

Albert, Max

Working Paper

Kuhn-Tucker conditions and linear homogeneity

Diskussionsbeiträge - Serie II, No. 226

Provided in Cooperation with:

Department of Economics, University of Konstanz

Suggested Citation: Albert, Max (1994) : Kuhn-Tucker conditions and linear homogeneity, Diskussionsbeiträge - Serie II, No. 226, Universität Konstanz, Sonderforschungsbereich 178 - Internationalisierung der Wirtschaft, Konstanz

This Version is available at:

<https://hdl.handle.net/10419/101611>

Standard-Nutzungsbedingungen:

Die Dokumente auf EconStor dürfen zu eigenen wissenschaftlichen Zwecken und zum Privatgebrauch gespeichert und kopiert werden.

Sie dürfen die Dokumente nicht für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, öffentlich zugänglich machen, vertreiben oder anderweitig nutzen.

Sofern die Verfasser die Dokumente unter Open-Content-Lizenzen (insbesondere CC-Lizenzen) zur Verfügung gestellt haben sollten, gelten abweichend von diesen Nutzungsbedingungen die in der dort genannten Lizenz gewährten Nutzungsrechte.

Terms of use:

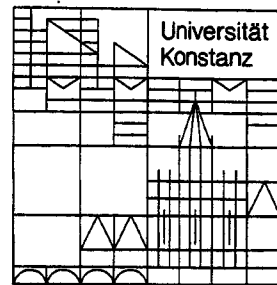
Documents in EconStor may be saved and copied for your personal and scholarly purposes.

You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.

**Sonderforschungsbereich 178
„Internationalisierung der Wirtschaft“**

Diskussionsbeiträge



**Juristische
Fakultät**

**Fakultät für Wirtschafts-
wissenschaften und Statistik**

Max Albert

**Kuhn-Tucker Conditions
and Linear Homogeneity**

Kuhn-Tucker Conditions and Linear Homogeneity

Max Albert*

Serie II - Nr. 226

Juni 1994

Correspondence:

Dr. Max Albert
Fakultät für Wirtschaftswissenschaften
und Statistik
Universität Konstanz
Postfach 55 60

D-78434 Konstanz, Germany

Tel.: 07531/88-2212

- * *I am grateful to Karl-Josef Koch, Nadine Leiner, Jürgen Meckl and Alan Woodland for helpful comments and discussions.*

Abstract

The paper proves a theorem on modified Kuhn–Tucker conditions for the case where a linearly homogenous function is part of the target function. These conditions are necessary and sufficient if target function and constraints satisfy certain requirements. Moreover, the envelope theorem can be applied as if the standard Kuhn–Tucker conditions were valid. An application to a problem in the theory of production illustrates the usefulness of the theorem.

1 Introduction

Optimization problems in economic theory often involve linearly homogenous functions. In general, the derivative of such a function is neither defined nor consistently definable at the origin. If the origin cannot be excluded as a solution, the Kuhn–Tucker (as well as the Lagrange) approach cannot be used to characterize solutions since this approach makes use of the derivative. The present paper derives *necessary and sufficient* first-order conditions (FOC) describing the optimum in these cases. The FOC are modifications of the well-known Kuhn–Tucker conditions.

The paper proceeds as follows. The next section states the problem. The third section formulates and proves the theorem and shows that the envelope theorem still applies for parametrized variants of the problem. The fourth section applies the theorem to the maximization problem defining the GNP function. Appendix A summarizes the Kuhn–Tucker approach as far as necessary to make the paper self-contained. Appendix B contains some short remarks on the differentiability problem.

2 The Problem

Consider a linearly homogenous function $h(\mathbf{x})$ and assume that it is differentiable everywhere except for the origin. Then it is easy to show that it is impossible to define $Dh(\mathbf{0})$ in a consistent way. Consider the behavior of h on two different rays through the origin, i.e. consider $h(t \cdot \mathbf{a}^*)$ and $h(t \cdot \bar{\mathbf{a}})$ where \mathbf{a}^* and $\bar{\mathbf{a}}$ are not collinear. Because of h 's linear homogeneity, we have $Dh(t \cdot \mathbf{a}) = Dh(\mathbf{a})$ if $t \neq 0$. In general, however, we have $Dh(\mathbf{a}^*) \neq Dh(\bar{\mathbf{a}})$. Thus there is no common limit to the differentials when one approaches the origin from different directions. Obviously, for a linearly homogenous function, $Dh(\mathbf{0})$ is defined only if h is linear.¹

¹ If h is concave, it is possible to replace differentiability by *subdifferentiability* (see appendix B). However, this is no real solution to the problem considered here. It is true that one can state the same Kuhn–Tucker conditions as before, using subdifferentiability instead of differentiability. However, *implicitly* one has to go through considerations analogous to those of the next section when interpreting these conditions.

A quite general standard formulation for the case where this makes trouble is

$$(1) \quad \max_{\mathbf{x}, \mathbf{y}} \{f(h(\mathbf{x}), \mathbf{x}, \mathbf{y}) : \mathbf{g}(\mathbf{x}, \mathbf{y}) \leq \mathbf{0}, \mathbf{x} \geq \mathbf{0}, \mathbf{y} \geq \mathbf{0}\},$$

where $f: \mathbb{R}_+ \times \mathbb{R}_+^n \times \mathbb{R}^m \mapsto \mathbb{R}$ is non-decreasing in h and where $h: \mathbb{R}_+^n \mapsto \mathbb{R}_+$ is linearly homogenous (but *not* linear) and non-decreasing in \mathbf{x} . Linear homogeneity of h implies that the Kuhn–Tucker approach is not applicable to problem (1) since at $\mathbf{x} = \mathbf{0}$ the FOC

$$(2) \quad \begin{aligned} (a) \quad & D_h f(h(\mathbf{x}), \mathbf{x}, \mathbf{y}) \cdot Dh(\mathbf{x}) + \\ & D_x f(h(\mathbf{x}), \mathbf{x}, \mathbf{y}) - \boldsymbol{\mu}^T D_x \mathbf{g}(\mathbf{x}, \mathbf{y}) \leq \mathbf{0}, \mathbf{x} \geq \mathbf{0} \quad w.c.s. \\ (b) \quad & D_y f(h(\mathbf{x}), \mathbf{x}, \mathbf{y}) - \boldsymbol{\mu}^T D_y \mathbf{g}(\mathbf{x}, \mathbf{y}) \leq \mathbf{0}, \mathbf{y} \geq \mathbf{0} \quad w.c.s. \\ (c) \quad & \mathbf{g}(\mathbf{x}, \mathbf{y}) \leq \mathbf{0}, \boldsymbol{\mu} \geq \mathbf{0} \quad w.c.s. \end{aligned}$$

are not defined. Therefore, if we cannot exclude $\mathbf{x} = \mathbf{0}$ as a possible solution, the Kuhn–Tucker conditions are neither necessary nor sufficient. In the following we derive an *alternative* set of modified Kuhn–Tucker conditions that are necessary and sufficient in a wide class of cases. These conditions are obtained by strengthening the standard conditions deriving from the following equivalent reformulation of problem (1):

$$(3) \quad \begin{aligned} \max_{t, \mathbf{a}, \mathbf{y}} \quad & \{f(t, t \cdot \mathbf{a}, \mathbf{y}) : \mathbf{g}(t \cdot \mathbf{a}, \mathbf{y}) \leq \mathbf{0}, \mathbf{a} \geq \mathbf{0}, \mathbf{y} \geq \mathbf{0}, \\ & t \geq 0, h(\mathbf{a}) \geq 1\} \end{aligned}$$

(3) rests on the idea to split up the original problem of choosing the optimal \mathbf{x} into two problems: the choice of a ray through the origin with direction \mathbf{a} and $h(\mathbf{a}) \geq 1$, and the choice of the distance t to the origin along that ray. This implies $\mathbf{x} = t \cdot \mathbf{a}$ since, as will be shown below, we have $h(\mathbf{a}) = 1$ at the maximum. (3) avoids the problem at the origin. However, the admissible set is not convex in (t, \mathbf{a}) despite its

being convex in t and \mathbf{a} separately. The standard FOC

$$\begin{aligned}
& (a) \quad \lambda \cdot Dh(\mathbf{a}) + \\
& \quad D_x f(t, t \cdot \mathbf{a}, \mathbf{y}) \cdot t - \boldsymbol{\mu}^T D_x \mathbf{g}(t \cdot \mathbf{a}, \mathbf{y}) \cdot t \leq 0, \quad \mathbf{a} \geq \mathbf{0} \quad w.c.s. \\
& (b) \quad D_h f(t, t \cdot \mathbf{a}, \mathbf{y}) + \\
(4) \quad & \quad D_x f(t, t \cdot \mathbf{a}, \mathbf{y}) \mathbf{a} - \boldsymbol{\mu}^T D_x \mathbf{g}(t \cdot \mathbf{a}, \mathbf{y}) \mathbf{a} \leq 0, \quad t \geq 0 \quad w.c.s. \\
& (c) \quad D_y f(t, t \cdot \mathbf{a}, \mathbf{y}) - \boldsymbol{\mu}^T D_y \mathbf{g}(t \cdot \mathbf{a}, \mathbf{y}) \leq 0, \quad \mathbf{y} \geq \mathbf{0} \quad w.c.s. \\
& (d) \quad \mathbf{g}(t \cdot \mathbf{a}, \mathbf{y}) \leq 0, \quad \boldsymbol{\mu} \geq \mathbf{0} \quad w.c.s. \\
& (e) \quad h(\mathbf{a}) \geq 1, \quad \lambda \geq 0 \quad w.c.s.
\end{aligned}$$

are sometimes necessary but never sufficient. In the next section, we derive necessary and sufficient conditions for problem (1). Moreover, we show that these conditions imply (4), and that problems (1) and (3) are indeed equivalent.

All proofs in the next section can be adapted without difficulty if the function h is replaced by a vector-valued linearly homogenous function $\mathbf{h}: \mathbb{R}_+^n \mapsto \mathbb{R}_+^s$.

3 Modified Kuhn–Tucker Conditions

We will assume that the function $h(\mathbf{x})$ is *regular* in the following sense.

Def. 1 *A nonlinear but linearly homogenous function $h: \mathbb{R}_+^n \mapsto \mathbb{R}_+$ is regular if and only if it is differentiable everywhere in $\mathbb{R}_+^n \setminus \{\mathbf{0}\}$, non-decreasing with $Dh(\mathbf{x}) \neq \mathbf{0}$ for $\mathbf{x} \neq \mathbf{0}$, and if the minimum problem*

$$(5) \quad z(\mathbf{q}) \stackrel{\text{def}}{=} \min_{\mathbf{x}} \{ \mathbf{q}^T \mathbf{x} : h(\mathbf{x}) \geq 1, \mathbf{x} \geq \mathbf{0} \}$$

has a solution for any $\mathbf{q} \in \mathbb{R}_+^n$.

Thus regularity requires that $z(\mathbf{q})$ is defined everywhere in \mathbb{R}_+^n . Beyond the differentiability requirement, this kind of regularity constrains the behavior of the function at the boundary of \mathbb{R}_+^n . According to theorem 8, the following FOC are *necessary* conditions for (5):

$$\begin{aligned}
(6) \quad & \mathbf{q}^T - \lambda \cdot Dh(\mathbf{x}) \geq \mathbf{0}, \quad \mathbf{x} \geq \mathbf{0} \quad w.c.s. \\
& h(\mathbf{x}) \geq 1, \quad \lambda \geq 0 \quad w.c.s.
\end{aligned}$$

Regularity guarantees that (6) has a solution for all non-negative \mathbf{q} . Note that for any solution $(\mathbf{x}^*, \lambda^*)$, we have $\lambda^* = (\mathbf{q}^*)^T \mathbf{x}^*$. Moreover we always can choose \mathbf{x}^* such that $h(\mathbf{x}^*) = 1$.

If we require h to be regular *and* concave, then the above FOC are not only necessary, but also sufficient according to theorem 9. As the following theorems show, the problem is simplified in this case.

Theorem 1 *Let $z: \mathbb{R}_+^n \mapsto \mathbb{R}_+$ be defined by (5) where $h: \mathbb{R}_+^n \mapsto \mathbb{R}_+$ is regular and concave. Then the following equivalence holds:*

$$(7) \quad h(\mathbf{x}) \equiv \min_{\mathbf{q}} \{ \mathbf{q}^T \mathbf{x} : z(\mathbf{q}) \geq 1, \mathbf{q} \geq \mathbf{0} \}$$

This theorem is well-known (cf. Diewert (1974: 112) and the comment by Shepard (1974: 200)). The following proof makes no direct use of these references.

Proof. Consider the function

$$g(\mathbf{x}) \stackrel{\text{def}}{=} \min_{\mathbf{q}} \{ \mathbf{q}^T \mathbf{x} : z(\mathbf{q}) \geq 1, \mathbf{q} \geq \mathbf{0} \},$$

which in contrast to h is *defined* by the minimization problem. Let $\bar{\mathbf{q}}$ be a solution corresponding to $\bar{\mathbf{x}}$. Because of the linear homogeneity of h and z , we can restrict our considerations w.l.o.g. to the case where $h(\bar{\mathbf{x}}) = 1$ and $z(\bar{\mathbf{q}}) = 1$. The proof proceeds in two steps: (i) We show that $g(\bar{\mathbf{x}}) \leq 1$. (ii) We show that $g(\bar{\mathbf{x}}) < 1$ is impossible.

(i) Since the upper contour set $\{ \mathbf{x} \in \mathbb{R}_+^n : h(\mathbf{x}) \geq 1 \}$ is convex, there exists a \mathbf{q}^* such that $\bar{\mathbf{x}}$ is a solution to problem (5). Again linear homogeneity allows us to choose $z(\mathbf{q}^*) = 1$ w.l.o.g. Since $\bar{\mathbf{q}}$ minimizes $\mathbf{q}^T \bar{\mathbf{x}}$ given the constraint $z(\mathbf{q}^*) = 1$, $g(\bar{\mathbf{x}})$ cannot be greater than 1 since \mathbf{q}^* with $(\mathbf{q}^*)^T \bar{\mathbf{x}} = 1$ is admissible.

(ii) Let $g(\bar{\mathbf{x}}) < 1$. Then $\bar{\mathbf{q}}^T \bar{\mathbf{x}} < z(\bar{\mathbf{q}}) = 1$. This, however, is impossible since $h(\bar{\mathbf{x}}) = 1$ and since $z(\bar{\mathbf{q}})$ is the minimum of $\bar{\mathbf{q}}^T \mathbf{x}$ for all \mathbf{x} fulfilling $h(\mathbf{x}) \geq 1$.

Thus we have shown that $g(\bar{\mathbf{x}}) = h(\bar{\mathbf{x}}) = 1$. Because of linear homogeneity, this generalizes to all values of h . *QED*

Since any linearly homogenous function is concave if it is quasiconcave (Takayama 1985: 114), the following theorem cannot be strengthened by assuming h to be quasiconcave.

Theorem 2 *Let the functions f , g be differentiable in their domains. Let f be non-decreasing in h with $D_h f \neq \mathbf{0}$ everywhere. Let h be regular and concave. Let $z(\mathbf{q})$ be defined by (5). Let $\pi(\mathbf{q})$ be defined by*

$$(8) \quad \pi(\mathbf{q}) \stackrel{\text{def}}{=} \max_{\mathbf{x}, \mathbf{y}} \{f(\mathbf{q}^T \mathbf{x}, \mathbf{x}, \mathbf{y}) : g(\mathbf{x}, \mathbf{y}) \leq \mathbf{0}, \mathbf{x} \geq \mathbf{0}, \mathbf{y} \geq \mathbf{0}\}.$$

Then problem (1) is equivalent to

$$(9) \quad \min_{\mathbf{q}} \{\pi(\mathbf{q}) : z(\mathbf{q}) \geq 1, \mathbf{q} \geq \mathbf{0}\}.$$

We proof the theorem in two steps.

1. *Let \mathbf{q}^* be a solution of (9) and $(\mathbf{x}^*, \mathbf{y}^*)$ a corresponding solution of (8). Then $(\mathbf{x}^*, \mathbf{y}^*)$ is a solution of (1).*

Proof. Because of $D_h f > 0$, \mathbf{q}^* has to minimize $\mathbf{q}^T \mathbf{x}^*$ and thus has to be a solution to problem (7) corresponding to \mathbf{x}^* . This implies that $(\mathbf{q}^*)^T = Dh(\mathbf{x}^*)$.

Because of the concavity of h , we have $Dh(\mathbf{x}^*)\mathbf{x} \geq h(\mathbf{x})$. Therefore $\pi(\mathbf{q}^*) \geq f(h(\mathbf{x}), \mathbf{x}, \mathbf{y})$ for all admissible (\mathbf{x}, \mathbf{y}) . Thus there is no admissible $(\bar{\mathbf{x}}, \bar{\mathbf{y}})$ such that $f(h(\bar{\mathbf{x}}), \bar{\mathbf{x}}, \bar{\mathbf{y}}) > f(h(\mathbf{x}^*), \mathbf{x}^*, \mathbf{y}^*)$. *QED*

2. *Let $(\mathbf{x}^*, \mathbf{y}^*)$ be a solution of (1). Then $\mathbf{q}^* = [Dh(\mathbf{x}^*)]^T$ is a solution of (9) and $(\mathbf{x}^*, \mathbf{y}^*)$ is a corresponding solution of (8).*

Proof. Again we use the fact that a solution of (9) must be equal to $Dh(\mathbf{x})$ for some \mathbf{x} . Let $\mathbf{s}(\mathbf{x}) \stackrel{\text{def}}{=} (Dh(\mathbf{x}))^T$ and $\phi(\mathbf{x}) \stackrel{\text{def}}{=} \pi(\mathbf{s}(\mathbf{x}))$. By definition of the functions, the minimum of π is the minimum of $\phi(\mathbf{x})$, and the solution of (9) must be $\mathbf{q} = [Dh(\hat{\mathbf{x}})]^T$ where $\hat{\mathbf{x}}$ minimizes ϕ .

Because of the concavity of h , we have $Dh(\bar{\mathbf{x}})\mathbf{x} \geq h(\mathbf{x})$ for all $\bar{\mathbf{x}}, \mathbf{x}$. Therefore $\phi(\mathbf{x}) \geq \pi^*$ where $\pi^* \stackrel{\text{def}}{=} f(h(\mathbf{x}^*), \mathbf{x}^*, \mathbf{y}^*)$. Thus \mathbf{x}^* is a minimum of $\phi(\mathbf{x})$, and \mathbf{q}^* is a solution of (9). Moreover, since $\min_{\mathbf{x}} \phi(\mathbf{x}) = \pi^*$, $(\mathbf{x}^*, \mathbf{y}^*)$ is a solution of (8) corresponding to \mathbf{q}^* . *QED*

The standard FOC of problems (8) and (9) are defined everywhere. If the admissible set of the original problem (1) is convex, the same goes for problem (8). Therefore theorem 2 solves our problem. But the combined FOC of problems (8) and (9) are quite messy. Moreover, it can be shown that these combined FOC are equivalent to those we will derive below without making use of the concavity of h . Nevertheless, theorem 2 may be useful in some contexts.

In the following, we make use of two further subproblems of (1). One is given by

$$(10) \quad \max_{\mathbf{x}, \mathbf{y}} \{f(h(\mathbf{x}), \mathbf{x}, \mathbf{y}) : \mathbf{g}(\mathbf{x}, \mathbf{y}) \leq \mathbf{0}, \mathbf{x} \geq \mathbf{0}, \mathbf{y} \geq \mathbf{0}, \boldsymbol{\tau}^T \mathbf{x} \geq m\}$$

with FOC

$$(11) \quad \begin{aligned} (a) \quad & D_h f(h(\mathbf{x}), \mathbf{x}, \mathbf{y}) \cdot D h(\mathbf{x}) + D_x f(h(\mathbf{x}), \mathbf{x}, \mathbf{y}) - \\ & \quad \boldsymbol{\mu}^T D_x \mathbf{g}(\mathbf{x}, \mathbf{y}) + \theta \cdot \boldsymbol{\tau}^T \leq \mathbf{0}, \mathbf{x} \geq \mathbf{0} \quad w.c.s. \\ (b) \quad & D_y f(h(\mathbf{x}), \mathbf{x}, \mathbf{y}) - \boldsymbol{\mu}^T D_y \mathbf{g}(\mathbf{x}, \mathbf{y}) \leq \mathbf{0}, \mathbf{y} \geq \mathbf{0} \quad w.c.s. \\ (c) \quad & \mathbf{g}(\mathbf{x}, \mathbf{y}) \leq \mathbf{0}, \boldsymbol{\mu} \geq \mathbf{0} \quad w.c.s. \\ (d) \quad & \boldsymbol{\tau}^T \mathbf{x} \geq m, \theta \geq 0 \quad w.c.s., \end{aligned}$$

where $\tau_j = 1$ for all j and $m > 0$. The additional constraint $\boldsymbol{\tau}^T \mathbf{x} \geq m > 0$ requires that at least one of the x_j is greater than zero. The second subproblem results from fixing a ray through the origin:

$$(12) \quad \max_{t, \mathbf{y}} \{f(t, t \cdot \mathbf{a}, \mathbf{y}) : \mathbf{g}(t \cdot \mathbf{a}, \mathbf{y}) \leq \mathbf{0}, t \geq 0, \mathbf{y} \geq \mathbf{0}\}$$

The FOC are given by

$$(13) \quad \begin{aligned} (a) \quad & D_h f(t, t \cdot \mathbf{a}, \mathbf{y}) + \\ & D_x f(t, t \cdot \mathbf{a}, \mathbf{y}) \mathbf{a} - \boldsymbol{\mu}^T D_x \mathbf{g}(t \cdot \mathbf{a}, \mathbf{y}) \mathbf{a} \leq 0, t \geq 0 \quad w.c.s. \\ (b) \quad & D_y f(t, t \cdot \mathbf{a}, \mathbf{y}) - \boldsymbol{\mu}^T D_y \mathbf{g}(t \cdot \mathbf{a}, \mathbf{y}) \leq 0, \mathbf{y} \geq \mathbf{0} \quad w.c.s. \\ (c) \quad & \mathbf{g}(t \cdot \mathbf{a}, \mathbf{y}) \leq \mathbf{0}, \boldsymbol{\mu} \geq \mathbf{0} \quad w.c.s. \end{aligned}$$

Theorem 3 *Let the functions f, g be differentiable everywhere in their domains. Let f be non-decreasing in h . Let h be regular. Consider the following FOC:*

$$\begin{aligned}
 (14) \quad & (a) \quad \lambda \cdot Dh(\mathbf{a}) + \\
 & \quad D_x f(t, t \cdot \mathbf{a}, \mathbf{y}) - \boldsymbol{\mu}^T D_x g(t \cdot \mathbf{a}, \mathbf{y}) \leq \mathbf{0}, \mathbf{a} \geq \mathbf{0} \quad w.c.s. \\
 & (b) \quad D_y f(t, t \cdot \mathbf{a}, \mathbf{y}) - \boldsymbol{\mu}^T D_y g(t \cdot \mathbf{a}, \mathbf{y}) \leq \mathbf{0}, \mathbf{y} \geq \mathbf{0} \quad w.c.s. \\
 & (c) \quad g(t \cdot \mathbf{a}, \mathbf{y}) \leq \mathbf{0}, \boldsymbol{\mu} \geq \mathbf{0} \quad w.c.s. \\
 & (d) \quad h(\mathbf{a}) \geq 1, \lambda \geq 0 \quad w.c.s. \\
 & (e) \quad D_h f(t, t \cdot \mathbf{a}, \mathbf{y}) \leq \lambda, t \geq 0 \quad w.c.s.
 \end{aligned}$$

- a) *Let the FOC (11) and (13) be necessary conditions for the subproblems (10) and (12), respectively. Then, if $(\mathbf{x}^*, \mathbf{y}^*)$ is a solution to (1), there exists a vector $(t^*, \mathbf{a}^*, \mathbf{y}^*, \boldsymbol{\mu}^*, \lambda^*)$ with $\mathbf{x}^* = t^* \cdot \mathbf{a}^*$ satisfying the FOC (14).*
- b) *Let the FOC (11) and (13) be sufficient conditions for the subproblems (10) and (12), respectively. Then, if there is a vector $(t^*, \mathbf{a}^*, \mathbf{y}^*, \boldsymbol{\mu}^*, \lambda^*)$ with $\mathbf{x}^* = t^* \cdot \mathbf{a}^*$ satisfying the FOC (14), $(\mathbf{x}^*, \mathbf{y}^*)$ is a solution to (1).*

We give separate proofs for several steps of the argument.

1. (2) and (11) are equivalent if $\boldsymbol{\tau}^T \mathbf{x} > m$.

Proof. In (11) $\boldsymbol{\tau}^T \mathbf{x} > m$ implies $\theta = 0$. The result follows immediately from a comparison of the two sets of inequalities. *QED*

2. If $(t, \mathbf{a}, \mathbf{y}, \boldsymbol{\mu}, \lambda)$ satisfies (14), then it satisfies (13).

Proof. From (14) we get

$$(15) \quad \lambda + D_x f(t, t \cdot \mathbf{a}, \mathbf{y}) \mathbf{a} - \boldsymbol{\mu}^T D_x g(t \cdot \mathbf{a}, \mathbf{y}) \mathbf{a} = 0.$$

Since $\lambda \geq D_h f(t, t \cdot \mathbf{a}, \mathbf{y}) \geq 0$, (15) is stronger than the first pair of inequalities in (13). The other pairs are equal or occur only in (14). *QED*

3. *Let the FOC of the subproblems be necessary. Then for every solution $(\mathbf{x}^*, \mathbf{y}^*)$ of problem (1) with $\mathbf{x}^* \neq \mathbf{0}$ there exist $t^*, \mathbf{a}^*, \boldsymbol{\mu}^*, \lambda^*$ such that $\mathbf{x}^* = t^* \cdot \mathbf{a}^*$ and the vector $(t^*, \mathbf{a}^*, \mathbf{y}^*, \boldsymbol{\mu}^*, \lambda^*)$ satisfies (14).*

Proof. If $\mathbf{x}^* \neq \mathbf{0}$, the conditions (2) are necessary and sufficient (step 1.). Thus there exist $\boldsymbol{\mu}^*$ such that $(\mathbf{x}^*, \mathbf{y}^*, \boldsymbol{\mu}^*)$ satisfies (2). Let $t^* = h(\mathbf{x}^*)$, $\mathbf{a}^* = \mathbf{x}^*/t^*$ and $\lambda^* = D_h f(t^*, \mathbf{x}^*, \mathbf{y}^*)$. Since $Dh(t \cdot \mathbf{a}^*) = Dh(\mathbf{a}^*)$, $(t^*, \mathbf{a}^*, \mathbf{y}^*, \boldsymbol{\mu}^*, \lambda^*)$ satisfies (14). *QED*

4. Let the FOC of the subproblems be necessary. Then for every solution $(\mathbf{x}^*, \mathbf{y}^*)$ of problem (1) with $\mathbf{x}^* = \mathbf{0}$ there exist $\mathbf{a}^*, \boldsymbol{\mu}^*, \lambda^*$ such that the vector $(t^*, \mathbf{a}^*, \mathbf{y}^*, \boldsymbol{\mu}^*, \lambda^*)$ with $t^* = 0$ satisfies (14).

Proof. Since $(\mathbf{0}, \mathbf{y}^*)$ is at least as good as any other point in $\mathbb{R}_+^n \times \mathbb{R}^m$, $(\mathbf{0}, \mathbf{y}^*)$ is a solution of Problem (12) for any \mathbf{a} . Thus for every \mathbf{a} there exists a $\boldsymbol{\mu}^*$ such that $(\mathbf{0}, \mathbf{a}, \mathbf{y}^*, \boldsymbol{\mu}^*)$ satisfies (13). Moreover, because of $t = 0$, $\boldsymbol{\mu}^*$ is independent from \mathbf{a} . The part of (14) that goes beyond (13), and therefore must be verified, is given by

$$(16) \quad \begin{aligned} (a) \quad & (\boldsymbol{\mu}^*)^T D_x g(\mathbf{0}, \mathbf{y}^*) - \\ & D_x f(\mathbf{0}, \mathbf{0}, \mathbf{y}^*) - \lambda \cdot Dh(\mathbf{a}) \geq \mathbf{0}, \mathbf{a} \geq \mathbf{0} \quad w.c.s. \\ (b) \quad & h(\mathbf{a}) \geq 1, \lambda \geq 0 \quad w.c.s. \\ (c) \quad & D_h f(\mathbf{0}, \mathbf{0}, \mathbf{y}^*) \leq \lambda., \end{aligned}$$

We show: (i) (16a,b) always have a solution $\lambda^* = \mathbf{q}^T \mathbf{a}^*$ and $h(\mathbf{a}^*) = 1$. (ii) λ^* satisfies (16c).

(i) With the help of the definition

$$\mathbf{q}^T \stackrel{\text{def}}{=} (\boldsymbol{\mu}^*)^T D_x g(\mathbf{0}, \mathbf{y}^*) - D_x f(\mathbf{0}, \mathbf{0}, \mathbf{y}^*) \geq \mathbf{0},$$

and because of $t = 0$, condition (13a) can be written in the following form:

$$(17) \quad \mathbf{q}^T \mathbf{a} - D_h f(\mathbf{0}, \mathbf{0}, \mathbf{y}^*) \geq 0$$

Because of $D_h f \geq 0$ this implies $\mathbf{q} \geq \mathbf{0}$, since (17) must hold for all $\mathbf{a} \geq \mathbf{0}$, $\mathbf{a} \neq \mathbf{0}$. We consider the set of \mathbf{a} satisfying $h(\mathbf{a}) = 1$. For such \mathbf{a} , (17) transforms into

$$(18) \quad [\mathbf{q}^T - D_h f(\mathbf{0}, \mathbf{0}, \mathbf{y}^*) \cdot Dh(\mathbf{a})] \mathbf{a} \geq 0.$$

Thus any solution of

$$(19) \quad \begin{aligned} (a) \quad & \mathbf{q}^T - \lambda \cdot Dh(\mathbf{a}) \geq \mathbf{0}, \mathbf{a} \geq \mathbf{0} \quad w.c.s. \\ (b) \quad & h(\mathbf{a}) \geq 1, \lambda \geq 0 \quad w.c.s. \\ (c) \quad & D_h f(\mathbf{0}, \mathbf{0}, \mathbf{y}^*) \leq \lambda, \end{aligned}$$

is a solution of (16). Since h is regular, there is a solution to (19a,b) satisfying $\lambda^* = \mathbf{q}^T \mathbf{a}^*$ and $h(\mathbf{a}^*) = 1$.

(ii) Since (18) is fulfilled for all \mathbf{a} with $h(\mathbf{a}) = 1$, (18) holds for \mathbf{a}^* , too, which implies

$$(20) \quad \lambda^* = \mathbf{q}^T \mathbf{a}^* \geq D_h f(0, \mathbf{0}, \mathbf{y}^*)$$

as required by (19c).

Thus $(0, \mathbf{a}^*, \mathbf{y}^*, \mu^*, \lambda^*)$ satisfies (14). *QED*

5. *Let the FOC of the subproblems be sufficient. If the vector $(t^*, \mathbf{a}^*, \mathbf{y}^*, \mu^*, \lambda^*)$ satisfies (14), then $(\mathbf{x}^*, \mathbf{y}^*)$ with $\mathbf{x}^* = t^* \cdot \mathbf{a}^*$ is a solution to problem (1).*

Proof. Assume that $(\mathbf{x}^*, \mathbf{y}^*)$ is *not* a solution. Then there must exist a $(\bar{\mathbf{x}}, \bar{\mathbf{y}})$ which is better than $(\mathbf{x}^*, \mathbf{y}^*)$. This leads to a contradiction. There are two possibilities.

(i) $\bar{\mathbf{x}}$ or \mathbf{x}^* , or both, are equal to $\mathbf{0}$. Then both are on the same ray through the origin defined by \mathbf{a}^* . Since (14) implies (13) (step 2.) and since the latter conditions guarantee that $(\mathbf{x}^*, \mathbf{y}^*)$ is optimal on that ray, $(\bar{\mathbf{x}}, \bar{\mathbf{y}})$ cannot be better. Contradiction!

(ii) $\bar{\mathbf{x}}$ and \mathbf{x}^* are both not equal to $\mathbf{0}$. In this case, (14) implies (2) (step 1.). The latter conditions guarantee that no (\mathbf{x}, \mathbf{y}) with $\bar{\mathbf{x}} \neq \mathbf{0}$ can be better than $(\mathbf{x}^*, \mathbf{y}^*)$. Contradiction! *QED*

6. *Steps 3. and 4. imply that the FOC (14) are necessary conditions if the FOC of the subproblems are necessary. Step 5. implies sufficiency if the FOC of the subproblems are sufficient. This completes the proof of theorem 3.*

The FOC (14) can be derived mechanically from the standard FOC (2) by the following procedure: In the *left-hand inequalities* of (2), replace $h(\mathbf{x})$ by t , \mathbf{x} by $t \cdot \mathbf{a}$, $Dh(\mathbf{x})$ by $Dh(\mathbf{a})$ and $D_h f(t, t \cdot \mathbf{a}, \mathbf{y})$ by an additional Lagrange multiplier λ . In the *right-hand inequalities*, replace \mathbf{x} by \mathbf{a} . Moreover, add the last two pairs of inequalities in order to determine t and λ .

The following theorem avoids explicit reference to the two subproblems of problem (1). It is analogous to theorem 10 in that it gives conditions under which the FOC (14) are *necessary and sufficient*.

Theorem 4 Let h be regular in the sense of Def. 1. Let f, g be differentiable everywhere in their domains. Let the function $m(\mathbf{x}, \mathbf{y}) \stackrel{\text{def}}{=} f(h(\mathbf{x}), \mathbf{x}, \mathbf{y})$ be quasiconcave. Let the admissible set $\{(\mathbf{x}, \mathbf{y}) : g(\mathbf{x}, \mathbf{y}) \leq 0, \mathbf{x} \geq 0, \mathbf{y} \geq 0\}$ be convex and contain an interior point. Let $(\mathbf{x}^*, \mathbf{y}^*)$ be admissible and $Dg_i(\mathbf{x}^*, \mathbf{y}^*) \neq 0$ for all constraints active at $(\mathbf{x}^*, \mathbf{y}^*)$. Let one of the following requirements be fulfilled:

- $Df(h(\mathbf{x}^*), \mathbf{x}^*, \mathbf{y}^*) \neq 0$
- m ist konkav.

Then $(\mathbf{x}^*, \mathbf{y}^*)$ is a solution of problem (1) if and only if there exists a $(t^*, \mathbf{a}^*, \mathbf{y}^*, \boldsymbol{\mu}^*, \lambda^*)$ with $\mathbf{x}^* = t^* \cdot \mathbf{a}^*$ satisfying (14).

Proof. The theorem follows immediately from theorems 10 and 3. *QED*

We now prove our claims about the modified problem (3). We need the following lemma.

Lemma 1 If problem (3) has a solution, it has a solution with $h(\mathbf{a}) = 1$.

Proof. Let (t^*, \mathbf{a}^*) be a solution with $h(\mathbf{a}^*) > 1$. Let $\mathbf{x}^* \stackrel{\text{def}}{=} t^* \cdot \mathbf{a}^*$. Then we may substitute t^* by $\bar{t} \stackrel{\text{def}}{=} h(\mathbf{x}^*) > t^*$ and \mathbf{a}^* by $\bar{\mathbf{a}} \stackrel{\text{def}}{=} \mathbf{a}^*/h(\mathbf{x}^*)$. This never yields a lower value of f since f is non-decreasing in h . Moreover, since we neither change $\mathbf{x}^* = \bar{t} \cdot \bar{\mathbf{a}} = t^* \cdot \mathbf{a}^*$ nor \mathbf{y}^* , the new point is still admissible. However, by definition, $h(\bar{\mathbf{a}}) = 1$. *QED*

Theorem 5 $(\mathbf{x}^*, \mathbf{y}^*)$ is a solution of problem (1) if and only if $(t^*, \mathbf{a}^*, \mathbf{y}^*)$ with $\mathbf{x}^* = t^* \cdot \mathbf{a}^*$ is a solution of problem (3).

Proof. By lemma 1, we can assume $h(\mathbf{a}) = 1$ w.l.o.g. for problem 3. Then any admissible point of one problem yields an admissible point for the other problem where the respective values of f are equal. *QED*

Theorem 6 If $(t, \mathbf{a}, \mathbf{y}, \boldsymbol{\mu}, \lambda)$ satisfies (14), it satisfies (4).

Proof. First note (4a,b) implies $D_h f(t, t \cdot \mathbf{a}, \mathbf{y}) \cdot t = \lambda$. Thus (4a) can be written as

$$(21) \quad \begin{aligned} & D_h f(t, t \cdot \mathbf{a}, \mathbf{y}) \cdot D h(\mathbf{a}) \cdot t + \\ & [D_x f(t, t \cdot \mathbf{a}, \mathbf{y}) - \boldsymbol{\mu}^T D_x g(t \cdot \mathbf{a}, \mathbf{y})] \cdot t \leq 0, \mathbf{a} \geq 0 \quad \text{w.c.s..} \end{aligned}$$

From (14e) we have $D_h f(t, t \cdot \mathbf{a}, \mathbf{y}) \cdot t = \lambda \cdot t$. This together with (14a) implies (21) and thus (4a).

Moreover, (14a) implies

$$(22) \quad \lambda + D_x f(t, t \cdot \mathbf{a}, \mathbf{y}) \mathbf{a} - \boldsymbol{\mu}^T D_x g(t \cdot \mathbf{a}, \mathbf{y}) \mathbf{a} = 0.$$

(22) together with (14e) implies (4b).

The other pairs of inequalities are either identical in both sets or occur only in (14). *QED*

Theorems (5) and (6) that the necessary and sufficient conditions (14) of problem (1) imply the standard FOC (4) resulting from the equivalent restatement (3).

We now show that the envelope theorem can be applied directly to problem (1) without modification. This has to be verified since the proof of the envelope theorem is based on the standard FOC.

Theorem 7 *Consider the maximum value function*

$$(23) \quad \pi(\boldsymbol{\theta}) \stackrel{\text{def}}{=} \max_{\mathbf{x}, \mathbf{y}} \{f(h(\mathbf{x}, \boldsymbol{\theta}), \mathbf{x}, \mathbf{y}, \boldsymbol{\theta}) : \mathbf{g}(\mathbf{x}, \mathbf{y}, \boldsymbol{\theta}) \leq \mathbf{0}, \mathbf{x} \geq \mathbf{0}, \mathbf{y} \geq \mathbf{0}\}$$

where $\boldsymbol{\theta} \in \mathbb{R}_+^l$ is a vector of parameters. Let the FOC (14)—modified by inclusion of the parameters $\boldsymbol{\theta}$ —be necessary and sufficient for all $\boldsymbol{\theta}$ in some open subset of \mathbb{R}^l containing $\boldsymbol{\theta}^*$. Let $\boldsymbol{\theta}^*$ be in the interior of a regime. Let the system of equalities consisting of the active conditions fulfill the requirements of the implicit-function theorem at $(t^*, \mathbf{a}^*, \mathbf{y}^*, \boldsymbol{\mu}^*, \lambda^*, \boldsymbol{\theta})$ where $(t^*, \mathbf{a}^*, \mathbf{y}^*, \boldsymbol{\mu}^*, \lambda^*)$ is a solution corresponding to $\boldsymbol{\theta}^*$. Define the Lagrangian

$$(24) \quad \mathbb{L}(\mathbf{x}, \mathbf{y}, \boldsymbol{\mu}, \boldsymbol{\theta}) \stackrel{\text{def}}{=} f(h(\mathbf{x}), \mathbf{x}, \mathbf{y}, \boldsymbol{\theta}) - \boldsymbol{\mu}^T \mathbf{g}(\mathbf{x}, \mathbf{y}, \boldsymbol{\theta}).$$

Under these conditions, we have $D\pi(\boldsymbol{\theta}^*) = D_{\boldsymbol{\theta}} \mathbb{L}(\mathbf{x}^*, \mathbf{y}^*, \boldsymbol{\mu}^*, \boldsymbol{\theta})$ where $\mathbf{x}^* \stackrel{\text{def}}{=} t^* \cdot \mathbf{a}^*$.

Proof. Consider the parametrized problem (3) together with its necessary and sufficient FOC (14) and the Lagrangian

$$(25) \quad \mathbb{L}(t, \mathbf{a}, \mathbf{y}, \boldsymbol{\mu}, \lambda, \boldsymbol{\theta}) \stackrel{\text{def}}{=} f(t, t \cdot \mathbf{a}, \mathbf{y}, \boldsymbol{\theta}) - \boldsymbol{\mu}^T \mathbf{g}(t \cdot \mathbf{a}, \mathbf{y}, \boldsymbol{\theta}) + \lambda \cdot (h(\mathbf{a}, \boldsymbol{\theta}) - 1).$$

Under the assumptions of the theorem, the envelope theorem can be proven as before using the FOC (14). Since problems (1) and problem (3) are equivalent, their optimal-value function are identical. We therefore get

$$(26) \quad D\pi(\boldsymbol{\theta}) \equiv D_{\boldsymbol{\theta}}\mathbb{L}(\cdot) \equiv D_{\boldsymbol{\theta}}f(t^*, \mathbf{x}^*, \mathbf{y}^*, \boldsymbol{\theta}) - (\boldsymbol{\mu}^*)^T D_{\boldsymbol{\theta}}\mathbf{g}(\mathbf{x}^*, \mathbf{y}^*) + \lambda^* \cdot D_{\boldsymbol{\theta}}h(\mathbf{a}^*, \boldsymbol{\theta}).$$

From the Lagrangian (24) we get

$$(27) \quad D_{\boldsymbol{\theta}}\mathbb{L}(\mathbf{x}^*, \mathbf{y}^*, \boldsymbol{\mu}^*, \boldsymbol{\theta}) = D_h f(h(\mathbf{x}^*), \mathbf{x}^*, \mathbf{y}^*, \boldsymbol{\theta}) \cdot D_{\boldsymbol{\theta}}h(\mathbf{a}^*, \boldsymbol{\theta}) + D_{\boldsymbol{\theta}}f(h(\mathbf{x}^*), \mathbf{x}^*, \mathbf{y}^*, \boldsymbol{\theta}) - (\boldsymbol{\mu}^*)^T D_{\boldsymbol{\theta}}\mathbf{g}(\mathbf{x}^*, \mathbf{y}^*).$$

This is equal to $D\pi(\boldsymbol{\theta})$ in (26), since, for all cases where $\mathbf{x}^* \neq \mathbf{0}$ and thus $t^* = h(\mathbf{x}^*) \neq 0$, (14) requires $\lambda^* = D_h f(h(\mathbf{x}^*), \mathbf{x}^*, \mathbf{y}^*, \boldsymbol{\theta})$. *QED*

4 A Problem in the Theory of Production

With perfect competition, concave production functions with constant returns to scale, and indifference of factor-owners between employments, the equilibrium conditions for the production sector of an economy can be stated, as is well-known, in the following form:

$$(28) \quad \begin{aligned} \mathbf{b}(\mathbf{w}) &\geq \mathbf{p}, \quad \mathbf{x} \geq \mathbf{0} \quad w.c.s. \\ \mathbf{x}^T D\mathbf{b}(\mathbf{w}) &\leq \mathbf{v}, \quad \mathbf{w} \geq \mathbf{0} \quad w.c.s. \end{aligned}$$

Here $\mathbf{p} \in \mathbb{R}_+^n$ denotes goods prices, $\mathbf{w} \in \mathbb{R}_+^m$ denotes factor prices, $\mathbf{x} \in \mathbb{R}_+^n$ denotes output, $\mathbf{v} \in \mathbb{R}_+^m$ denotes overall inputs, and $\mathbf{b} \in \mathbb{R}_+^n$ denotes unit costs. The unit cost functions $b_j(\mathbf{w})$ are defined by

$$(29) \quad b_j(\mathbf{w}) \stackrel{\text{def}}{=} \min_{\mathbf{v}^j} \{ \mathbf{w}^T \mathbf{v}^j : f_j(\mathbf{v}^j) \geq 1, \mathbf{v}^j \geq \mathbf{0} \}.$$

$b_j(\mathbf{w})$ is linearly homogenous, non-decreasing, and concave. By the envelope theorem, $D\mathbf{b}(\mathbf{w})$ is the matrix of input coefficients. On the question of differentiability, see appendix B. We assume the production function to be regular in the sense of def. 1 in section 3. This implies the existence of the cost functions for all non-negative factor prices but excludes, for example, CES production functions with an

elasticity of substitution smaller or equal to one. Woodland (1977) shows how to deal with non-regular production functions: the corresponding cost functions have to be defined by the *infimum* instead of the minimum. However, while it is surely to be appreciated if non-regular functions are covered, Pearce, in the context of the factor-price equalization controversy (e.g. Pearce 1970), argued convincingly that real-world production functions can never show the behavior we have excluded by the regularity assumption. Thus we opt for this assumption, since the analysis in section 3 is based on it.

Conditions (28) are the necessary and sufficient FOC of Samuelson's (1953) problem of factor-cost minimization:

$$(30) \quad r(\mathbf{p}, \mathbf{v}) \stackrel{\text{def}}{=} \min_{\mathbf{w}} \{ \mathbf{w}^T \mathbf{v} : \mathbf{b}(\mathbf{w}) \geq \mathbf{p}, \mathbf{w} \geq \mathbf{0} \}$$

Because of the "product exhaustion theorem", national factor costs are equal to the value of national production. Thus $r(\mathbf{p}, \mathbf{v})$ is the GNP function. This can also be seen from (28) directly.

Problem (30) is the dual to the more natural formulation

$$(31) \quad r(\mathbf{p}, \mathbf{v}) \stackrel{\text{def}}{=} \max_{\mathbf{v}^j} \left\{ \sum_{j=1}^n p_j \cdot f_j(\mathbf{v}^j) : \sum_{j=1}^n \mathbf{v}^j \leq \mathbf{v}, \mathbf{v}^j \geq \mathbf{0} \right\}.$$

The equivalence of problems (30) and (31) is not easily seen. Dixit & Norman (1980: 45f) give a proof that uses the optimality properties of equilibrium. There is, however, a simpler method of proof using necessary and sufficient FOC.

The usual FOC of problem (31), i.e.

$$(32) \quad \begin{aligned} (a) \quad & p_j \cdot Df_j(\mathbf{v}^j) \leq \boldsymbol{\mu}^T, \mathbf{v}^j \geq \mathbf{0} \quad w.c.s. \quad j = 1, \dots, n \\ (b) \quad & \sum_{j=1}^n \mathbf{v}^j \leq \mathbf{v}, \boldsymbol{\mu} \geq \mathbf{0} \quad w.c.s., \end{aligned}$$

are stated e.g. by Woodland (1982: 56). Obviously these conditions are correct descriptions of the production equilibrium if it is optimal to produce *all* the goods. The Lagrange multipliers $\boldsymbol{\mu}$ are to be interpreted as factor prices; the conditions, beyond restating the constraints, imply that factors not fully employed have zero prices, and factors are either paid their marginal value product or not used in a sector. However, the conditions cease to be meaningful if there is specialization on a subset of goods: If $\mathbf{v}^j = \mathbf{0}$, the marginal value products in sector j are not defined. Since

the production functions are concave, it is possible to define a subdifferential at the origin (see appendix B). However, as already mentioned above, any interpretation along these lines *implicitly* uses the same sort of arguments that are made explicit by the theorems proved in section 3. The FOC used by Woodland are defined in the admissible set, necessary and sufficient in this interpretation, but the FOC derived below are much easier to handle.

Probably to avoid these difficulties, Silberberg (1978: 446ff) considers the following equivalent restatement of (31):

$$(33) \quad \max_{\mathbf{x}, \mathbf{a}^j} \left\{ \mathbf{p}^T \mathbf{x} : f_j(\mathbf{a}^j) \geq 1, \sum_{j=1}^n \mathbf{a}^j \cdot \mathbf{x}_j \leq \mathbf{v}, \mathbf{x}, \mathbf{a}^j \geq \mathbf{0} \right\}$$

This is analogous to problem (3). However, this reformulation causes trouble instead of avoiding it. Silberberg discusses the standard FOC

$$(34) \quad \begin{aligned} (a) \quad & \lambda_j \cdot Df_j(\mathbf{a}^j) \leq \boldsymbol{\mu}, \mathbf{a}^j \geq \mathbf{0} \quad w.c.s. \quad j = 1, \dots, n \\ (b) \quad & \sum_{j=1}^n \mathbf{a}^j \cdot \mathbf{x}_j \leq \mathbf{v}, \boldsymbol{\mu} \geq \mathbf{0} \quad w.c.s. \\ (c) \quad & f_j(\mathbf{a}^j) \geq 1, \lambda_j \geq 0 \quad w.c.s. \quad j = 1, \dots, n \\ (d) \quad & \boldsymbol{\mu}^T \mathbf{a}^j \geq p_j, \mathbf{x}_j \geq 0 \quad w.c.s. \quad j = 1, \dots, n, \end{aligned}$$

which are defined in the admissible set and necessary. He fails to mention, however, that they are insufficient because they never enforce diversification. It is always possible to satisfy them by specialization on any good. In order to see this, just let $x_1 = f_1(\mathbf{v})$, $x_{j \neq 1} = 0$, $\lambda_1 = p_1 \cdot x_1$, $\lambda_{j \neq 1} = 0$ and $\mu_i = p_1 \cdot D_i f_1(\mathbf{v})$. Then it is possible to satisfy $\boldsymbol{\mu}^T \mathbf{a}^j \geq p_j$ for the goods not produced. In many cases, this can be done even if $f_j(\mathbf{a}^j) = 1$ is imposed. The crucial point is that conditions (34) require input coefficients to be cost-minimizing *only for goods that are actually produced*. This, however, yields no basis for an optimal decision as to whether a good should be produced or not.

Theorem 4, which obviously applies to (31), delivers the necessary and sufficient conditions:

$$(35) \quad \begin{aligned} (a) \quad & \lambda_j \cdot Df_j(\mathbf{a}^j) \leq \boldsymbol{\mu}^T, \mathbf{a}^j \geq \mathbf{0} \quad w.c.s. \quad j = 1, \dots, n \\ (b) \quad & \sum_{j=1}^n \mathbf{x}_j \cdot \mathbf{a}^j \leq \mathbf{v}, \boldsymbol{\mu} \geq \mathbf{0} \quad w.c.s. \\ (c) \quad & f_j(\mathbf{a}^j) \geq 1, \lambda_j \geq 0 \quad w.c.s. \quad j = 1, \dots, n \\ (d) \quad & \mathbf{p} \leq \boldsymbol{\lambda}, \mathbf{x} \geq \mathbf{0} \quad w.c.s. \end{aligned}$$

This results from successive applications of the theorem to all the production functions in the maximand of (31). These conditions give a complete picture of the production equilibrium; just note that λ is the unit-cost and μ the factor-price vector.

In order to prove the equivalence of the alternative definitions of the GNP function, we have to transform (28) such that the FOC of the cost-minimization problems (29) show up. The latter FOC are given by

$$(36) \quad \begin{aligned} \lambda_j \cdot Df_j(\mathbf{a}^j) &\leq \mathbf{w}^T, \mathbf{a}^j \geq \mathbf{0} \quad w.c.s. \\ f_j(\mathbf{a}^j) &\geq 1, \lambda_j \geq 0 \quad w.c.s. \end{aligned}$$

Since $b_j(\mathbf{w}) = \mathbf{w}^T \mathbf{a}^j = \lambda_j$ at the cost minimum, we can replace $b_j(\mathbf{w})$ in (28) by λ_j and add (36). The derivatives $Db_j(\mathbf{w})$ can be replaced by the input coefficients \mathbf{a}^j . The resulting FOC

$$(37) \quad \begin{aligned} (a) \quad &\lambda_j \cdot Df_j(\mathbf{a}^j) \leq \mathbf{w}^T, \mathbf{a}^j \geq \mathbf{0} \quad w.c.s. \quad j = 1, \dots, n \\ (b) \quad &\sum_{j=1}^n \theta_j \cdot \mathbf{a}^j \leq \mathbf{v}, \mathbf{w} \geq \mathbf{0} \quad w.c.s. \\ (c) \quad &f_j(\mathbf{a}^j) \geq 1, \lambda_j \geq 0 \quad w.c.s. \quad j = 1, \dots, n \\ (d) \quad &\mathbf{p} \leq \lambda, \boldsymbol{\theta} \geq \mathbf{0} \quad w.c.s. \end{aligned}$$

are again necessary and sufficient for problem (30). Obviously, (35) and (37) are equivalent: just let $\mu = \mathbf{w}$ and $\boldsymbol{\theta} = \mathbf{x}$. Since $\mathbf{w}^T \mathbf{v} = \mathbf{p}^T \mathbf{x}$ in both sets of FOC (by product exhaustion or by a simple manipulation of the FOC), the optimal-value functions of both problems are identical.

Section 3 demonstrated that the envelope theorem applies directly to problem (31) without any modification. Thus the well-known interpretation of the derivatives of the GNP function, i.e. $D_p r(\mathbf{p}, \mathbf{v}) = \mathbf{x}^T(\mathbf{p}, \mathbf{v})$ and $D_v r(\mathbf{p}, \mathbf{v}) = \mathbf{w}^T(\mathbf{p}, \mathbf{v})$, follows as usual.

5 Conclusion

The general problem treated in this paper may also turn up in other contexts than the theory of production. If preferences are assumed to be homothetic, the same problem may arise in a welfare-theoretical context even if a non-homogenous representation of preferences is chosen. A homothetic function $h(\mathbf{x})$ can be represented

as $f(g(\mathbf{x}))$ where g is linearly homogenous. If $Df(0) = 0$, or if g is linear, then h can be differentiable at the origin, otherwise not. It is not always convenient to use a problem dual to the original one to avoid problems at the origin. Moreover, the proof of the duality relationship between two problems can be difficult. In such cases, the theorems proven here might be helpful.

A The Kuhn–Tucker Approach

We start from a standard maximization problem

$$(38) \quad \max_{\mathbf{x}} \{f(\mathbf{x}) : \mathbf{g}(\mathbf{x}) \leq \mathbf{0}, \mathbf{x} \geq \mathbf{0}\}$$

with Lagrangian

$$(39) \quad \mathbb{L}(\mathbf{x}, \boldsymbol{\mu}) = f(\mathbf{x}) - \boldsymbol{\mu}^T \mathbf{g}(\mathbf{x})$$

and FOC

$$(40) \quad \begin{aligned} D_{\mathbf{x}} \mathbb{L}(\mathbf{x}, \boldsymbol{\mu}) &\leq \mathbf{0}, \mathbf{x} \geq \mathbf{0} \quad w.c.s. \\ D_{\boldsymbol{\mu}} \mathbb{L}(\mathbf{x}, \boldsymbol{\mu}) &\geq \mathbf{0}, \boldsymbol{\mu} \geq \mathbf{0} \quad w.c.s., \end{aligned}$$

yielding

$$(41) \quad \begin{aligned} Df(\mathbf{x}) - \boldsymbol{\mu}^T D\mathbf{g}(\mathbf{x}) &\leq \mathbf{0}, \mathbf{x} \geq \mathbf{0} \quad w.c.s. \\ \mathbf{g}(\mathbf{x}) &\leq \mathbf{0}, \boldsymbol{\mu} \geq \mathbf{0} \quad w.c.s. \end{aligned}$$

“w.c.s.” is short for “with complementary slackness”. Let us assume that the functions $f: \mathbb{R}_+^n \mapsto \mathbb{R}$ and $\mathbf{g}: \mathbb{R}_+^n \mapsto \mathbb{R}^m$ are differentiable everywhere in \mathbb{R}_+^n . Sometimes we will refer to the set $\{\mathbf{x} \in \mathbb{R}_+^n : \mathbf{g}(\mathbf{x}) \leq \mathbf{0}\}$ as the admissible set. The n conditions $\mathbf{x} \geq \mathbf{0}$ will be called *non-negativity conditions* while the term “constraints” will refer only to the m conditions $\mathbf{g}(\mathbf{x}) \leq \mathbf{0}$.

There are theorems specifying requirements for the functions f and \mathbf{g} such that the FOC (41) are necessary and/or sufficient for problem (38). The following three theorems apply to many of the cases interesting in economics:

Theorem 8 *Let \mathbf{x}^* be a solution to problem (38). Let f, \mathbf{g} be differentiable everywhere in \mathbb{R}_+^n . Let one of the following conditions hold:*

- *All the functions $g_i, i = 1, \dots, m$ are concave.*

- All the functions g_i , $i = 1, \dots, m$ are convex. Moreover, there exists a point \bar{x} in the admissible set such that $g(\bar{x}) < 0$.
- The admissible set is convex and contains an interior point. Moreover, if the i th constraint is not slack at x^* , $Dg_i(x^*) \neq 0$ holds.

Then there exists a vector μ^* such that (x^*, μ^*) satisfies the FOC (41).

Theorem 9 Let (x^*, μ^*) satisfy the FOC (41). Let the functions f , g be differentiable in \mathbb{R}_+^n . Let f be quasiconcave. Moreover, let the admissible set be convex and contain a point $\bar{x} \in \mathbb{R}_{++}^n$. Then x^* is a solution to problem (38) if one of the following conditions hold:

- $Df(x^*) \neq 0$
- f is concave.

Theorems 8 and 9 follow immediately from the Arrow–Hurwicz–Uzawa theorem or the Arrow–Enthoven theorem, respectively (Takayama 1985: 97f, 114f). The next theorem, which can be strengthened easily, is an obvious consequence of theorems 8 and 9:

Theorem 10 Let the functions f , g be differentiable in \mathbb{R}_+^n with f quasiconcave. Let the admissible set be convex and let it contain an interior point. Consider an admissible x^* with $Dg_i(x^*) \neq 0$ for all constraints that are active at x^* . Let one of the following conditions hold:

- $Df(x^*) \neq 0$
- f is concave.

Then x^* is a solution to problem (38) if and only if there is a vector μ^* such that (x^*, μ^*) satisfies the FOC (41).

In section 3, we consider the applicability of the envelope theorem. For this purpose we have to look at the proof of this theorem. We start from the following parametrized maximum problem and the resulting maximum value function:

$$(42) \quad \pi(\theta) \stackrel{\text{def}}{=} \max_x \{f(x, \theta) : g(x, \theta) \leq 0, x \geq 0\}$$

$\theta \in \mathbb{R}_+^l$ is a vector of l parameters. (42) includes the case that some of the functions f, g_i depend only on some (or none) of the parameters. We make use of the Lagrangian

$$(43) \quad \mathbb{L}(\mathbf{x}, \boldsymbol{\mu}, \boldsymbol{\theta}) = f(\mathbf{x}, \boldsymbol{\theta}) - \boldsymbol{\mu}^T \mathbf{g}(\mathbf{x}, \boldsymbol{\theta})$$

and of the FOC

$$(44) \quad \begin{aligned} D_{\mathbf{x}}f(\mathbf{x}, \boldsymbol{\theta}) - \boldsymbol{\mu}^T D_{\mathbf{x}}\mathbf{g}(\mathbf{x}, \boldsymbol{\theta}) &\leq 0, \mathbf{x} \geq 0 \quad w.c.s. \\ \mathbf{g}(\mathbf{x}, \boldsymbol{\theta}) &\leq 0, \boldsymbol{\mu} \geq 0 \quad w.c.s. \end{aligned}$$

From (44) it follows that at the maximum $\boldsymbol{\mu}^T \mathbf{g}(\mathbf{x}, \boldsymbol{\theta}) = 0$; thus with solutions $\mathbf{x}^*, \boldsymbol{\mu}^*$ we have $\pi(\boldsymbol{\theta}) \equiv \mathbb{L}(\mathbf{x}^*, \boldsymbol{\mu}^*, \boldsymbol{\theta})$, i.e. the optimal value function is identical to the Lagrangian at the maximum.

In order to state the envelope theorem, we use the concept of a *regime* of parameter values (cf. Dixit & Norman 1980: 320–1). A regime is a set of parameter values $\boldsymbol{\theta}$ for which the same constraints and non-negativity conditions are active or slack at the maximum.

Theorem 11 (Envelope Theorem) *Consider the function π as defined by (42). Assume that the FOC (44) are necessary and sufficient. Let $\boldsymbol{\theta}^*$ be in the interior of a regime. Let the system of equalities consisting of the active conditions satisfy the requirements of the implicit-function theorem at $(\mathbf{x}^*, \boldsymbol{\mu}^*, \boldsymbol{\theta}^*)$ where $(\mathbf{x}^*, \boldsymbol{\mu}^*)$ is a solution corresponding to $\boldsymbol{\theta}^*$. Then we have $D\pi(\boldsymbol{\theta}^*) = D_{\boldsymbol{\theta}}\mathbb{L}(\mathbf{x}^*, \boldsymbol{\mu}^*, \boldsymbol{\theta}^*)$.*

Proof. The system of equalities consisting of the *active* conditions implicitly defines differentiable solution functions $\mathbf{x}(\boldsymbol{\theta})$ and $\boldsymbol{\mu}(\boldsymbol{\theta})$ for problem (42) in a neighborhood of $(\mathbf{x}^*, \boldsymbol{\mu}^*, \boldsymbol{\theta}^*)$. From the chain rule we get

$$(45) \quad \begin{aligned} D\pi(\boldsymbol{\theta}) &\equiv D_{\boldsymbol{\theta}}\mathbb{L}(\mathbf{x}^*, \boldsymbol{\mu}^*, \boldsymbol{\theta}) && + \\ &D_{\mathbf{x}}\mathbb{L}(\mathbf{x}^*, \boldsymbol{\mu}^*, \boldsymbol{\theta}) \cdot D\mathbf{x}^*(\boldsymbol{\theta}) && + \\ &D_{\boldsymbol{\mu}}\mathbb{L}(\mathbf{x}^*, \boldsymbol{\mu}^*, \boldsymbol{\theta}) \cdot D\boldsymbol{\mu}^*(\boldsymbol{\theta}) \quad . \end{aligned}$$

The Kuhn–Tucker conditions imply that the derivatives of the Lagrangian (a) w.r.t. x_j and (b) w.r.t. μ_j are equal to zero (a) if $x_j > 0$ or (b) if the constraint belonging to μ_j is active, respectively. On the other hand, (c) if $x_j = 0$ or (d) if the constraint belonging to μ_j is slack, we have (c) $D_i \mathbf{x}^*(\boldsymbol{\theta}) = 0$ or (d) $D_i \boldsymbol{\mu}^*(\boldsymbol{\theta}) = 0$, respectively,

since we assumed that there is no change in regime in a neighborhood of $(\mathbf{x}^*, \boldsymbol{\mu}^*, \boldsymbol{\theta}^*)$. Thus the second and third term in (45) are zero. *QED*

At the border between regimes, this argument does not hold. However, the result $D\pi(\boldsymbol{\theta}^*) = D_\theta \mathbb{L}(\mathbf{x}^*, \boldsymbol{\mu}^*, \boldsymbol{\theta}^*)$ might hold nevertheless (cf. Takayama (1985: 137f) for sufficient conditions).

B Differentiability

A general remark on differentiability is in order since the paper deals with a problem of non-differentiability. In all cases considered here, differentiability may be replaced by *subdifferentiability* (cf. e.g. Diewert 1974: 158).

Subdifferentiability is a generalization of differentiability. The set of points on and below the graph of a concave function $f(\mathbf{x})$, called *epi* f , has supporting hyperplanes everywhere (except possibly at the boundary of the function's domain where even a concave function might not be continuous). If there is only one supporting hyperplane at \mathbf{x}^* , the function is differentiable at \mathbf{x}^* ; if there are more, it is subdifferentiable. The subgradient at \mathbf{x}^* is a vector \mathbf{s} such that $\begin{pmatrix} 1 \\ -\mathbf{s} \end{pmatrix}$ is normal to one of the supporting hyperplanes at \mathbf{x}^* , and

$$(46) \quad f(\mathbf{x}) - f(\mathbf{x}^*) \geq \mathbf{s}^\top (\mathbf{x} - \mathbf{x}^*)$$

for all \mathbf{x} in the domain of f . The subdifferential $Df(\mathbf{x}^*)$ denotes the set of all (transposed) subgradients. Equations involving Df are satisfied if there is at least one member in the set satisfying them.

If a function is not concave, subgradients still may exist in the sense that (46) holds for all \mathbf{x} in a neighborhood of \mathbf{x}^* . While subdifferentiability is guaranteed for concave functions (and an analogue obviously holds for *convex* functions), functions that are neither concave nor convex may be sub- or superdifferentiable wherever they are not differentiable. For most of the considerations here, this generalization of differentiability will do.

Non-differentiability of linearly homogenous functions at the origin sometimes can be treated by analogy. If the function $h(\mathbf{x})$ defined in \mathbb{R}_+^n is linearly homogenous and *concave*, the set of subgradients at the origin is the union of all the subgradients

in $\mathbb{R}_+^n \setminus \{0\}$. This set can be fairly large. While it is possible to make use of the subdifferential here, the interpretation of the FOC of an optimization problem is much easier if one does not.

References

- Diewert, W. E. (1974), "Applications of duality theory", in: Intriligator & Kendrick (1974), 106–171
- Dixit, A. K./ Norman, V. D. (1980), *Theory of international trade*, Welwyn: Nisbet & Co./ Cambridge University Press
- Intriligator, M. D./ Kendrick, D. A. (eds.) (1974), *Frontiers of quantitative economics, Vol. II*, Amsterdam: North-Holland
- Pearce, I. (1970), *International trade, Book II*, London: Macmillan
- Samuelson, P. A. (1953), "Prices of factors and goods in general equilibrium", repr. in: Stiglitz, J. E. (ed.) (1966), *The collected scientific papers of Paul A. Samuelson, Vol. II*, Cambridge/Mass.: MIT Press, 888–908
- Shepard R. W. (1974), "Comments", in: Intriligator & Kendrick (1974), 200–206
- Silberberg, E. (1978), *The structure of economics*, New York: McGraw-Hill
- Takayama, A. (1985), *Mathematical economics*, 2. ed., Cambridge: Cambridge University Press
- Woodland, A. D. (1977), "A dual approach to equilibrium in the production sector in international trade theory", *Canadian Journal of Economics* 10, 50–68
- Woodland, A. D. (1982), *International trade and resource allocation*, Amsterdam: North-Holland