.

Bijzondere gevallen van stelling 3.1

1. Wij merkten boven reeds op, dat in plaats van voorwaarde (3.6), voorwaarde (3.9) komt, wanneer men de eis $x_j \geq 0$ laat vervallen. Laat men alle variabelen vrij, dan kan bovendien (3.7) weggelaten worden, omdat er automatisch aan is voldaan.

Luidt voorwaarde (3.4) niet $g_i(X) \leq a_{i0}$, maar

$$g_i(X) = a_{i0} \quad (3.14),$$

voor iedere i , dan is (3.8) overbodig en behoeft men ook niet $Y \geq 0$ te eisen. Vervangt men tegelijkertijd (3.4) door (3.14) en $X \geq 0$ door: X vrij te kiezen, dan vervallen zowel (3.7) als (3.8) en gaat (3.6) over in (3.9), waardoor men het oude probleem van LAGRANGE heeft teruggekregen.

2. Zijn zowel f als de functies g_i lineair, dan volgen uit de formules (3.6) t/m (3.8) de voorwaarden voor een minimum bij een lineair programmeringsprobleem.

Laat $f(X)$ gelijk zijn aan $C \cdot X$ en $g_i(X) \leq a_{i0}$ de vorm $\sum_{j=1}^n a_{ij} x_j \leq a_{i0}$ bezitten, of, alle voorwaarden samengenomen (vgl. (3.7) uit hoofdstuk XII)

$$\sum_{j=1}^n A_j x_j \leq A_0 \quad (3.15).$$

De notatie wijkt enigszins af van die in hoofdstuk XII: de echte variabelen zijn hier genummerd van 1 t/m n , terwijl de in te voeren verschilvariabelen met $v_1, \dots, v_m$ aangegeven zullen worden. De formules (3.6) t/m (3.8) gaan nu over in

$$c_j + \sum_{i=1}^m \bar{y}_i \cdot a_{ij} \geq 0 \quad (j=1, \dots, n) \quad (3.16)$$

$$\sum_{j=1}^n \left[c_j + \sum_{i=1}^m \bar{y}_i a_{ij} \right] \cdot \bar{x}_j = 0 \quad (3.17)$$

$$\sum_{i=1}^m \left[\sum_{j=1}^n a_{ij} \bar{x}_j - a_{i0} \right] \cdot \bar{y}_i = 0 \quad (3.18).$$

In hoofdstuk XII werd de grootheid z_j ingevoerd, welke in iedere stap van het iteratieproces van de simplexmethode gelijk was aan (vgl. (11.4) uit hoofdstuk XII)

$$z_j = C_0 B^{-1} A_j \quad (3.19);$$

hierin is B^{-1} inverse van de matrix van basisvectoren en C_0 dat deel van de vector C , dat correspondeerde met de basisvectoren. Geeft men de waarden van de z_j in de optimale oplossing aan met $\bar{z}_j$, dan kan aangetoond worden, dat hiervoor geldt

$$\bar{z}_j = - \sum_{i=1}^m \bar{y}_i a_{ij} = -\bar{Y} \cdot A_j \quad (3.20).$$

Het bewijs van deze eigenschap valt buiten het bestek van deze syllabus.³⁾ Substitueren wij echter (3.20) in (3.16), dan vinden wij, dat in het minimum van $f(X)$ moet gelden

$$c_j - \bar{z}_j \geq 0 \quad (j=1, \dots, n) \quad (3.21)$$

en dat is juist in overeenstemming met de inhoud van stelling 7.1 uit hoofdstuk XII.

Op blz. 34 van hetzelfde hoofdstuk werd opgemerkt, dat in iedere stap de verschillen $z_j - c_j$ nul waren voor de in de basis opgenomen vectoren A_j ; voor de niet in de basis opgenomen vectoren zijn de bijbehorende variabelen altijd nul. Dus geldt zowel voor de variabelen x_j als de verschilvariabelen v_i in iedere stap, dat het produkt van de variabele met het bijbehorende verschil $z_j - c_j$ nul is. Substitueren wij nu in (3.17) en (3.18) resp. (3.20) en

$$\bar{v}_i = a_{i0} - \sum_{j=1}^n a_{ij} \bar{x}_j \quad (3.22),$$

3) Men kan het vinden in b.v. S. VAJDA, Mathematical Programming, Addison-Wesley Publishing Company, (1961), hoofdstuk 3, § 4.

dan blijkt

$$\sum_{j=1}^n (c_j - \bar{z}_j) \bar{x}_j = 0 \quad (3.23)$$

en

$$\sum_{i=1}^m \bar{v}_i \bar{y}_i = 0 \quad (3.24)$$

te zijn. De genoemde eigenschap wordt dus door (3.17) en (3.18) voor de optimale oplossing tot uitdrukking gebracht.

3. De voorwaarden (3.6) t/m (3.8) vormen ook het uitgangspunt van het algoritme van Ph. WOLFE om kwadratische programmeringsproblemen op te lossen.

Hiertoe is het doelmatig, evenals in § 2 het tweedegraadsdeel van $f(X)$ in een symmetrische vorm te schrijven en wij gaan daarom uit van

$$f(X) = \sum_{j=1}^n c_j x_j + \frac{1}{2} \sum_{j=1}^n \sum_{k=1}^n c_{jk} x_j x_k \quad (3.25),$$

of in vector- en matrixnotatie

$$f(X) = C \cdot X + \frac{1}{2} X^T E X \quad (3.26),$$

waarin E de symmetrische matrix (2.6) voorstelt.

Verder worden in het kwadratische programmeringsprobleem de functies $g_i(X)$ lineair ondersteld. Volgens stelling 3.1 is $X=\bar{X}$ dan en slechts dan een optimale oplossing, wanneer een niet negatieve vector $Y=\bar{Y}$ bestaat, zodanig, dat

$$c_j + \sum_{k=1}^n c_{jk} \bar{x}_k + \sum_{i=1}^m \bar{y}_i a_{ij} \geq 0 \quad (j=1, \dots, n) \quad (3.27),$$

$$\sum_{j=1}^n \left[c_j + \sum_{k=1}^n c_{jk} \bar{x}_k + \sum_{i=1}^m \bar{y}_i a_{ij} \right] \cdot \bar{x}_j = 0 \quad (3.28)$$

en

$$\sum_{i=1}^m \left[\sum_{j=1}^n a_{ij} \bar{x}_j - a_{i0} \right] \bar{y}_i = 0 \quad (3.29).$$

Afgezien van (3.28) en (3.29) is het vinden van vectoren $\bar{X}$ en $\bar{Y}$, die hieraan voldoen aan gewoon lineair programmeringsprobleem. Door een kleine wijziging in het iteratieproces van de simplexmethode aan te brengen, blijkt het dan ook als zodanig te kunnen worden opgelost.

In § 5 zal dit verder worden uitgewerkt. Er wordt daar ondersteld, dat de lineaire bijvoorwaarden vergelijkingen zijn van de vorm

$$\sum_{j=1}^n a_{ij} x_j = a_{i0} \quad (i=1, \dots, m; a_{i0} \geq 0) \quad (3.30).$$

Zoals reeds eerder werd opgemerkt, behoeven de $\bar{y}_i$ dan niet ≥ 0 te zijn en kan voorwaarde (3.29) vervallen. Het probleem luidt dan dus: zoek een vector $X=\bar{X} \geq 0$ en een (willekeurige) vector $Y=\bar{Y}$, zodanig, dat voldaan is aan (3.27), (3.28) en (3.30).

Voorbeeld 3.1

Voorbeeld 1.4 uit hoofdstuk VII komt neer op het maximaliseren van een lineaire functie (vgl. (1.42)) onder een lineaire bijvoorwaarde ($0 \leq \lambda \leq 1$) en een kwadratische bijvoorwaarde (vgl. (1.52)).

Ga na of de daar gegeven oplossingen voldoen aan de voorwaarden van stelling 3.1.

4. Voorbeelden van kwadratische programmeringsproblemen

Kwadratische programmeringsproblemen komen onder andere voor in de volgende situaties.

a. Produktieproblemen ¹⁾

We beschouwen eerst een produktiemodel van een automobielfabriek, welke produceert voor een markt met volledige mededinging (d.w.z. de verkoopprijs van een auto is onafhankelijk van de grootte van de eigen productie). Dit markttype geeft bij het hier beschouwde produktiemodel aanleiding tot een lineair programmeringsmodel.

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

De genoemde fabriek heeft vier afdelingen, waarvoor de grondstoffen onbeperkt toegeleverd kunnen worden. De productie 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 4.I aangegeven.

Tabel 4.I

Produktiecapaciteiten per maand van de afdelingen van een autofabriek

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

¹⁾ Dit voorbeeld is ontleend aan R.DORFMAN, P.A. SAMUELSON and R.M. SOLOW, Linear Programming and Economic Analysis, Mac Graw Hill (1958), hfdst. 6, § 2 en hfdst. 8, § 1.

Voor de produktie van één luxe wagen (truck) per maand heeft men $\frac{1}{25.000}^e$ ($\frac{1}{35.000}^e$) 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 onder bovenstaande restricties gevraagd wordt 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):

$$\begin{aligned} \frac{1}{25.000} x_1 + \frac{1}{35.000} x_2 &= \text{deel gebruikt van de produktiecapaciteit} \\ &\text{van afd. I} \leq 1 \\ \frac{1}{33.333} x_1 + \frac{1}{16.667} x_2 &= \text{" " " " " capaciteit} \\ &\text{van afd. II} \leq 1 \\ x_1 &= \text{produktie van afdeling III} \leq 22.500 \\ x_2 &= \text{" " " " " IV} \leq 15.000. \end{aligned}$$

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:

Maximaliseer $f = 300x_1 + 250x_2$ (4.1)

onder de restricties: $\frac{x_1}{25.000} + \frac{x_2}{35.000} \leq 1$ (4.2)

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

$$x_1 \leq 22.500$$

$$x_2 \leq 15.000$$

$$x_1, x_2 \geq 0.$$

De nevenvoorwaarden (4.2) kunnen grafisch uitgezet worden. Uit fig. 4.1 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 C geeft de maximale winst en punt C de optimale productiecombinatie aan. 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 $x_2=6.481$.

TRUCKS (IN
DUIZENDTALLEN)

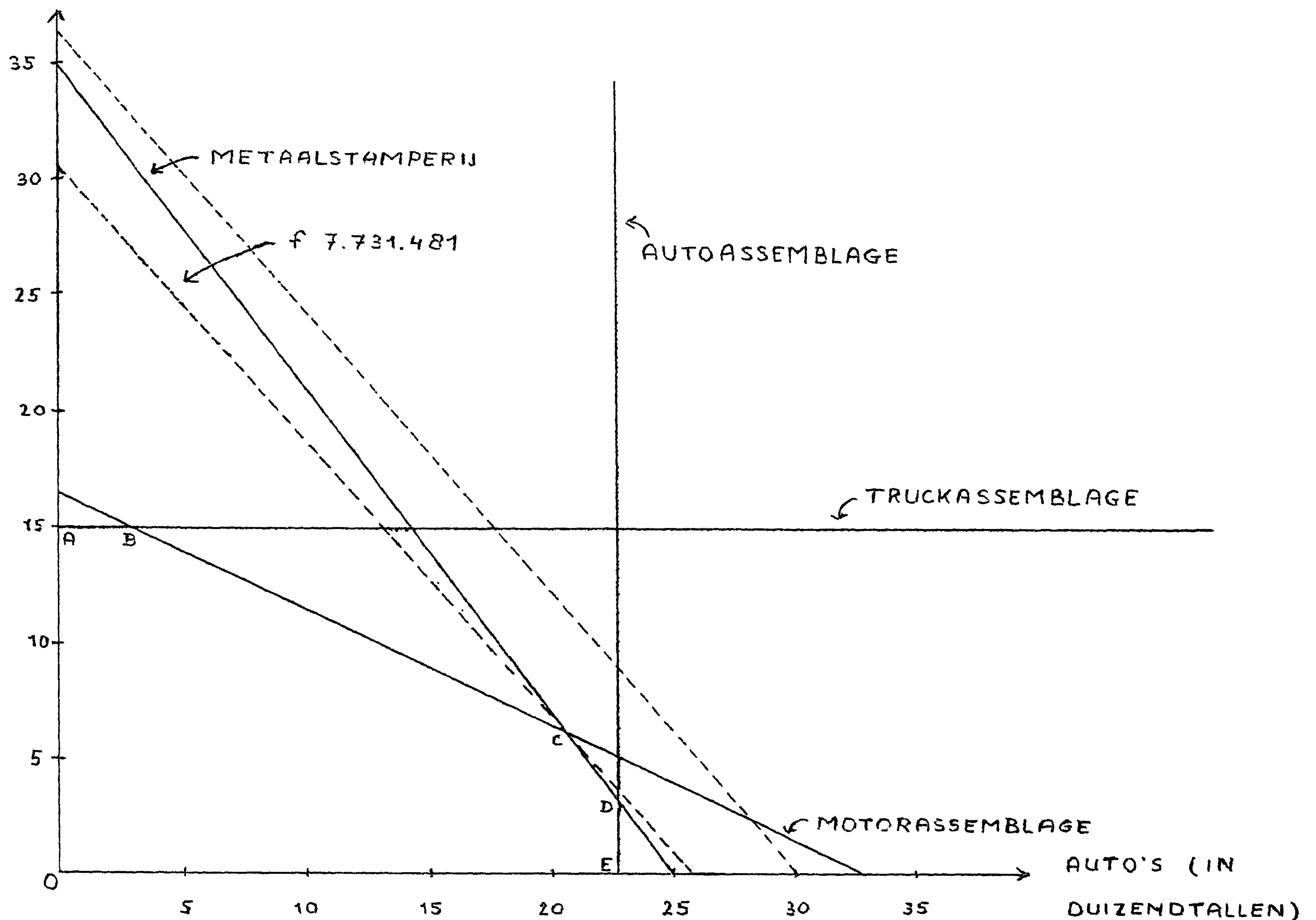


fig. 4.1

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

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:

$$v_1 = 625 - \frac{x_1}{60} \text{ en } v_2 = 250 \quad (4.3).$$

De winstfunctie krijgt dan de vorm:

$$f = v_1 x_1 + v_2 x_2 = (625 - \frac{x_1}{60}) x_1 + 250 x_2 \quad (4.4).$$

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

$$-\frac{1}{60} x_1^2 + 625x_1 + 250x_2 = \text{const.} \quad (4.5).$$

Uit fig. 4.2 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.

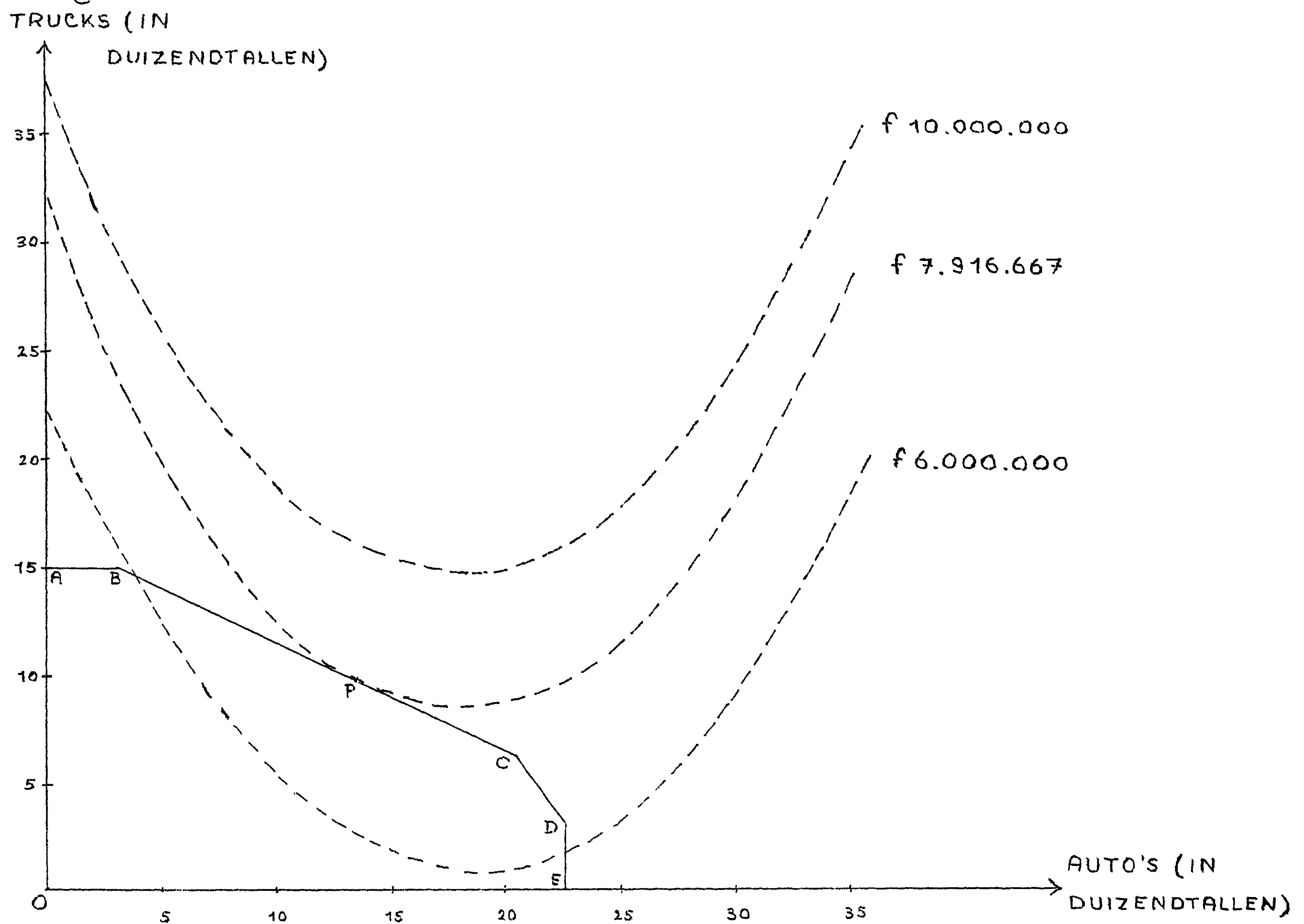


fig. 4.2

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

Opmerking 4.1.

De optimale productiecombinatie in het kwadratische model

behoeft - in tegenstelling tot die in 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 α_1 , α_2 , α_3 en α_4 met verwachtingen en spreidingen van resp. 4,5,6,3 en 1,2,2, 1 gulden.

De belegger wil gemiddeld minstens 4,5% rente van zijn kapitaal maken. Verder wenst hij in de bedrijven B en C niet meer dan 60% van zijn kapitaal te investeren en zal hij trachten onder deze voorwaarden zijn risico zo klein mogelijk te maken. Dit laatste 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 voorgesteld door de stochastische variabele $\underline{d}$:

$$\underline{d} = \alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_3 + \alpha_4 x_4 \quad (4.6).$$

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

$$\mathbb{E} \underline{d} = x_1 \mathbb{E} \alpha_1 + x_2 \mathbb{E} \alpha_2 + x_3 \mathbb{E} \alpha_3 + x_4 \mathbb{E} \alpha_4,$$

of in getallen:

$$\mathbb{E} \underline{d} = 4x_1 + 5x_2 + 6x_3 + 3x_4 \quad (4.7).$$

De eis, dat gemiddeld minstens 4,5% rente gemaakt moet worden houdt de volgende voorwaarde in:

$$4x_1 + 5x_2 + 6x_3 + 3x_4 \geq 45.000 \quad (4.8).$$

Uit stelling 4.4 van hoofdstuk III volgt voor de variantie van $\underline{d}$

$$\sigma^2(\underline{d}) = \sigma^2(\alpha_1 x_1) + \sigma^2(\alpha_2 x_2) + \sigma^2(\alpha_3 x_3) + \sigma^2(\alpha_4 x_4),$$

of

$$\sigma^2(\underline{d}) = x_1^2 \sigma^2(\alpha_1) + x_2^2 \sigma^2(\alpha_2) + x_3^2 \sigma^2(\alpha_3) + x_4^2 \sigma^2(\alpha_4).$$

Vullen wij de opgegeven varianties in, dan is dus:

$$\sigma^2(\underline{d}) = x_1^2 + 4x_2^2 + 4x_3^2 + x_4^2 \quad (4.9).$$

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

$$x_1 + x_2 + x_3 + x_4 = 10.000 \quad (4.10).$$

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

$$x_2 + x_3 \leq 6.000 \quad (4.11).$$

Het kwadratische programmeringsprobleem luidt nu

$$\text{minimaliseer } f = x_1^2 + 4x_2^2 + 4x_3^2 + x_4^2 \quad (4.12)$$

$$\text{onder de voorwaarden } x_1 + x_2 + x_3 + x_4 = 10.000 \quad (4.10)$$

$$x_2 + x_3 \leq 6.000 \quad (4.11)$$

$$4x_1 + 5x_2 + 6x_3 + 3x_4 \geq 45.000 \quad (4.8)$$

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

Opmerking 4.2

Het probleem kan gemakkelijk realistischer gemaakt worden door bijv. koerswaarden in te voeren en de voorwaarde van onafhankelijkheid van $\alpha_1, \dots, \alpha_4$ te laten vervallen. 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:

$$\mathcal{E}(y|x) = \alpha + \beta x \quad (4.13).$$

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

$$f = \sum_{i=1}^n (y_i - \alpha - \beta x_i)^2 \quad (4.14)$$

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

$$\beta \geq 0 \quad (4.15).$$

Na bepaling van α en β kan (4.13) voor de samenstelling van de gevraagde tabel gebruikt worden. α en β volgen uit het kwadratische programmeringsprobleem: minimaliseer (4.14) onder de voorwaarde (4.15).

Men kan f ook iets anders schrijven, als volgt

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

Probleem (4.14), (4.15) luidt dan:

$$\text{Minimaliseer } f = n\alpha^2 + A\beta^2 + 2B\alpha\beta + 2C\alpha + 2D\beta + E \quad (4.16)$$

$$\text{onder de restrictie: } \beta \geq 0 \quad (4.17).$$

In dit eenvoudige voorbeeld kan het antwoord ook zonder kwadratische programmeringstechnieken gevonden worden. Dit is niet meer het geval bij regressieproblemen in een groter aantal variabelen, waarbij de coëfficiënten aan een aantal lineaire bijvoorwaarden moeten voldoen.

5. Het algoritme van WOLFE

In § 3 hebben wij gezien, dat het kwadratische programme-ringsprobleem:

$$\text{minimaliseer } f(x_1, \dots, x_n) = \sum_{j=1}^n c_j x_j + \frac{1}{2} \sum_{j=1}^n \sum_{k=1}^n c_{jk} x_j x_k \quad (5.1),$$

waarin de c_{jk} een positief definitie matrix vormen, onder de voor-waarden

$$\sum_{j=1}^n a_{ij} x_j = a_{i0} \quad (i=1, \dots, m) \quad (5.2)$$

en

$$X = (x_1 \dots x_n) \geq 0 \quad (5.3)$$

opgelost wordt door de vectoren $X \geq 0$ en Y (vrij), welke voldoen aan (5.2),

$$c_k + \sum_{j=1}^n c_{jk} x_j + \sum_{i=1}^m a_{ik} y_i \geq 0 \quad (k=1, \dots, n) \quad (5.4)$$

en

$$\sum_{k=1}^n \left[c_k + \sum_{j=1}^n c_{jk} x_j + \sum_{i=1}^m a_{ik} y_i \right] x_k = 0 \quad (5.5).$$

Het algoritme van WOLFE lost het kwadratische programmerings-probleem langs deze weg op. Er bestaan verschillende versies van het algoritme, waarvan wij er hier één zullen behandelen.

Het uitgangspunt vormen de voorwaarden (5.2) t/m (5.5); even-als in hoofdstuk XII worden de a_{i0} in (5.2) weer ≥ 0 ondersteld. Wij transformeren de ongelijkheden (5.4) tot vergelijkingen door de niet negatieve verschilvariabelen v_k er aan toe te voegen:

$$c_k + \sum_{j=1}^n c_{jk} x_j + \sum_{i=1}^m a_{ik} y_i - v_k = 0 \quad (5.6).$$

Substitutie van (5.6) in (5.5) leidt tot

$$\sum_{k=1}^n v_k x_k = 0 \quad (5.7).$$

In (5.6) worden de constanten c_k naar het rechter lid ge-bracht, terwijl wij in die vergelijkingen van (5.6), waarin $-c_k$ negatief is, het teken van linker- en rechterlid omdraaien. Een-voudigheidshalve onderstellen wij hier, dat alle $-c_k \geq 0$ zijn.

Om in het stelsel (5.2), (5.6) over een volledige eenheidsmatrix te kunnen beschikken voegen wij aan (5.2) de kunstmatige variabelen μ_i ($i=1, \dots, m$) toe en aan (5.6) de kunstmatige variabelen z_{1k} . De vorm van de voorwaarden wordt hierdoor

$$\sum_{j=1}^n a_{ij}x_j + \mu_i = a_{i0} \quad (5.8),$$

resp.

$$\sum_{j=1}^n c_{jk}x_j + \sum_{i=1}^m a_{ik}y_i - v_k + z_{1k} = -c_k \quad (5.9).$$

Bovendien voegen wij aan (5.9) nog niet negatieve variabelen z_{2k} toe met een minteken. Deze zijn niet van belang om een eenheidsmatrix te krijgen, doch spelen verderop een rol. Voorwaarde (5.9) wordt dus

$$\sum_{j=1}^n c_{jk}x_j + \sum_{i=1}^m a_{ik}y_i - v_k + z_{1k} - z_{2k} = -c_k \quad (5.10).$$

Het zoeken naar vectoren X en Y , die aan (5.2) t/m (5.5) voldoen, verloopt nu in twee fasen. In de eerste fase wordt nagegaan of er een vector $X \geq 0$ bestaat, die aan (5.2) voldoet. Hiertoe minimaliseren wij

$$f_1 = \sum_{i=1}^m \mu_i \quad (5.11),$$

uitgaande van de basisoplossing

$$\mu_i = a_{i0}, \quad z_{1k} = -c_k \quad (5.12)$$

van het stelsel (5.8), (5.10). De variabelen y_i en v_k , welke in de beginoplossing alle nul zijn, houden wij tijdens de gehele eerste fase nul, waardoor aan de voorwaarde (5.7) steeds is voldaan.

Is het minimum van $f_1 > 0$, dan kan aan (5.2) niet door een $X \geq 0$ worden voldaan. Deze conclusie is gerechtvaardigd, omdat het niet kan voorkomen, dat er wel een $X \geq 0$ bestaat, die aan (5.8) voldoet, maar niet aan (5.8) en (5.10) tezamen. Bij iedere waarde van het eerste deel van (5.10) kunnen namelijk niet negatieve waarden van de variabelen z_{1k} en z_{2k} gevonden worden, zodanig dat aan (5.10) wordt voldaan, dit tengevolge van de ingevoerde varia-

belen z_{2k} . Vinden wij voor het minimum van f_1 de waarde nul, dan wordt met de tweede fase begonnen.

Opmerking 5.1

Voor hetgeen men moet doen, wanneer er μ -variabelen in de basis blijven met waarde 0, verwijzen wij naar § 9 van hoofdstuk XII.

Opmerking 5.2

Wanneer de oplossing van de eerste fase meer dan n van de variabelen z_{1k} en z_{2k} bevat, kan dit aantal altijd teruggebracht worden tot n door z -variabelen te vervangen door x -variabelen.

In de tweede fase wordt

$$f_2 = \sum_{k=1}^n z_{1k} + \sum_{k=1}^n z_{2k} \quad (5.13)$$

geminimaliseerd onder de voorwaarden (5.2), (5.7) en (5.10). Alle μ -variabelen worden geschrapt en bovendien de z -variabelen, die aan het einde van de eerste fase niet in de basis zijn opgenomen. De variabelen y_i en v_k worden niet langer gelijk aan nul gehouden.

Afgezien van (5.7) bestaan de voorwaarden nu uit $m + n$ vergelijkingen met n variabelen v_k , n variabelen x_j , m variabelen y_i en een aantal z variabelen. Daar de stelling van KUHN en TUCKER zowel nodige als voldoende voorwaarden geeft, moeten er waarden $v_k = \bar{v}_k$ (≥ 0), $x_j = \bar{x}_j$ (≥ 0) en $y_i = \bar{y}_i$ bestaan, die aan (5.2), (5.6) en (5.7) voldoen. Het is dus mogelijk f_2 tot nul te reduceren. Bovendien volgt uit (5.7), dat hoogstens n van de variabelen v_k en x_j in de oplossing $\neq 0$ zijn. Samen met de m variabelen y_i vormen deze een basisoplossing van (5.2) en (5.6). Wij behoeven dus alleen de basisoplossingen van dit stelsel af te tasten om de gezochte oplossing te vinden. In hoofdstuk XII hebben wij gezien, dat dit kan met de simplexmethode.

In de tweede fase gebruiken wij deze methode dan ook om (5.13) onder de voorwaarden (5.2), (5.7) en (5.10) te minimaliseren. Hierbij kan men de vrije y -variabelen eerst uit het probleem trachten te verwijderen, door ze zoveel mogelijk uit de vergelijkingen (5.10) te elimineren. Tenzij alle uit de a_{ik} te vormen $m \times m$ matrices singulier zijn, betekent dit, dat alle y 's verdwijnen

en het aantal voorwaarden met m wordt verminderd.

Men kan echter ook de y variabelen in het probleem houden en beginnen ze in de basis op te nemen met hetzij een positief teken, hetzij een negatief teken en zodra ze in de basis zijn, ze daar te laten.

In beide gevallen gaat men daarna één van de in hoofdstuk XII beschreven methoden toepassen, met dien verstande, dat een variabele v_k niet in de basis wordt opgenomen, wanneer x_k er al in zit, en omgekeerd. Door deze keuzeregels blijft steeds aan voorwaarde (5.7) voldaan. Hoewel de bij de gewone simplexmethoden geldende bewijzen niet meer rechtstreeks opgaan, kan men ook hier aantonen, dat het minimum in een eindig aantal stappen wordt bereikt. Men heeft het gevonden, zodra f_2 niet meer verkleind kan worden, rekening houdend met de keuzeregels.

Samenvatting

Men beschikt bij de methode van WOLFE over een groot tableau, afgeleid uit de vergelijkingen (5.8) en (5.10) en over twee te minimaliseren functies. In tabel 5.I zijn alle bij de variabelen $x_j, y_i, v_k, \mu_i, z_{1k}$ en z_{2k} behorende vectoren opgenomen, benevens de

Tabel 5.I

Het volledige tableau, behorende bij de methode van WOLFE

f_1	0		0	0		0	0		0	1		1	0		0	0		0
f_2	0		0	0		0	0		0	0		0						
	x_1		x_n	y_1		y_m	v_1		v_n	μ_1		μ_m	z_{11}		z_{1n}	z_{21}		z_{2n}
A_0	A_1		A_n	A_{n+1}		A_{n+m}	A_{n+m+1}		A_{2n+m}	A_{2n+m+1}		A_{2n+2m}	$A_{2n+2m+1}$		A_{3n+2m}	$A_{3n+2m+1}$		A_{4n+2m}
a_{10}	a_{11}		a_{1n}							1								
$\vdots$	$\vdots$		$\vdots$							$\diagdown$								
a_{m0}	a_{m1}		a_{mn}									1						
$-c_1$	c_{11}		c_{n1}	a_{11}		a_{m1}	-1						1			-1		
$\vdots$	$\vdots$		$\vdots$	$\vdots$		$\vdots$	$\diagdown$						$\diagdown$			$\diagdown$		
$-c_n$	c_{1n}		c_{nn}	a_{1n}		a_{mn}	-1								1			-1
		n			m			n			m			n			m	

coëfficiënten van deze variabelen in de functies f_1 en f_2 . De $4n + 2m$ vectoren zijn alle aangegeven met een A en doorlopend genummerd.

Het deeltableau voor de eerste fase ontstaat door de vectoren A_{n+1} t/m A_{2n+m} te schrappen en f_1 te minimaliseren, dat voor de tweede fase door de vectoren A_{2n+m+1} t/m A_{2n+2m} en die vectoren van $A_{2n+2m+1}$ t/m A_{4n+2m} , welke aan het eind van de eerste fase niet in de basis zitten te schrappen en f_2 te minimaliseren.

Voorbeeld 5.1

Wij illustreren één en ander aan de hand van een eenvoudig probleem:

$$\text{maximaliseer } x_1(4-x_1) + (2-x_2)x_2 \quad (5.14)$$

onder de voorwaarden

$$x_1 + x_2 \leq 2 \quad (5.15)$$

en

$$x_1 \geq 0, x_2 \geq 0 \quad (5.16).$$

Om het probleem in de vorm (5.1) t/m (5.3) te krijgen, vervangen wij (5.14) t/m (5.16) door

$$\text{minimaliseer } f(x_1, x_2) = -4x_1 - 2x_2 + \frac{1}{2}(2x_1^2 + 2x_2^2) \quad (5.17)$$

onder de voorwaarden

$$x_1 + x_2 + x_3 = 2 \quad (5.18)$$

en

$$x_1, x_2, x_3 \geq 0 \quad (5.19).$$

Met behulp van stelling 2.2 kan men direct zien, dat $f(x_1, x_2)$ inderdaad convex is. De vergelijkingen (5.8) en (5.10) krijgen de vorm

$$x_1 + x_2 + x_3 + \mu = 2 \quad (5.20)$$

en

$$\begin{aligned} 2x_1 + y_1 - v_1 + z_{11} - z_{21} &= 4 \\ 2x_2 + y_2 - v_2 + z_{12} - z_{22} &= 2 \\ y_3 - v_3 + z_{13} - z_{23} &= 0 \end{aligned} \quad (5.21)$$

Het volledige ziet er dus als volgt uit:

Tabel 5.II

Het volledige tableau, behorende bij voorbeeld 5.1

f_1	0	0	0	0	0	0	0	1	0	0	0	0	0	0
f_2	0	0	0	0	0	0	0	0						
	x_1	x_2	x_3	y	v_1	v_2	v_3	μ	z_{11}	z_{12}	z_{13}	z_{21}	z_{22}	z_{23}
A_0	A_1	A_2	A_3	A_4	A_5	A_6	A_7	A_8	A_9	A_{10}	A_{11}	A_{12}	A_{13}	A_{14}
2	1	1	1	0	0	0	0	1	0	0	0	0	0	0
4	2	0	0	1	-1	0	0	0	1	0	0	-1	0	0
2	0	2	0	1	0	-1	0	0	0	1	0	0	-1	0
0	0	0	0	1	0	0	-1	0	0	0	1	0	0	-1

In de eerste fase wordt

$$f_1 = \mu \quad (5.22)$$

geminimaliseerd onder de voorwaarden

$$\begin{array}{rclcl}
 x_1 + x_2 + x_3 + \mu & & & & = 2 \\
 2x_1 & + z_{11} & -z_{21} & & = 4 \\
 & & & & (5.23). \\
 2x_2 & + z_{12} & -z_{22} & & = 2 \\
 & +z_{13} & -z_{23} & & = 0
 \end{array}$$

Gezien het grote aantal nullen, ligt het voor de hand de herziene simplexmethode te gebruiken, voor welke methode wij, evenals voor de notatie verwijzen naar § 11 van hoofdstuk XII. De bijbehorende tableau's zijn opgenomen in de tabellen 5.III en 5. IV; lege vakjes geven nullen aan. In tabel 5.III zijn ook de $z_j - c_j$ waarden vermeld; de vector die in de basis komt is aangegeven door het bijbehorende $z_j - c_j$ hokje te arcen. Daar de μ variabele in de eerste stap nul wordt, is het tweede tableau in tabel 5. IV het eindtableau van de eerste fase.

Tabel 5. III

De vectoren in de eerste fase

c_j				1						
	A_1	A_2	A_3	A_8	A_9	A_{10}	A_{11}	A_{12}	A_{13}	A_{14}
1	1	1	1	1						
2					1			-1		
		2				1			-1	
							1			-1
$z_j - c_j$		1	1							

Tabel 5. IV

De iteratiestappen in de eerste fase

C_o	B	X_o	B^{-1}	X_1
1	μ	2	1	1
	z_{11}	4	1	2
	z_{12}	2	1	
	z_{13}	0	1	
$C_o B^{-1}$			1	

C_o	B	X_o	B^{-1}	X
	x_1	2	1	
	z_{11}	0	-2 1	
	z_{12}	2	1	
	z_{13}	0	1	
$C_o B^{-1}$				

De tweede wordt begonnen met het schrappen van de vectoren A_8, A_{12}, A_{13} en A_{14} . Wij minimaliseren nu

$$f_2 = z_{11} + z_{12} + z_{13} \quad (5.24)$$

onder de voorwaarden

$$\begin{array}{rcll} x_1 + x_2 + x_3 & & & = 2 \\ 2x_1 & +y & -v_1 & +z_{11} = 4 \\ & 2x_2 & +y & -v_2 +z_{12} = 2 \\ & & +y & -v_3 +z_{13} = 0 \end{array} \quad (5.25)$$

en

$$v_1 x_1 + v_2 x_2 + v_3 x_3 = 0 \quad (5.26).$$

Als eerste basis gebruiken wij de vectoren A_1, A_9, A_{10} en A_{11} , waarvan de inverse staat in de laatste tabel van fase I. Verder worden in deze tabel de nieuwe C_0 en $C_0 B^{-1}$ ingevuld. Bij het transformeren houden wij ons aan de keuzeregels, dat een v_k niet in de basis kan komen als x_k er in is en omgekeerd. Tabel 5. V bevat de vectoren en de $z_j - c_j$ waarden van de tweede fase, tabel 5. VI de bijbehorende tableau's; blanco vakjes geven weer nullen aan.

Tabel 5.V

De vectoren in de tweede fase

c_j								1	1	1
	A_1	A_2	A_3	A_4	A_5	A_6	A_7	A_9	A_{10}	A_{11}
1	1	1	1							
2				1	-1			1		
		2		1		-1			1	
				1			-1			1
$z_j - c_j$			-2	3	-1	-1	-1			
		6	4		2	-1	-1	-3		
			-2		-1	-1	2			-3

Tabel 5.VI

De iteratiestappen in de tweede fase

→

C_o	B	X_o	B^{-1}				X_4
	x_1	2	1				0
1	z_{11}	0	-2	1			1
1	z_{12}	2			1		1
1	z_{13}	0				1	1
$C_o B^{-1}$			-2	1	1	1	

→

C_o	B	X_o	B^{-1}				X_2
	x_1	2	1				1
	y	0	-2	1			-2
1	z_{12}	2	2	-1	1		4
1	z_{13}	0	2	-1		1	2
$C_o B^{-1}$			4	-2	1	1	

→

C_o	B	X_o	B^{-1}				X_7
	x_1	2		$\frac{1}{2}$	$-\frac{1}{2}$		$\frac{1}{2}$
	y	0			1		-1
1	z_{12}	2	-2	1	1	-2	2
	x_2	0	1	$-\frac{1}{2}$		$\frac{1}{2}$	$-\frac{1}{2}$
$C_o B^{-1}$			-2	1	1	-2	

C_o	B	X_o	B^{-1}				
	x_1	$\frac{3}{2}$	$\frac{1}{2}$	$\frac{1}{4}$	$-\frac{1}{4}$		
	y	1	-1	$\frac{1}{2}$	$\frac{1}{2}$		
	v_3	1	-1	$\frac{1}{2}$	$\frac{1}{2}$	-1	
	x_2	$\frac{1}{2}$	$\frac{1}{2}$	$-\frac{1}{4}$	$\frac{1}{4}$		
$C_o B^{-1}$							

In het laatste tableau van tabel 5.VI zijn alle z variabelen uit de basis verdwenen en f_2 is dus 0, waarmee de gezochte oplossing is gevonden. De waarden, die (5.14) maximaliseren zijn $x_1 = \frac{3}{2}$ en $x_2 = \frac{1}{2}$; de waarde van het maximum bedraagt $4\frac{1}{2}$.

Opmerking 5.3

Wanneer er ongelijkheden onder de bijvoorwaarden voorkomen, zoals (5.15) in voorbeeld 5.1, dan bespaart men variabelen en vergelijkingen, wanneer men uitgaat van de KUHN - TUCKER condities in de vorm (3.27) t/m (3.29). Luiden de bijvoorwaarden bijvoorbeeld

$$\sum_{j=1}^n a_{ij}x_j \leq a_{i0} \quad (i=1, \dots, l) \quad (5.27)$$

$$\sum_{j=1}^n a_{ij}x_j = a_{i0} \quad (i=l+1, \dots, m) \quad (5.28),$$

dan krijgen (3.27) t/m (3.29) de vorm

$$c_k + \sum_{j=1}^n c_{jk}x_j + \sum_{i=1}^m a_{ik}y_i \geq 0 \quad (k=1, \dots, n) \quad (5.29)$$

$$\sum_{k=1}^n \left[c_k + \sum_{j=1}^n c_{jk}x_j + \sum_{i=1}^m a_{ik}y_i \right] x_k = 0 \quad (5.30)$$

$$\sum_{i=1}^l \left[\sum_{j=1}^n a_{ij}x_j - a_{i0} \right] y_i = 0 \quad (5.31)$$

$$y_1, \dots, y_l \geq 0; \quad y_{l+1}, \dots, y_m \text{ vrij} \quad (5.32).$$

Wij voegen aan (5.27) de verschilvariabelen $x_{n+i} (\geq 0)$ toe en aan (5.29) de verschilvariabelen $v_k (\geq 0)$. De voorwaarden (5.27), (5.29), (5.30) en (5.31) worden dan

$$\sum_{j=1}^n a_{ij}x_j + x_{n+i} = a_{i0} \quad (i=1, \dots, l) \quad (5.33)$$

$$c_k + \sum_{j=1}^n c_{jk}x_j + \sum_{i=1}^m a_{ik}y_i - v_k = 0 \quad (k=1, \dots, n) \quad (5.34)$$

$$\sum_{k=1}^n v_k x_k = 0 \quad (5.35)$$

$$\sum_{i=1}^l x_{n+i} y_i = 0 \quad (5.36).$$

Om een volledige eenheidsmatrix te krijgen behoeven nu slechts $m-1$

μ variabelen aan (5.28) toegevoegd te worden en n variabelen z_{1j} aan (5.34). In de tweede fase wordt de keuzeregel gewijzigd in: wanneer x_k in de basis zit, kan v_k er niet in, en omgekeerd voor $k=1, \dots, n$; wanneer x_{n+i} in de basis zit, kan y_i er niet in, en omgekeerd voor $i=1, \dots, l$. Bovendien moeten $y_1, \dots, y_l \geq 0$ gekozen worden.

Voorbeeld 5.2

Wij passen de geschetste methode toe op voorbeeld 5.1. De voorwaarden, waaraan x_1, x_2 en y moeten voldoen, zijn

$$\begin{aligned} x_1 + x_2 &\leq 2 \\ 2x_1 + y &\geq 4 \\ 2x_2 + y &\leq 2 \\ x_1, x_2, y &\geq 0 \end{aligned} \tag{5.37}.$$

Toevoegen van de verschilvariabelen x_3, v_1, v_2 en de variabelen $z_{11}, z_{12}, z_{21}, z_{22}$ leidt tot

$$\begin{aligned} x_1 + x_2 + x_3 &= 2 \\ 2x_1 + y - v_1 + z_{11} - z_{21} &= 4 \\ 2x_2 + y - v_2 + z_{12} - z_{22} &= 2 \end{aligned} \tag{5.38},$$

terwijl bovendien moet gelden

$$v_1 x_1 + v_2 x_2 = 0, \quad y x_3 = 0 \tag{5.39}.$$

Fase I kan worden overgeslagen, omdat er geen μ variabelen zijn en fase II begint met de basisoplossing $x_3=2$, $z_{11}=4$ en $z_{12}=2$. De van belang zijnde vectoren zijn opgenomen in tabel 5.VII, evenals de $z_j - c_j$ waarden. De nummering van de vectoren is dezelfde als in voorbeeld 5.1. Van het begin af aan zijn er 3 vectoren minder (A_8, A_{11}, A_{14}) en één vergelijking. De y -variabele kan hier pas in de laatste stap in de basis komen, aangezien x_3 er steeds in zit.

Tabel 5.VII

De vectoren van voorbeeld 5.2

c_j							1	1
	A_1	A_2	A_3	A_4	A_5	A_6	A_9	A_{10}
	1	1	1					
	2			1	-1		1	
		2		1		-1		1
$z_j - c_j$	2	2		2	-1	-1		
		2		1		-1	-1	
			-2	2	-1	-1		

Tabel 5.VIII

De iteratiestappen in voorbeeld 5.2

C_o	B	X_o	B^{-1}	X_1
	x_3	2	1	1
1	z_{11}	4	1	2
1	z_{12}	2	1	
$C_o B^{-1}$			0 1 1	

C_o	B	X_o	B^{-1}	X_2
	x_3	0	1 $-\frac{1}{2}$	1
	x_1	2	$\frac{1}{2}$	0
1	z_{12}	2	1	2
$C_o B^{-1}$			0 0 1	

C_0	B	X_0	B^{-1}			X_4
	x_2	0	1	$-\frac{1}{2}$		$-\frac{1}{2}$
	x_1	2		$\frac{1}{2}$		$\frac{1}{2}$
→ 1	z_{12}	2	-2	1	1	2
$C_0 B^{-1}$			-2	1	1	

C_0	B	X_0	B^{-1}			
	x_2	$\frac{1}{2}$	$\frac{1}{2}$	$-\frac{1}{4}$	$\frac{1}{4}$	
	x_1	$\frac{3}{2}$	$\frac{1}{2}$	$\frac{1}{4}$	$-\frac{1}{4}$	
	y	1	-1	$\frac{1}{2}$	$\frac{1}{2}$	
$C_0 B^{-1}$						

Opmerking 5.4

Wil men alle variabelen op dezelfde wijze behandelen, dus alle met de voorwaarde, dat ze ≥ 0 moeten zijn, dan kan men de y's, welke behoren bij voorwaarden in de vorm van vergelijkingen splitsen in twee variabelen en schrijven

$$y = y_1 - y_2 \quad (5.40),$$

waarin zowel y_1 als $y_2 \geq 0$ dienen te zijn.

§6 Het algoritme van BEALE

Naast het algoritme van WOLFE om kwadratische programme-ringsproblemen op te lossen, bestaan er nog verschillende an-dere algoritmen. Eén van de oudste is dat van BEALE, gepubli-ceerd in 1955 ¹⁾. Ook hier wordt uitgegaan van convexe, diffe-rentiëerbare functies.

Wij zullen de gedachtengang eerst illustreren aan de hand van een functie van 2 variabelen.

$$y = f(x_1, x_2) . \quad (6.1)$$

Wanneer $\frac{\partial f}{\partial x_1}$ 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 $X=(0\ 0)$ ²⁾. Zijn echter beide, of één der partiële afgeleiden niet nul, dan bestaan er lagere waarden van $f(x_1, x_2)$ dan $f(0,0)$. Wij kunnen deze bijvoorbeeld vinden door een variatie vanuit nul van een x_i , waarvoor geldt $\frac{\partial f}{\partial x_i} \neq 0$; hierbij laten wij de andere variabele gelijk aan nul.

In de in fig. 6.1 getekende situatie kan f verminderd worden door x_2 een positieve variatie te geven en x_1 nul te laten. De methode waarbij eerst nagegaan wordt of de waarde van de functie vanuit $X=(0\ 0)$ verminderd kan worden en zo ja, een vermindering wordt aangebracht door slechts één variabele te veranderen, werken wij nu uit voor een aantal numerieke voorbeelden.

1) E.M.L. BEALE, On Minimizing a Convex Function subject to Lineair Inequalities, Journal of the Royal Statistical Society, Series B, 17 (1955), 173-184.

2) Dat bij het trekken van deze conclusie van de convexiteit van de functie gebruik moet worden gemaakt, blijkt uit het voorbeeld, genoemd op blz. 42 van hoofdstuk II.

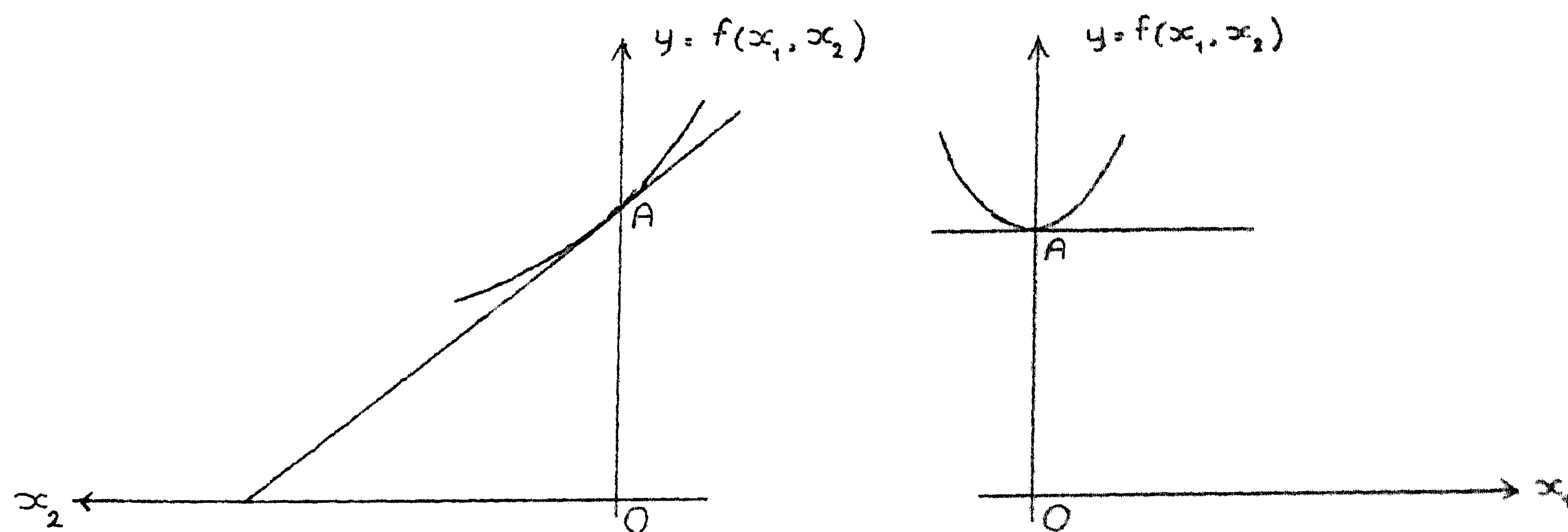
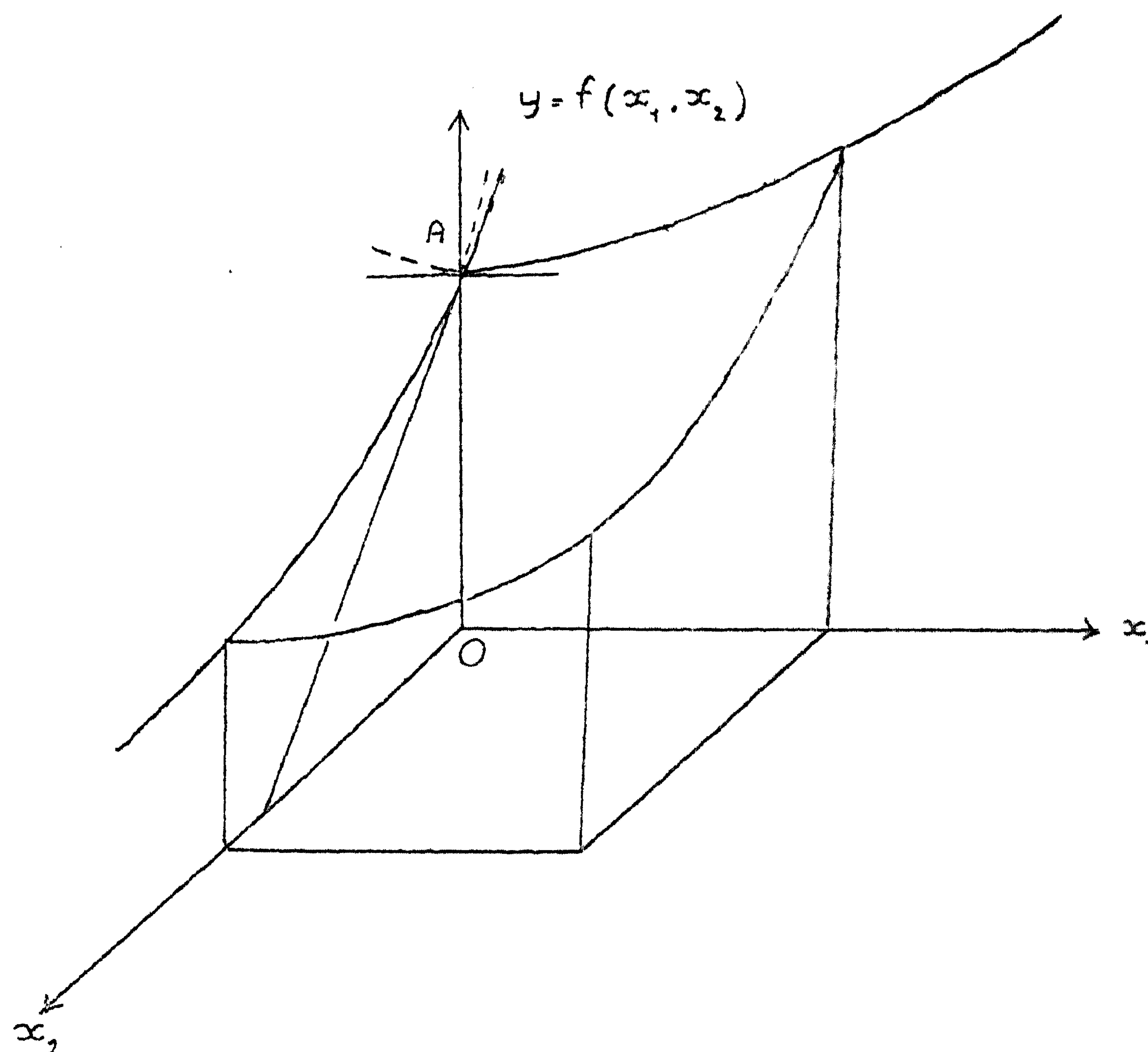


fig. 6.1

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

Voorbeeld 6.1

Bepaal het minimum van

$$y = f(x_1, x_2) = (x_1 - 16)^2 + (x_2 + 4)^2. \quad (6.2)$$

Met behulp van de in hoofdstuk II, § 6 aangegeven methode is

dit minimum direct te bepalen; wij vinden voor $x_1=16$ en $x_2=-4$ de waarde nul. De berekening kan echter ook geschieden met bovengenoemd principe.

Wij bepalen eerst de partiële afgeleiden in $X=(0\ 0)$.

Uit

$$\frac{1}{2} \frac{\partial f}{\partial x_1} = x_1 - 16 \quad (6.3)$$

en

$$\frac{1}{2} \frac{\partial f}{\partial x_2} = x_2 + 4 \quad (6.4)$$

volgt

$$\frac{1}{2} \left(\frac{\partial f}{\partial x_1} \right)_{x_1=x_2=0} = -16 \text{ en } \frac{1}{2} \left(\frac{\partial f}{\partial x_2} \right)_{x_1=x_2=0} = 4. \quad (6.5)$$

f kan dus verminderd worden door een positieve variatie van x_1 , waarbij x_2 nul gehouden wordt en door een negatieve variatie van x_2 , waarbij x_1 nul gehouden wordt. Doen wij het eerste, dan kan x_1 groeien tot 16, alvorens $f(x_1, x_2)$ weer toe gaat nemen. Wij definiëren nu een nieuwe variabele u_1 door

$$u_1 = \frac{1}{2} \frac{\partial f}{\partial x_1} = x_1 - 16 \quad (6.6)$$

en elimineren x_1 uit f met behulp van (6.6). De nieuwe functie

$$f_1(u_1, x_2) = u_1^2 + (x_2 + 4)^2 \quad (6.7)$$

neemt $u_1=0$, $x_2=0$ dezelfde waarde aan als de oude voor $x_1=16$, $x_2=0$, te weten 16. De convexiteit van de functie blijft bij de uitgevoerde transformatie behouden.

Op de nieuwe functie $f_1(u_1, x_2)$ wordt nu hetzelfde procédé toegepast. De afgeleiden naar u_1 resp. x_2 zijn

$$\frac{1}{2} \frac{\partial f_1}{\partial u_1} = u_1 \quad (6.8)$$

en

$$\frac{1}{2} \frac{\partial f_1}{\partial x_2} = x_2 + 4 \quad (6.9)$$

en bezitten in $u_1=0$, $x_2=0$ de waarden

$$\frac{1}{2} \left(\frac{\partial f_1}{\partial u_1} \right)_{u_1=x_2=0} = 0 \text{ en } \frac{1}{2} \left(\frac{\partial f_1}{\partial x_2} \right)_{u_1=x_2=0} = 4. \quad (6.10)$$

De functie f_1 kan dus verminderd worden door een negatieve variatie van x_2 waarbij u_1 nul gehouden wordt. De partiële afgeleide (6.9) wordt nul voor $x_2 = -4$; wij verminderen x_2 daarom tot -4. Door middel van de nieuwe variabele

$$u_2 = \frac{1}{2} \frac{\partial f_1}{\partial x_2} = x_2 + 4 \quad (6.11)$$

wordt x_2 geëlimineerd uit $f_1(x_1, x_2)$, hetgeen leidt tot

$$f_2(u_1, u_2) = u_1^2 + u_2^2. \quad (6.12)$$

De waarde van f_2 in $u_1 = 0, u_2 = 0$ is weer dezelfde als die van $f_1(u_1, x_2)$ in $u_1 = 0, x_2 = -4$, namelijk nul.

Daar de functie $f_2(u_1, u_2)$ wederom convex is en de beide partiële afgeleiden in $u_1 = u_2 = 0$ nul zijn, kan f_2 niet verder verminderd worden en dus is

$$\min f(x_1, x_2) = \min f_1(u_1, x_2) = \min f_2(u_1, u_2) = 0, \quad (6.13)$$

welk minimum bereikt wordt voor $u_1 = u_2 = 0$ of $x_1 = +16$ en $x_2 = -4$.

Voorbeeld 6.2

Bepaal het minimum van (6.2) onder de voorwaarden

$$x_1 \geq 0, x_2 \geq 0. \quad (6.14)$$

De methode uit hoofdstuk II kan nu niet toegepast worden, daar de oplossing niet aan (6.14) voldoet.

De berekening met de methode van BEALE is tot en met (6.10) identiek aan die in voorbeeld 6.1. Uit (6.10) blijkt, dat f_1 alleen door een negatieve variatie van x_2 kan worden verminderd. Door voorwaarde (6.14) is dit echter niet toegestaan en dus is het minimum van f onder deze voorwaarde gelijk aan 16; het wordt bereikt voor $u_1 = 0, x_2 = 0$ of $x_1 = 16, x_2 = 0$.

Voorbeeld 6.3

Bepaal het minimum van (6.2) onder de voorwaarden (6.14) en en

$$x_1 + x_2 \leq 8. \quad (6.15)$$

De laatste voorwaarde kan ook herschreven worden als

$$x_3 = 8 - x_1 - x_2 \geq 0. \quad (6.16)$$

De multiplicatorenmethode van LAGRANGE, uitgaande van de functie

$$F(x_1, x_2, x_3, y) = (x_1 - 16)^2 + (x_2 + 4)^2 + y(x_1 + x_2 + x_3 - 8) \quad (6.17)$$

levert geen toegelaten oplossing.

Bij de methode van BEALE blijkt uit (6.5), dat f vermindert kan worden door een positieve variatie van x_1 vanuit nul, waarbij x_2 nul gelaten wordt. Laten wij x_1 toenemen, dan moeten wij er op letten, dat

$$u_1 = \frac{1}{2} \frac{\partial f}{\partial x_1} = x_1 - 16$$

niet positief wordt, en

$$x_3 = 8 - x_1$$

niet negatief. Wij zien, dat x_3 het eerst nul wordt, n.l. voor $x_1 = 8$. Daarom wordt x_1 uit (6.16) opgelost en geëlimineerd uit (6.2). De variabele u_1 kan verder weggelaten worden, omdat noch het teken, noch de grootte voor ons van belang zijn. Het probleem (6.2), (6.14), (6.15) wordt dus vervangen door het probleem:

$$\text{minimaliseer } f_1(x_2, x_3) = (x_2 + x_3 + 8)^2 + (x_2 + 4)^2 \quad (6.18)$$

onder de voorwaarden

$$x_1 = 8 - x_2 - x_3 \quad (6.19)$$

en

$$x_1 \geq 0, x_2 \geq 0, x_3 \geq 0. \quad (6.20)$$

De waarde van f_1 in $x_2=x_3=0$ is dezelfde als die van f in $x_1=8, x_2=0$. De afgeleiden van $f_1(x_2, x_3)$ naar x_2 resp. x_3 zijn

$$\frac{1}{2} \frac{\partial f_1}{\partial x_2} = 2x_2 + x_3 + 12 \quad (6.21)$$

en

$$\frac{1}{2} \frac{\partial f_1}{\partial x_3} = x_2 + x_3 + 8 \quad (6.22)$$

en nemen in $x_2=x_3=0$ de waarden

$$\frac{1}{2} \left(\frac{\partial f}{\partial x_2} \right)_{x_2=x_3=0} = 12 \text{ en } \frac{1}{2} \left(\frac{\partial f}{\partial x_3} \right)_{x_2=x_3=0} = 8 \quad (6.23)$$

aan. Beide partiële afgeleiden zijn positief, zodat slechts door negatieve variaties vanuit $x_2=x_3=0$ vermindering van $f_1(x_2, x_3)$ is te bereiken. Daar deze door (6.20) niet zijn toegestaan is het minimum van (6.2) reeds gevonden. Het bedraagt dus 80 in het punt $x_2=x_3=0$, of $x_1=8$ en $x_2=0$.

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 6.1 - zie figuur 6.2 - wordt de x_1 eerst tot 16 gevarieerd, waardoor wij in het punt E komen, en daarna constant gehouden. Vervolgens wordt x_2 gevarieerd tot -4, waardoor het punt M bereikt wordt. In voorbeeld 6.2 wordt x_2 gevarieerd tot 16. Verdere vermindering van f is niet mogelijk, zoals meetkundig gemakkelijk is in te zien. In voorbeeld 6.3 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.

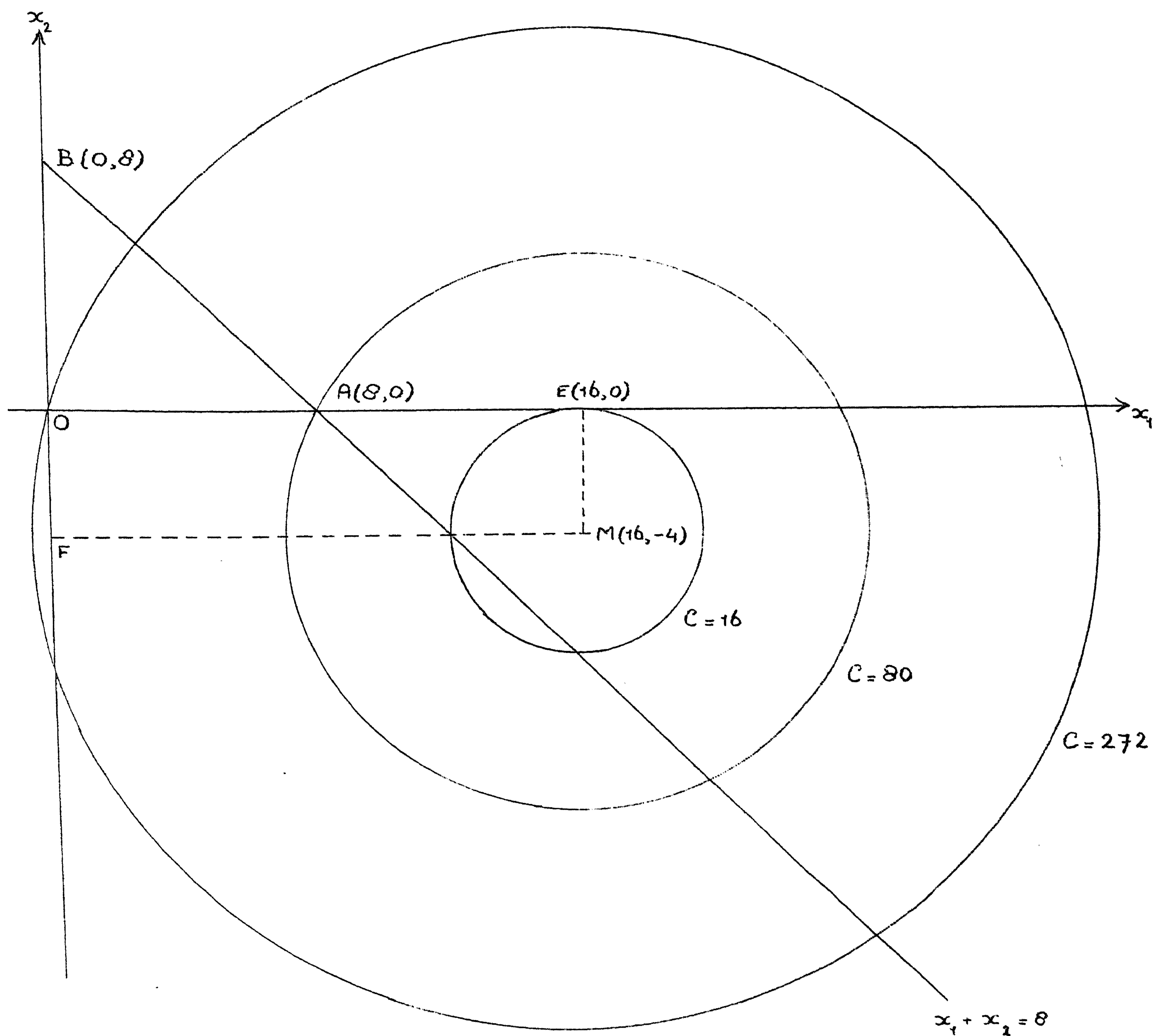


fig. 6.2

Voorbeeld 6.1.	toegelaten gebied:	gehele vlak		minimum $f=0$
"	6.2	"	"	kwadrant $x_1 \geq 0, x_2 \geq 0$; minimum $f=16$
"	6.3	"	"	OAB; minimum $f=80$

Opmerking 6.1

Wij herinneren er aan, dat het nul zijn van de partiële afgeleiden in $X = (0 \ 0)$ wel een noodzakelijke, maar geen voldoende voorwaarde is voor een minimum van een differentiëerbare functie in de oorsprong. Het is wel voldoende bij een convexe, differentiëerbare functie. Zijn echter de variabelen niet negatief, dan behoeven de partiële afgeleiden in een punt niet altijd nul te zijn, om daar een minimum van de functie te hebben; vgl. fig. 3.1 uit § 3.

Wij gaan nu over tot het algoritme van BEALE in de algemene vorm. Hiertoe wordt het kwadratische programmeringsprobleem als volgt opgeschreven.

$$\text{Minimaliseer } y=f(x)=\sum_{j=0}^n \sum_{k=0}^n \bar{c}_{jk} x_j x_k, \quad x_0=1, \quad \bar{c}_{jk}=\bar{c}_{kj} \quad (6.24)$$

$$\text{onder de voorwaarden } \sum_{j=1}^n \bar{a}_{ij} x_j = \bar{a}_{i0} \quad (6.25)$$

en

$$x_1, \dots, x_n \geq 0. \quad (6.26)$$

Vergeleken met de formulering in (5.1) is

$$\bar{c}_{j0} = \bar{c}_{0j} = \frac{1}{2} c_j \quad (j=1, \dots, n)$$

en

$$\bar{c}_{jk} = \frac{1}{2} c_{jk} \quad (j=1, \dots, n; k=1, \dots, n).$$

Het symbool x_0 wordt in f alleen ingevoerd om de notatie iets te vereenvoudigen. Wij onderstellen eenvoudigheidshalve, dat de $\bar{a}_{i0}$ alle ≥ 0 zijn en dat de matrix $\bar{A}$ van coëfficiënten $\bar{a}_{ij}$ een volledige eenheidsmatrix bevat, b.v. gevormd door de coëfficiënten van de variabelen x_{n-m+1} t/m x_n . Voor hetgeen men moet doen, wanneer dit niet het geval is, verwijzen wij naar de paragrafen 8 en 9 van hoofdstuk XII.

De voorwaarden (6.25) worden herschreven als

$$x_{n-m+i} = \bar{a}_{i0} - \sum_{j=1}^{n-m} \bar{a}_{ij} x_j \quad (i=1, \dots, m) \quad (6.27)$$

en vervolgens gebruikt om de variabelen x_{n-m+i} te elimineren uit (6.24). Vervangen wij nu

$$x_0, x_1, \dots, x_{n-m} \text{ door } z_0, z_1, \dots, z_{n-m}$$

$$x_{n-m+1}, \dots, x_n \quad " \quad x_1, \dots, x_m$$

$$\bar{a}_{i0}, -\bar{a}_{i1}, \dots, -\bar{a}_{i, n-m} \text{ door } a_{i0}, a_{i1}, \dots, a_{i, n-m} \quad (i=1, \dots, m),$$

dan gaat het probleem (6.24) t/m (6.26) over in:

$$\text{minimaliseer } y=f_1(z_1, \dots, z_{n-m}) = \sum_{j=0}^{n-m} \sum_{k=0}^{n-m} c_{jk} z_j z_k \quad (6.28)$$

onder de voorwaarden

$$x_i = a_{i0} + \sum_{j=1}^{n-m} a_{ij} z_j \quad (i=1, \dots, m) \quad (6.29)$$

en

$$x_i \geq 0 \quad (i=1, \dots, m), \quad z_j \geq 0 \quad (j=1, \dots, n-m). \quad (6.30)$$

Door de substitutie met behulp van (6.27) uit te voeren, kan men eenvoudig verifiëren, dat ook voor de c_{jk} geldt

$$c_{jk} = c_{kj} \quad (j, k = 0, \dots, n-m). \quad (6.31)$$

De hier toegepaste transformatie is dezelfde als die, welke in voorbeeld 6.3 werd uitgevoerd.

Aan de voorwaarden (6.29), (6.30) wordt voldaan door

$$x_i = a_{i0} \quad (i=1, \dots, m), \quad z_j = 0 \quad (j=1, \dots, n-m). \quad (6.32)$$

Daar de bij de x_i behorende vectoren van coëfficiënten weer een basis van de ruimte vormen, wordt ook hier van een basisoplossing gesproken. In overeenstemming met de in hoofdstuk XII gebruikte terminologie, noemt men $x_1, \dots, x_m$ basisvariabelen en $z_1, \dots, z_{n-m}$ niet-basisvariabelen.

De bij (6.32) behorende waarde van f_1 is gelijk aan c_{00} . Het betreffende punt valt, evenals in de voorbeelden 6.1, 6.2 en 6.3 het geval was, samen met de oorsprong, wanneer wij de coördinaten

$z_1, \dots, z_{n-m}$ gebruiken. Uitgaande van de basisoplossing, zoeken wij een nieuwe basisoplossing, waarin de waarde van f_1 lager is. Na iedere stap wordt het probleem weer in de vorm (6.28) t/m (6.30) getransformeerd.

Om na te gaan of er een toegelaten punt is met een lagere waarde van f_1 dan c_{j0} , differentiëren wij f_1 partiël naar de variabelen z_j :

$$\frac{1}{2} \frac{\partial f_1}{\partial z_j} = c_{j0} + \sum_{k=1}^{n-m} c_{jk} z_k \quad (6.33)$$

De waarde van de partiële afgeleide naar z_j in

$$z_1 = z_2 = \dots = z_{n-m} = 0 \quad (6.34)$$

is $2c_{j0}$. Zijn alle $c_{j0} \geq 0$, dan is het minimum bereikt. Zijn echter één of meer van de $c_{j0} < 0$, dan neemt f_1 af, wanneer men één van de betreffende variabelen positief maakt en de overige z -variabelen nul houdt. Meetkundig gezien is $2c_{j0}$ de richtingscoëfficiënt van de raaklijn in het punt $z_j = 0$, aan de doorsnijdingskromme van het oppervlak $y = f_1(z_1, \dots, z_{n-m})$ met het vlak $z_j O y$; zie fig. 6.3. Aangezien $y = f_1(z_1, \dots, z_{n-m})$ in A steiler

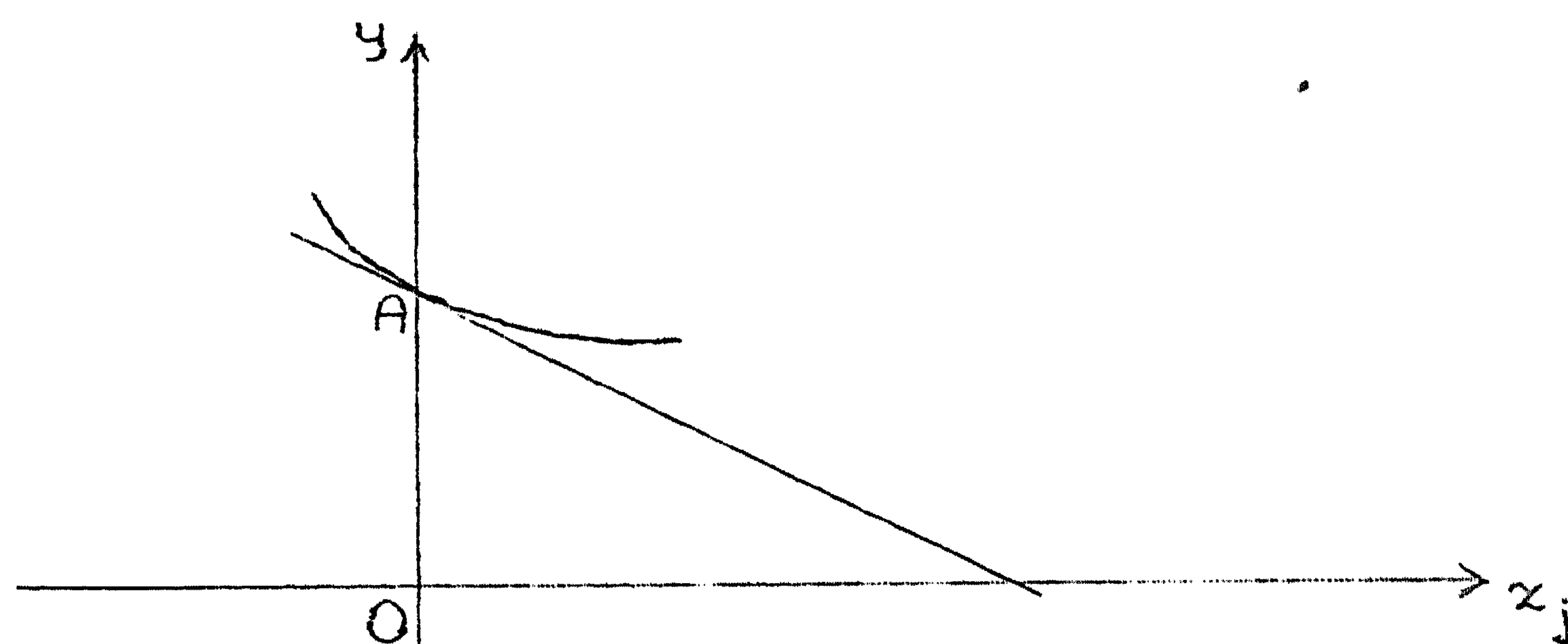


fig. 6.3

De doorsnijding van $y = f_1(z_1, \dots, z_{n-m})$ met het vlak $z_j O y$

loopt, naarmate c_{j0} kleiner is, ligt het voor de hand die variabele z_j positief te maken, waarvoor c_{j0} minimaal is, of $|c_{j0}|$ maximaal is. Wordt dit minimum voor verschillende waarden van j bereikt, dan kiezen wij hiervan degene met de kleinste index.

Stel dat z_1 positief gemaakt wordt. Bij variatie van z_1 vanuit 0 naar $+\infty$ kunnen zich drie verschillende gevallen

voordoen, in het vervolg aangegeven door A, B en C. Wij kunnen ze onderscheiden met de uit (6.29) en (6.33) volgende vergelijkingen

$$x_i = a_{i0} + a_{i1}z_1 \quad (i=1, \dots, m) \quad (6.35)$$

en

$$\frac{1}{2} \left(\frac{\partial f_1}{\partial z_1} \right)_{z_1=\dots=z_{l-1}=z_{l+1}=\dots=z_{n-m}=0} = c_{10} + c_{11}z_1, \quad (6.36)$$

welke de waarden van de x_i en de partiële afgeleide naar z_1 aangeven, wanneer z_1 verandert en $z_1, \dots, z_{l-1}, z_{l+1}, \dots, z_{n-m}$ nul blijven. In (6.35) is $a_{i0} \geq 0$ en a_{i1} willekeurig; in (6.36) is $c_{10} < 0$ en $c_{11} \geq 0$, omdat uit f_1 convex volgt, dat ook

$$c_{00} + 2c_{10}z_1 + c_{11}z_1^2 \quad (6.37)$$

convex moet zijn.

Geval A; $c_{11}=0$ en alle $a_{i1} \geq 0$

Wanneer z_1 toeneemt blijven alle $x_i \geq 0$ en behoudt de afgeleide het negatieve teken, zodat wij steeds in het toegelaten gebied blijven en f_1 steeds afneemt. Uit

$$f_1 = c_{00} + 2c_{10}z_1 \quad (6.38)$$

volgt dat het minimum van f_1 onder de voorwaarden (6.29) en (6.30) $-\infty$ is.

Bevinden wij ons niet in geval A, dan zal bij toenemende z_1 op een gegeven ogenblik, hetzij een x_i , hetzij de afgeleide gelijk aan nul worden. Bij verdere toeneming van z_1 , zouden wij in het eerste geval het toegelaten gebied verlaten, terwijl in het tweede geval f_1 weer groter zou worden.

Geval B; de basisvariabele x_r wordt nul, voordat de andere basisvariabelen en de partiële afgeleide naar z_1 nul worden.

Uit (6.35) en (6.36) volgt, dat in dit geval

$$\min_{a_{il} < 0} \left(-\frac{a_{io}}{a_{il}}, -\frac{c_{lo}}{c_{ll}} \right) = -\frac{a_{ro}}{a_{rl}} \quad (6.39)$$

is. Wij geven z_1 de waarde $-\frac{a_{ro}}{a_{rl}}$, voegen deze variabele aan de basisvariabelen toe en de variabele x_r , welke voor

$$z_1 = -\frac{a_{ro}}{a_{rl}} \quad (6.40)$$

nul wordt, aan de niet-basisvariabelen. Om het probleem in de vorm (6.28) t/m (6.29) terug te krijgen, lossen wij z_1 op uit de r^{de} vergelijking van (6.29):

$$z_1 = -\frac{a_{ro}}{a_{rl}} + \frac{1}{a_{rl}} x_r - \sum_{\substack{j=1 \\ j \neq l}}^{n-m} \frac{a_{rj}}{a_{rl}} z_j \quad (6.41)$$

en elimineren z_1 vervolgens uit de overige vergelijkingen van (6.29) en uit f_1 . De vergelijkingen (6.29) gaan, uitgezonderd de r^{de} , over in

$$x_i = a_{io} - \frac{a_{il}}{a_{rl}} \cdot a_{ro} + \frac{a_{il}}{a_{rl}} x_r + \sum_{\substack{j=1 \\ j \neq l}}^{n-m} \left(a_{ij} - \frac{a_{il}}{a_{rl}} a_{rj} \right) z_j \quad (i \neq r) \quad (6.42)$$

of

$$x_i = a'_{io} + a'_{il} x_r + \sum_{\substack{j=1 \\ j \neq l}}^{n-m} a'_{ij} z_j \quad (6.43)$$

met

$$a'_{ij} = a_{ij} - \frac{a_{il}}{a_{rl}} a_{rj} \quad \begin{matrix} (i=1, \dots, r-1, r+1, \dots, m) \\ (j=0, 1, \dots, l-1, l+1, \dots, n-m) \end{matrix} \quad (6.44)$$

en

$$a'_{il} = \frac{a_{il}}{a_{rl}} \quad (i=1, \dots, r-1, r+1, \dots, m). \quad (6.45)$$

De r^{de} vergelijking wordt

$$z_1 = a'_{ro} + a'_{rl} x_r + \sum_{\substack{j=1 \\ j \neq l}}^{n-m} a'_{rj} z_j \quad (6.46)$$

met

$$a'_{rj} = - \frac{a_{rj}}{a_{r1}} \quad (j=0,1,\dots,l-1,l+1,\dots,n-m) \quad (6.47)$$

en

$$a'_{r1} = \frac{1}{a_{r1}} \quad (6.48)$$

Definiëren wij nu

$$z'_j = z_j \quad (j=1,\dots,l-1,l+1,\dots,n-m), \quad z'_1 = x_r \quad (6.49)$$

en

$$x'_i = x_i \quad (i=1,\dots,r-1,r+1,\dots,m), \quad x'_r = z_1, \quad (6.50)$$

dan hebben de nieuwe bijvoorwaarden reeds de oude vorm (6.29) gekregen.

Teneinde eenvoudige transformatieformules voor f_1 te verkrijgen, schrijven wij (6.28) in de vorm

$$y = \sum_{j=0}^{n-m} \left(\sum_{k=0}^{n-m} c_{jk} z_k \right) z_j \quad (6.51)$$

en elimineren z_1 eerst uit de vorm tussen haken, hetgeen leidt tot

$$\sum_{j=0}^{n-m} \left[\sum_{\substack{k=0 \\ k \neq 1}}^{n-m} \left(c_{jk} - \frac{a_{rk}}{a_{r1}} c_{j1} \right) z_k + \frac{c_{jk}}{a_{r1}} x_r \right] z_j.$$

Vervolgens substitueren wij binnen de vierkante haken (6.49) en schrijven

$$\sum_{j=0}^{n-m} \left[\sum_{k=0}^{n-m} c'_{jk} z'_k \right] z_j \quad (6.52)$$

met

$$c'_{jk} = c_{jk} - \frac{a_{rk}}{a_{r1}} c_{j1} \quad (j=0,1,\dots,n-m) \quad (6.53)$$

$$(k=0,1,\dots,l-1,l+1,\dots,n-m)$$

en

$$c'_{j1} = \frac{c_{j1}}{a_{r1}} \quad (j=0,1,\dots,n-m). \quad (6.54)$$

De variabelen z_j worden nu binnen de haken gebracht en de variabelen z'_k er buiten:

$$\sum_{k=0}^{n-m} \left[\sum_{j=0}^{n-m} c'_{jk} z_j \right] z'_k \quad (6.55)$$

De matrix van coëfficiënten in (6.55) is de gespiegelde van die in (6.52). Daar uit (6.53) en (6.54) volgt, dat niet behoeft te gelden

$$c'_{jk} = c'_{kj} , \quad (6.56)$$

zijn deze matrices in het algemeen niet identiek. Door substitutie van (6.41) wordt z_1 weer geëlimineerd en gaat (6.55) over in

$$\sum_{k=0}^{n-m} \left[\sum_{\substack{j=0 \\ j \neq 1}}^{n-m} (c'_{jk} - \frac{a_{rj}}{a_{r1}} \cdot c'_{lk}) z_j + \frac{c'_{lk}}{a_{r1}} x_r \right] z'_k \quad (6.57)$$

Ook hierin vervangen wij de z_j en x_r door z'_j ; de coëfficiënten worden

$$c''_{jk} = c'_{jk} - \frac{a_{rj}}{a_{r1}} c'_{lk} \quad \begin{matrix} (j=0,1,\dots,l-1,l+1,\dots,n-m) \\ (k=0,1,\dots,n-m) \end{matrix} \quad (6.58)$$

en

$$c''_{lk} = \frac{c'_{lk}}{a_{r1}} \quad (k=0,1,\dots,n-m). \quad (6.59)$$

Vergelijkt men deze transformatieformules met (6.53) en (6.54), dan blijkt j de rol van de k -index te hebben overgenomen en k die van j . Als wij ze weer dezelfde rol laten vervullen, dan worden (6.58) en (6.59)

$$c''_{kj} = c'_{kj} - \frac{a_{rk}}{a_{r1}} c'_{lj} \quad \begin{matrix} (j=0,1,\dots,n-m) \\ (k=0,1,\dots,l-1,l+1,\dots,n-m) \end{matrix} \quad (6.60)$$

en

$$c''_{lj} = \frac{c'_{lj}}{a_{r1}} \quad (j=0,1,\dots,n-m). \quad (6.61)$$

Vergelijken wij deze formules met (6.53) en (6.54), dan zien wij, dat de c'' coëfficiënten op dezelfde wijze uit de c' coëfficiënten volgen als de c' coëfficiënten uit de c 's, mits men de c_{jk} vervangt door de c'_{kj} ; de a_{rk} coëfficiënten blijven hierbij dezelfde. Door de substitutie uit te voeren, kan men aantonen, dat voor de twee keer getransformeerde coëfficiënten geldt

$$c''_{jk} = c''_{kj} \quad (6.62)$$

Na deze transformaties bezit ook de objectfunctie de oude vorm (6.28). Wij kunnen dus op het nieuwe stelsel dezelfde iteratie toepassen als op (6.28) t/m (6.30) is gebeurd.

De coëfficiënten a_{ij} en c_{jk} kan men in tableau's opschrijven als volgt:

A						C					
	a_{10}	.	.	.	$a_{1,n-m}$		c_{00}	.	.	.	$c_{0,n-m}$
	.				.		.				.
	.				.		.				.
	.				.		.				.
	a_{m0}	.	.	.	$a_{m,n-m}$		$c_{n-m,0}$				$c_{n-m,n-m}$

Bij bovengenoemde transformaties wordt A dus één keer getransformeerd met de formules (6.44), (6.45) en (6.47), (6.48). Het C-tableau wordt getransformeerd met (6.53) en (6.54) en vervolgens gespiegeld ten opzichte van de hoofddiagonaal. Wij hernoemen de elementen door voor c'_{jk} te lezen c'_{kj} en transformeren daarna nog eens met (6.53) en (6.54), dezelfde a_{ij} en de hernoemde c'_{jk} gebruikend.

Geval C; de afgeleide naar z_1 wordt nul, voordat één der basisvariabelen x_i nul wordt.

Uit (6.35) en (6.36) volgt nu, dat

$$\min_{a_{i1} < 0} \left(-\frac{a_{i0}}{a_{i1}}, -\frac{c_{10}}{c_{11}} \right) = -\frac{c_{10}}{c_{11}} \quad (6.63)$$

is. Bij verdere toename van z_1 , zal f_1 niet verder afnemen, maar weer groter worden. Wij geven z_1 daarom de waarde $-\frac{c_{10}}{c_{11}}$

en voegen aan de voorwaarden (6.29) toe de voorwaarde

$$u_1 = c_{10} + \sum_{k=1}^{n-m} c_{1k} z_k \quad (6.64)$$

Het A tableau wordt dus met een rij uitgebreid en men leest de coëfficiënten c_{10} en c_{1k} als nieuwe a_{ij} coëfficiënten. De variabele z_1 lost men op uit (6.64) en vervolgens elimineert men z_1 uit alle vergelijkingen van (6.29) en uit (6.28). De hiervoor nodige transformatieformules zijn dezelfde als de onder B

gevonden formules.

Wij krijgen op deze wijze $m+1$ basisvariabelen, te weten $x_1, \dots, x_m$ en z_1 , terwijl het aantal niet-basisvariabelen gelijk aan $n-m$ is gebleven, n.l. $z_1, \dots, z_{l-1}, z_{l+1}, \dots, z_{n-m}$ en u_1 . Herdopen wij de variabelen met

$$z'_j = z_j \quad (j=1, \dots, l-1, l+1, \dots, n-m), z'_1 = u_1 \quad (6.65)$$

en

$$x'_i = x_i \quad (i=1, \dots, m), x'_{m+1} = z_1, \quad (6.66)$$

dan krijgt de nieuwe objectfunctie de vorm

$$f_2(z'_1, \dots, z'_{n-m}) = c'_{11}(z'_1)^2 + \sum_{j=0}^{n-m} \sum_{\substack{k=0 \\ j \neq 1, k \neq 1}}^{n-m} c'_{jk} z'_j z'_k \quad (6.67)$$

met

$$c'_{11} = \frac{1}{c_{11}}.$$

De coëfficiënten $c'_{j1} = c'_{1j}$ zijn dus alle 0 voor $j \neq 1$.

Het probleem heeft weer vrijwel dezelfde vorm als in (6.28) t/m (6.30). Het belangrijkste verschil is, dat de variabele z'_1 niet ≥ 0 behoeft te zijn, maar een zogenaamde vrije variabele is. Bovendien is het aantal bijvoorwaarden $m+1$ geworden.

Het zou voor de hand liggen om in de volgende stap eerst de vrije variabele z'_1 tot basisvariabele te maken. Immers, is de coëfficiënt $c'_{10} > 0$, dan krijgen wij lagere waarden van f_2 door z'_1 negatief te maken en is $c'_{10} < 0$, dan geschiedt dit door z'_1 een positieve waarde te geven. Aan (6.67) zien wij echter, dat $c'_{10} = 0$ is; verder is $c'_{11} > 0$, waardoor z'_1 ongelijk nul maken een verhoging van f_2 impliceert.

Wij bepalen derhalve op de gewone wijze de partiële afgeleiden en gaan na of er een negatieve afgeleide in het punt $z'_1 = \dots = z'_{n-m} = 0$ is. Zo nee, dan is het minimum bereikt. Zo ja, dan wordt op de reeds beschreven wijze een niet-basisvariabele gekozen, welke positief gemaakt wordt en wordt nagegaan in welk van de gevallen A, B of C men verkeert. Is dit geval A, dan is

het minimum weer $-\infty$. Verkeren wij in situatie B, dan wordt een x_i' variabele niet-basisvariabele en een z_j' ($j \neq 1$) variabele basisvariabele. Het aantal voorwaarden blijft $m+1$ en transformaties vinden plaats met behulp van (6.44), (6.45), (6.47), (6.48) en (6.53), (6.54). Hierna geldt voor de coëfficiënten van $z_1' z_j'$ in het algemeen niet meer, dat ze, afgezien van de coëfficiënt van $(z_1')^2$, alle nul zijn. In de volgende stap kan de functie dan wel kleiner gemaakt worden door z_1' hetzij positief, hetzij negatief te kiezen.

Zijn wij in geval C, dan wordt op de reeds beschreven wijze een nieuwe u-variabele en een nieuwe voorwaarde ingevoerd. Het aantal basisvariabelen wordt dus $m+2$; het aantal niet-basisvariabelen blijft $n-m$, waaronder 2 vrije variabelen. De objectfunctie krijgt de vorm

$$f_3(z_1''', \dots, z_{n-m}''') = \frac{1}{c_{11}'''} (z_1''')^2 + \frac{1}{c_{tt}'''} (z_t''')^2 + \sum_{j=0}^{n-m} \sum_{\substack{k=0 \\ j, k \neq 1, t}}^{n-m} c_{jk}''' z_j''' z_k''' \quad (6.68)$$

Men kan op deze manier doorgaan tot het minimum is bereikt. Hierbij kunnen vrije variabelen pas basisvariabelen gemaakt worden, nadat een niet negatieve variabele aan de z-variabelen is toegevoegd.

Opmerking 6.2

Het aantal basisvariabelen is minimaal m en maximaal n . Zodra een vrije variabele uit de niet basisvariabelen verwijderd wordt, kan deze en de bijbehorende vergelijking zonder meer geschrapt worden. Het heeft n.l. geen zin deze verder mee te transformeren, daar zowel het teken als de waarde van een u-variabele ons interesseert.

Opmerking 6.3

Wanneer een vrije variabele z_1 in de objectfunctie voorkomt met coëfficiënt $c_{10} > 0$, dan wordt z_1 negatief gemaakt en moet het onder A, B en C vermelde als volgt worden aangevuld.

A. Is $c_{11} = 0$ en zijn alle $a_{11} \leq 0$, dan kan de functie wille-

keurig klein gemaakt worden door z_1 voldoende klein te kiezen.

B. Wordt de basisvariabele $x_r = 0$, voordat de andere basisvariabelen of u_1 van teken veranderen, dan volgt uit (6.35) en (6.36), dat moet gelden

$$\min_{a_{il} > 0} \left(-\frac{a_{io}}{a_{il}}, -\frac{c_{lo}}{c_{ll}} \right) = -\frac{a_{ro}}{a_{rl}}, \quad (6.69)$$

aangezien een basisvariabele bij negatieve variatie van z_1 alleen negatief kan worden als $a_{il} > 0$ is. De transformatieformules blijven alle dezelfde.

C. Wordt u_1 nul, voordat één van de basisvariabelen van teken verandert, dan is

$$\min_{a_{il} > 0} \left(-\frac{a_{io}}{a_{il}}, -\frac{c_{lo}}{c_{ll}} \right) = -\frac{c_{lo}}{c_{ll}}. \quad (6.70)$$

Ook hier ondergaan de transformaties geen wijzigingen.

Opmerking 6.4

Men zegt dat de functie in standaardvorm (s.v.) is, wanneer $c_{jo} = 0$ is voor alle vrije variabelen z_j . De functie kan dus alleen zijn minimum aannemen in een s.v. Tijdens de transformaties trachten wij steeds f zo spoedig mogelijk in een s.v. te transformeren.

Opmerking 6.5

Het minimum van de functie wordt in een eindig aantal stappen bereikt. Men kan dit bewijzen door aan te tonen, dat a) het aantal verschillende standaardvormen eindig is, b) na een aantal iteraties nooit een s.v. bereikt wordt, die al eerder is aangenomen en c) met behulp van het algoritme de functie steeds in een eindig aantal stappen in een s.v. geschreven kan worden. Voor een volledig bewijs verwijzen wij naar de literatuur.³⁾

³⁾ Zie b.v. het artikel, genoemd in voetnoot 1).

Opmerking 6.6

In het begin van deze paragraaf is ondersteld, dat f een convexe functie is. Het gevonden lokale minimum is dus een absoluut minimum. Is de functie niet convex, dan heeft het beschreven algoritme niet tot het absolute minimum te leiden, zelfs niet tot een lokaal minimum. Het algoritme kan echter met enige regels aangevuld worden, waardoor hetzij een lokaal minimum bereikt wordt, hetzij een punt, dat een lokaal minimum is van een functie, die uit f ontstaat door de coëfficiënten van f willekeurig weinig te veranderen. ⁴⁾

Voorbeeld 6.5

Wij passen het algoritme toe op het probleem uit voorbeeld 5.1:

$$\text{minimaliseer } f(x_1, x_2) = -4x_1 - 2x_2 + x_1^2 + x_2^2 \quad (6.71)$$

onder de voorwaarden

$$x_1 + x_2 + x_3 = 2 \quad (6.72)$$

en

$$x_1, x_2, x_3 \geq 0. \quad (6.73)$$

Omdat x_3 een verschilvariabele is, die niet in f voorkomt, staat het probleem reeds in de vorm (6.28) t/m (6.30), wanneer wij schrijven:

$$\text{minimaliseer } y = f(z_1, z_2) = -4z_0 z_1 - 2z_0 z_2 + z_1^2 + z_2^2 \quad (6.74)$$

onder de voorwaarden

$$x_3 = 2 - z_1 - z_2 \quad (6.75)$$

en

$$x_3 \geq 0, z_1 \geq 0, z_2 \geq 0. \quad (6.76)$$

⁴⁾ Dit is uitgewerkt in o.a. E.M.L. BEALE, On Quadratic Programming, Naval Research Logistics Quarterly, 6, (1959), 227-243.

De A en C tableau's luiden respectievelijk

$$\begin{array}{c|cc} A & z_1 & z_2 \\ \hline x_3 & 2 & -1 & -1 \end{array} \quad \text{en} \quad \begin{array}{c|ccc} C & z_0 & z_1 & z_2 \\ \hline z_0 & 0 & -2 & -1 \\ z_1 & -2 & 1 & 0 \\ z_2 & -1 & 0 & 1 \end{array} \quad (6.77)$$

Aangezien het aantal variabelen erg klein is, worden alle formules volledig uitgeschreven en is het dus niet nodig de variabelen steeds te henummeren.

De eerste stap

Wij gaan uit van de tableau's in (6.77). Er zijn negatieve c_{j0} , n.l. c_{10} en c_{20} en

$$\max_{c_{j0} < 0} |c_{j0}| = |c_{10}| = 2. \quad (6.78)$$

Wij nemen z_1 dus in de basis op: $l=1$. Om na te gaan of een basisvariabele, dan wel de afgeleide het eerst nul wordt, bepalen wij

$$\min_{a_{i1} < 0} \left(-\frac{a_{i0}}{a_{i1}}, -\frac{c_{10}}{c_{11}} \right) = \min \left(-\frac{a_{10}}{a_{11}}, -\frac{c_{10}}{c_{11}} \right) = \min(2, 2) = 2. \quad (6.79)$$

Om geen vrije variabele aan de z -variabelen toe te behoeven voegen, kiezen wij $-\frac{a_{10}}{a_{11}}$ als quotiënt, waarvoor het minimum bereikt wordt, dus $r=1$. Het A tableau wordt vervolgens met (6.47) en (6.48) getransformeerd, waardoor het overgaat in

$$\begin{array}{c|cc} A_1 & z_3 & z_2 \\ \hline x_1 & 2 & -1 & -1 \end{array} \quad (6.80)$$

Het C-tableau wordt met (6.53) en (6.54) getransformeerd; $a_{rk}=a_{1k}$ en $a_{r1}=a_{11}$. Wij vinden

$$\begin{array}{c|ccc}
 C_1' & z_0 & z_3 & z_2 \\
 \hline
 z_0 & -4 & 2 & 1 \\
 z_1 & 0 & -1 & -1 \\
 z_2 & -1 & 0 & 1
 \end{array} , \quad (6.81)$$

spiegelen C_1' om de hoofddiagonaal en passen (6.53) en (6.54) nog eens toe:

$$\begin{array}{c|ccc}
 C_1'' & z_0 & z_3 & z_2 \\
 \hline
 z_0 & -4 & 0 & -1 \\
 z_3 & 0 & 1 & 1 \\
 z_2 & -1 & 1 & 2
 \end{array} . \quad (6.82)$$

In overeenstemming met hetgeen reeds eerder is opgemerkt, blijkt C_1' niet symmetrisch te zijn, terwijl C_1'' dit wel is.

De tweede stap

Wij gaan uit van de tableau's (6.80) en (6.82). Er is één negatieve c_{j0} , n.l. c_{20} en dus wordt z_2 in de basis opgenomen: $l=2$. Verder is

$$\min_{a_{i2} < 0} \left(-\frac{a_{i0}}{a_{i2}}, -\frac{c_{20}}{c_{22}} \right) = \min \left(-\frac{a_{10}}{a_{12}}, -\frac{c_{20}}{c_{22}} \right) = \min \left(2, \frac{1}{2} \right) = \frac{1}{2} , \quad (6.83)$$

waaruit volgt, dat de afgeleide het eerst 0 is. Wij voeren daarom de vrije variabele u_2 in, als volgt

$$u_2 = -1 + z_3 + 2z_2 \quad (6.84)$$

en voegen aan (6.80) een rij toe:

$$\begin{array}{c|ccc}
 A_1' & z_3 & z_2 & \\
 \hline
 x_1 & 2 & -1 & -1 \\
 u_2 & -1 & 1 & 2
 \end{array} . \quad (6.85)$$

De waarde van r is 2 en met (6.44), (6.45), (6.47), (6.48) wordt A_1' getransformeerd tot

$$\begin{array}{c|ccc}
 A_2 & & z_3 & u_2 \\
 \hline
 x_1 & \frac{3}{2} & -\frac{1}{2} & -\frac{1}{2} \\
 x_2 & \frac{1}{2} & -\frac{1}{2} & \frac{1}{2}
 \end{array} \quad (6.86)$$

Het C-tableau wordt met (6.53), (6.54) getransformeerd tot $(a_{r1}=a_{22}=2, a_{rk}=a_{2k})$

$$\begin{array}{c|ccc}
 C_2' & z_0 & z_3 & u_2 \\
 \hline
 z_0 & -\frac{9}{2} & \frac{1}{2} & -\frac{1}{2} \\
 z_3 & \frac{1}{2} & \frac{1}{2} & \frac{1}{2} \\
 z_2 & 0 & 0 & 1
 \end{array} \quad (6.87)$$

gespiegeld en nog een getransformeerd tot

$$\begin{array}{c|ccc}
 C_2'' & z_0 & z_3 & u_2 \\
 \hline
 z_0 & -\frac{9}{2} & \frac{1}{2} & 0 \\
 z_3 & \frac{1}{2} & \frac{1}{2} & 0 \\
 u_2 & 0 & 0 & \frac{1}{2}
 \end{array} \quad (6.88)$$

Tableau (6.88) staat in standaardvorm; verder is $c_{30} > \frac{1}{2}$. Het minimum is dus bereikt en bedraagt $c_{00} = -\frac{9}{2}$; het wordt aangenomen voor $x_1 = \frac{3}{2}$, $x_2 = \frac{1}{2}$, $z_3 = 0$. Dit antwoord komt overeen met het in voorbeeld 5.1 gevonden antwoord. In afwijking van hetgeen hier gebeurd is, is het gebruikelijk ook voor u-variabelen z's te schrijven en ze dan te nummeren naar de volgorde, waarin ze voorkomen; u_2 zou dus z_4 geworden zijn.

Opmerking 6.7

Ook deze berekening kan eenvoudig meetkundig worden geïnterpreteerd. De krommen

$$y = -4x_1 - 2x_2 + x_1^2 + x_2^2 = \text{constant} = C \quad (6.89)$$

stellen namelijk voor $C \geq -5$ cirkels voor met middelpunt $(2,1)$ en straal $\sqrt{C+5}$. Gevraagd wordt dus naar de cirkel met de kleinste straal, die nog een punt gemeen heeft met het toegelaten gebied OAB ; zie fig. 6.4.

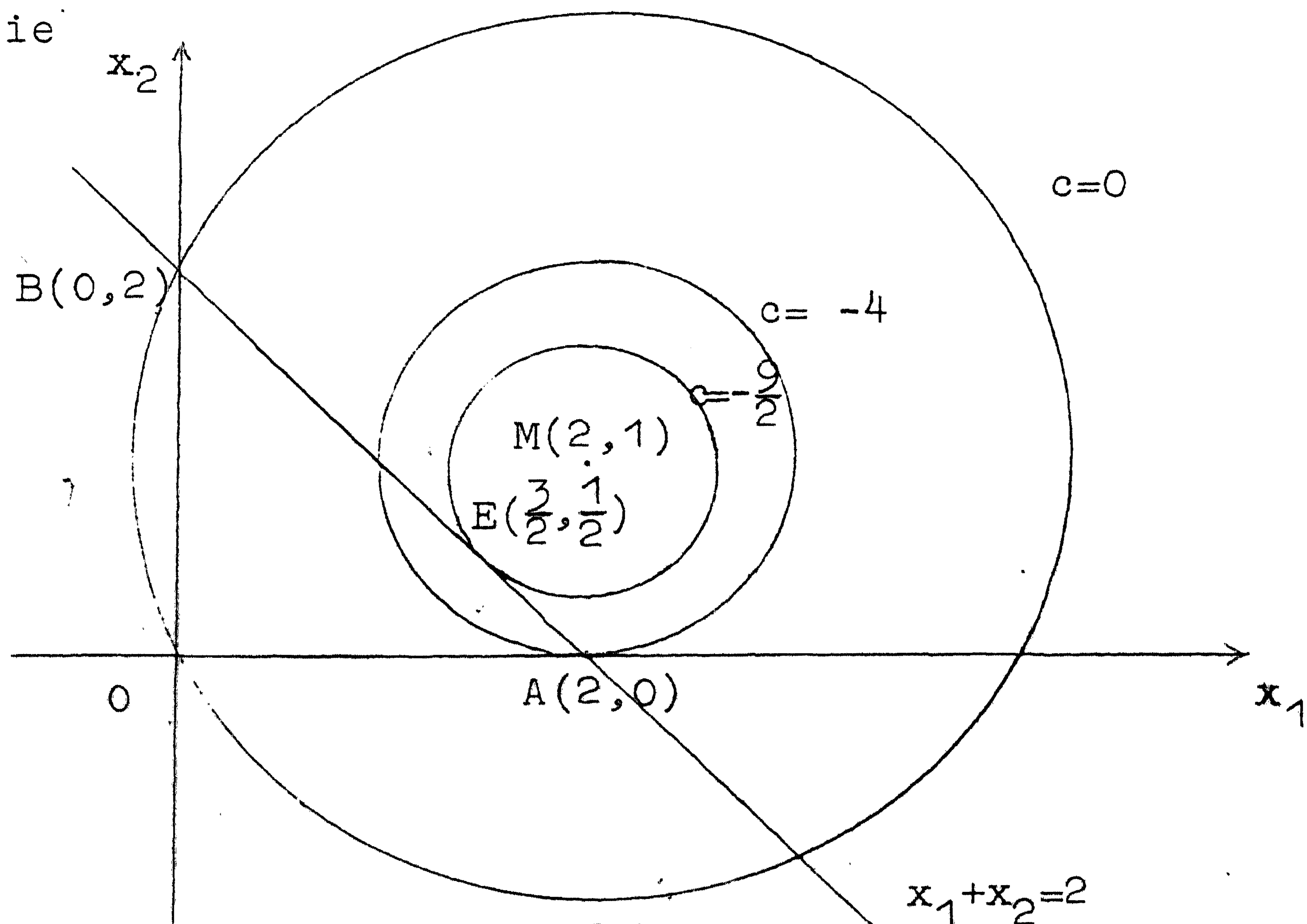


fig. 6.4

Toegelaten gebied $\triangle OAB$; kleinste afstand van M tot $\triangle OAB$ is ME

Bij de oplossing starten wij in $x_1=x_2=0$, dus O. In de eerste stap gaan wij naar $x_1=2$, $x_2=0$, dat is A. In de tweede stap wordt $x_3=0$ gehouden en lopen wij dus langs AB naar E, met coördinaten $x_1 = \frac{3}{2}$, $x_2 = \frac{1}{2}$.

Voor verdere literatuurverwijzingen en een overzicht van bestaande programma's om wiskundige programmeringsproblemen op te lossen, verwijzen wij naar een artikel van W.S. DORN. ⁵⁾

⁵⁾ W.S. DORN, Non-Linear Programming- A Survey, Management Science 9 (1963) 171-208.