



NIST
PUBLICATIONS

Applied and
Computational
Mathematics
Division

NISTIR 6074

Computing and Applied Mathematics Laboratory

Limits of SUMT Trajectories in Convex Programming

*Garth P. McCormick,
The George Washington University
Christoph Witzgall, NIST*

September 1997

U.S. DEPARTMENT OF COMMERCE
Technology Administration
National Institute of Standards and Technology
Gaithersburg, MD 20899

QC
100
.U56
NO. 6074
1997

Limits of SUMT Trajectories in Convex Programming

Garth P. McCormick

The George Washington University
Emeritus, SEAS
Department of Operations Research
Washington, DC 20052

Christoph Witzgall

U.S. DEPARTMENT OF COMMERCE
Technology Administration
National Institute of Standards
and Technology
Information Technology Laboratory
Mathematical and Computational Sciences Division
Gaithersburg, MD 20899-0001

September 1997



U.S. DEPARTMENT OF COMMERCE
William M. Daley, Secretary

TECHNOLOGY ADMINISTRATION
Gary Bachula, Acting Under Secretary for Technology

NATIONAL INSTITUTE OF STANDARDS
AND TECHNOLOGY
Robert E. Hebner, Acting Director

Abstract

The limits of a class of primal and dual solution trajectories associated with the Sequential Unconstrained Minimization Technique (SUMT) based on the logarithmic barrier function are investigated for convex programming problems in the presence of multiple optima and degeneracy conditions. For linear programming problems, such limits can be characterized as analytic centers of, loosely speaking, optimality regions. Of interest here are generalizations of those results.

A class of convex programming problems, identified by what is termed the “rank-integrity property”, is introduced for which primal trajectory limits can be characterized in analogy to the linear case and without differentiability conditions. This class of problems contains linear and convex quadratic programming problems as strict subsets. Similar, albeit more restricted, results are obtained for dual trajectories defined in the differentiable case.

Keywords: analytic center, barrier function, convex programming, degeneracy, interior point method, optimization, rank-integrity, SUMT, trajectory.

1. Introduction

Consider the following “*primal*” convex programming problem defined on a convex domain $C \subseteq R^n$,

$$(1.1) \quad \begin{array}{ll} \text{Minimize} & f(x) \\ & \text{subject to} \\ & g_i(x) \geq 0, \quad \text{for } i = 1, \dots, m, \quad x \in R^n, \end{array}$$

where the objective function $f(x)$ is continuous and convex on C , and the constraint functions $g_i(x)$, $i = 1, \dots, m$, are continuous and concave on C .

If an optimal solution x^* exists, let $t^* = f(x^*)$ denote the minimum value of the programming problem. The sets

$$X^F = \{x : g_i(x) \geq 0, i = 1, \dots, m\} \subseteq C, \quad X^* = \{x : f(x) = t^*, x \in X^F\} \subseteq X^F$$

of feasible and optimal solution will here be referred to as “*feasibility region*” and “*optimality region*”, respectively. Both are closed convex sets. We will restrict our analysis to convex programming problems for which

$$(1.2) \quad \text{the optimality region } X^* \text{ is nonempty, bounded and, consequently, compact,}$$

and we are particularly interested in those cases in which the optimality region contains more than one point.

An important consequence of the boundedness of the optimality region is that (see Fiacco and McCormick [8], Theorem 24)

(1.3) *all level sets $\{x \in X^F : f(x) \leq t\}$ are bounded.*

Once and for all, we will require the standard

(1.4) *Constraint qualification of Slater [20]: There exists a point $\bar{x} \in R^n$ such that for all constraints $g_i(x) \geq 0$, $i = 1, \dots, m$,*

$$g_i(\bar{x}) > 0.$$

That condition guarantees that the interior $\text{int}X^F$ of the feasibility region is nonempty, and that none of the constraint functions vanishes identically on X^F .

The Slater constraint qualification is clearly necessary and sufficient for the definition of the “*logarithmic barrier function*”

$$(1.5) \quad \sum_{i=1}^m \ln[g_i(x)],$$

which tends to $-\infty$ as x approaches the boundary of the feasibility region X^F . Use of the logarithmic barrier function for perturbing the objective function in order to approximate optimal solutions has been suggested early on by Frisch [9]. Following the general SUMT approach, a family of (*logarithmic*) *barrier perturbations* are introduced on $\text{int}X^F$:

$$(1.6) \quad L(x, r) = f(x) - r \left(\sum_{i=1}^m \ln[g_i(x)] \right), \quad r > 0,$$

These functions are clearly continuous and convex in the variable x for each specific parameter value r . They grow infinitely large close to the boundary of the feasibility region X^F . Any minimizer

$$x(r) := \operatorname{argmin}_x L(x, r)$$

of such a barrier perturbation is therefore an unconstrained minimizer. The idea is to reduce the perturbation successively in order for the unconstrained minimizers

$x(r)$ to converge towards a – constrained – optimal solution $x^* \in X^*$.

It is an important result that those minimizers always exist – even if the feasibility region X^F is not bounded – , provided the optimality region X^* is bounded (assumption (1.2)). It then follows that the minimizers accumulate towards the optimality region. Indeed the following statement (see Fiacco and McCormick [8], Theorem 25) – which extends also to classes of barrier functions other than the logarithmic ones – is the basis of barrier-driven Sequential Unconstrained Minimization Techniques (SUMT) pioneered by Fiacco and McCormick [5], [6],[7].

(1.7) *If the optimality region X^* of the convex programming problem (1.1) is nonempty and bounded, then*

- (i) *each barrier perturbation $L(x, r)$, $r > 0$ has a minimizer $x(r)$;*
- (ii) *the values of minimizers $x(r)$ converge, that is,*

$$f[x(r)] \rightarrow t^* \quad \text{as } r \rightarrow 0$$

holds, where $x(r)$ is any minimizer of the barrier perturbation $L(x, r)$;

- (iii) *$\underline{r} < \bar{r}$ implies $f[x(\underline{r})] \leq f[x(\bar{r})]$ for any minimizers $x(\underline{r})$ and $x(\bar{r})$, respectively.*

Under very general conditions, the minimizers $x(r)$ will be unique. Then those minimizers describe a continuous

(1.8) “(primal) trajectory” $x = x(r)$, $r > 0$.

Indeed, any sequence of minimizers $x(r_i), r_i \rightarrow r_0 > 0$, must have an accumulation point \bar{x} by (1.7,iii) and (1.3). Since $L(x, r)$ is continuous in both x and r , $L[x(r_i), r_i] \leq L[x(r_0), r_i]$ implies in the limit $L(\bar{x}, r_0) \leq L[x(r_0), r_0]$. As the minimizer $x(r_0)$ is unique, $\bar{x} = x(r_0)$. Thus $x(r_i)$ converges towards $x(r_0)$, whence the trajectory $x = x(r)$ is continuous at any $r = r_0 > 0$.

The trajectory $x = x(r)$ and its relatives have received considerable attention McCormick [13], [14], Anstreicher [2], Bayer and Lagarias [3], [4]. An algorithmic convergence theory based on the concept of “self-concordant” barrier functions has been developed by Nesterov and Nemirovskii [16], [17]. The same trajectories generally arise for the method of centers proposed originally by Huard [10],[11], investigated later by Sonnevend [21] and, for the case of linear programming problems, by Renegar [18],[19]. The projective interior point method proposed and proved to be of polynomial complexity by Karmarkar [12] is based on related concepts.

If there is a unique optimal solution x^* , that is, if the optimality region X^* consists of a single point, then the primal trajectory converges to x^* in view of (1.7). But what happens if there are multiple optimal points? Does a limit exist, and if so, how can that particular optimal point be characterized?

In the linear programming case, the feasibility region X^F is a closed convex polyhedron, and the optimality region X^* is either a single vertex or, in the case of multiple optimal solutions, an entire exposed set or (“*face*”) of that polyhedron such as an edge or a higher dimensional analogue (see for instance Stoer and Witzgall [22]). Megiddo [15] observed and, Adler and Monteiro [1], Witzgall, Boggs, and Domich [23]) proved, in more general settings, that the minimizers $x(r)$ converge towards the “analytic center” of the optimal face. One goal of this work is to investigate generalizations of this result.

We call a convex programming problem (1.1) “*differentiable*” if the functions $f(x)$ and $g_i(x), i = 1 \dots m$, are differentiable in the interior of their definition region C , and if the feasibility region is contained in that interior:

$$X^F \subset \text{int}C.$$

For such differentiable convex programming problems it is possible to define a dual trajectory associated with the primal trajectory $x = x(r)$. Indeed, in view of the convexity and differentiability of the logarithmic barrier perturbation $L(x, r)$, vanishing gradients are necessary and sufficient to characterize its minimizers $x(r)$. Considering gradients as column vectors, we may write that optimality condition as follows:

$$(1.9) \quad \nabla f[x(r)] = \sum_{i=1}^m \nabla g_i[x(r)] \frac{r}{g_i[x(r)]}.$$

Here the vectors $u(r)$ with nonnegative components

$$(1.10) \quad u_i(r) = \frac{r}{g_i[x(r)]}, \quad i = 1, \dots, m,$$

are usually unique and, if so, define a

$$\text{“dual trajectory”} \quad u = u(r)$$

which shadows the primal trajectory.

This trajectory relates to the “*Wolfe-dualization*” [24] of the convex programming problem (1.1):

(1.11) Maximize $f(x) - \sum_{i=1}^m g_i(x)u_i$ subject to

$$\begin{aligned} \nabla f(x) - \sum_{i=1}^m \nabla g_i(x)u_i &= 0, \\ u &= (u_1, \dots, u_m)^T \geq 0, \quad x \in C. \end{aligned}$$

The combination $[x(r), u(r)]$ constitutes for every parameter $r > 0$ a feasible solution of the Wolfe-dualization (compare [7]). Furthermore, consider the – necessary and sufficient – KKT optimality condition for an optimal solution x^* of the primal convex programming problem. It requires the existence of an m –vector $u^* \geq 0$ of KKT multipliers such that

$$\nabla f(x^*) = \sum \nabla g_i(x^*)u_i^*,$$

while satisfying the complementary slackness condition

$$\sum_{i=1}^m g_i(x^*)u_i^* = 0.$$

Note that those conditions also imply that (x^*, u^*) is an optimal solution to the Wolfe-dualized problem (1.11) and, conversely, that if (x^*, u^*) is an optimal solution, then x^* is an optimal solution of the primal problem and u^* is a corresponding vector of KKT multipliers.

For the unique limit $x^* \in X^*$ of the primal trajectory $x = x(r)$ – should it exist – , we may consider the convex set

$$(1.12) \quad U^* = U^*(x^*) = \{u \in R^m : (x^*, u) \text{ Wolfe dual optimal (1.11)}\}$$

as a “*dual optimality region*”.

Indeed, for linear programming problems, the Wolfe-dualization creates the familiar dual linear programming problem – with the primal variables dropping out. The set U^* then becomes the optimality region of that dual linear program, and the dual trajectory can be interpreted as a SUMT trajectory of the dual linear programming problem converging to the analytic center of the dual optimality region. We will examine generalizations of this dual result, too.

The material will be presented as follows: In Section 2, following the Introduction, analytic centers of primal optimality regions are introduced, followed by a section

of examples and counterexamples. In Section 4, rank-integrity is defined and key properties derived. The first theorem characterizes limits of the primal trajectory under rank-integrity. It will be proved in Section 5. The last four sections introduce analytic centers of dual optimality regions, and prove the second theorem about conditions under which those analytic centers are limits of dual trajectories.

2. Analytic Centers of the Primal Optimality Region

We are particularly interested in the case in which the optimality region X^* of the convex programming problem (1.1) consists of more than one point, and we wish to establish a concept of “analytic center of the optimality region”.

The classic ([10],[11],[21]) definition of the analytic center of a system of constraints

$$(2.1) \quad g_i(x) \geq 0, \quad i = 1, \dots, m,$$

is as maximizer of the product of the constraint functions or, equivalently if the interior $\text{int}X^F$ of the feasibility region is not empty (see 1.4), as maximizer of the logarithmic barrier function (1.5),

$$\text{argmax}_x \left(\prod_{i=1}^m g_i(x) \right) = \text{argmax}_x \left(\sum_{i=1}^m \ln[g_i(x)] \right).$$

Such analytic centers exist if the feasibility region X^F is bounded.

Note that this definition pertains to systems of constraints and not to their solution sets: Different systems of constraints with identical solution sets may give rise to different analytic centers, each maximizing a different sum of constraint logarithms. Adding a redundant constraint is a case in point.

In the linear programming case, a system of constraints which characterizes the optimality region can be obtained in canonical fashion from the constraints of the given linear programming problem: turn the inequalities which are “*universally binding*”, that is, satisfied as equalities for every optimal solution, into equations. The resulting system of linear inequalities and equations may not have interior solutions. As a result, the definition of analytic centers needs a slight generalization: only those constraints that are not universally binding are included in the sum of constraint logarithms to be maximized.

In our more general case of convex programming, there appears to be no natural

system of optimality constraints that derives from the feasibility constraints, and leads to a direct canonical characterization of the analytic center of optimality regions. Furthermore, the optimality region of programming problems that are linear is always an exposed set of the feasibility region. For nonlinear convex programming problems, the optimality region may not even be an extreme set.

Returning to the linear programming case, the analytic center of a bounded optimality region can also be characterized as the maximizer over the optimality region X^* of a logarithmic barrier function again constructed from those feasibility constraints which are not universally binding at optimality. That alternate definition differs conceptionally from the earlier one in that the analytic center is defined with respect to the optimality region X^* as a set rather than a characterizing system of constraints. It is this characterization of the analytic center of the optimality region that extends naturally to convex programming problems. Let thus

$$(2.2) \quad Z = \{1 \leq i \leq m : g_i(x) = 0 \text{ for all } x \in X^*\},$$

denote the set of the indices of constraints that are universally binding at optimal solutions of problem (1.1).

The concept of universally binding constraints is not as compelling here as for linear programming problems. Indeed, any system of constraints (2.1) can – in many ways – be replaced by a single constraint $g(x) \geq 0$ with a single concave function $g(x)$, and this single sufficient constraint will be universally binding, unless there are optimal solutions in the interior of the feasibility region X^F . By contrast, linear programming problems with bounded optimality regions will always have constraints that are not universally binding. We denote the index set associated with constraints that are not universally binding by

$$(2.3) \quad N = \{1 \leq i \leq m : i \notin Z\},$$

and define an “*analytic center of the (primal) optimality region*” by the following optimization problem:

$$(2.4) \quad \text{Maximize } \prod_{i \in N} g_i(x) \quad \text{subject to } x \in X^*.$$

If $N = \emptyset$, then we consider any $x \in X^*$ an analytic center of the optimality region. This convention may lead to multiple analytic centers. We will give an example later where the above maximization problem does not define a unique analytic center even if $N \neq \emptyset$. Note that by that definition

(2.5) *analytic centers x^* form a nonempty convex set and satisfy*

$$g_i(x^*) > 0 \text{ for } i \in N.$$

Proof: At least one analytic center x^* exists since X^* is supposed to be nonempty and bounded by assumption (1.2).

If the index set N is empty, nothing more needs to be shown. Otherwise, we observe that the optimality region X^* contains a point \hat{x} with $g_i(\hat{x}) > 0$ for all $i \in N$. Indeed, let i be some index in N . If $g_i(x) = 0$ for every $x \in X^*$, then $i \in Z$ by definition. Thus there exists $x^{(i)} \in X^*$ with $g_i(x^{(i)}) > 0$. Define

$$\hat{x} := \frac{1}{|N|} \left(\sum_{i \in N} x^{(i)} \right),$$

where $|N|$ denotes the cardinality N . Since X^* is convex, $\hat{x} \in X^*$ and $g_i(\hat{x}) > 0$ for all $i \in N$. The objective function of problem (2.4) assumes a positive value for \hat{x} and, therefore, for x^* . Thus $g_i(x^*) > 0$ for $i \in N$.

This observation also makes it possible to pass to the logarithm of the objective function of problem (2.4), that is, the corresponding sum of logarithms

$$\sum_{i \in N} \ln[g_i(x)],$$

without affecting the definition of analytic centers. That function is concave, whence the set of maximizers convex. \square

We observe, that

(2.6) *in the absence of constraints that are universally binding at optimality, that is, if $Z = \emptyset$, every accumulation point of minimizers $x(r)$ of the barrier perturbation $L(x, r)$ is an analytic center (2.4) of the optimality region.*

Proof: Because all constraint indices i are in N , there exists by (2.5) an analytic center x^* with $g_i(x^*) > 0, i = 1, \dots, m$, that is, $x^* \in \text{int}X^F$. Thus $L(x^*, r)$ is defined for $r > 0$, and

$$L[x(r), r] = f[x(r)] - r \sum_{i=1}^m \ln[g_i[x(r)]] \leq f(x^*) - r \sum_{i=1}^m \ln[g_i[x^*]] = L(x^*, r).$$

Since $f(x^*) = t^*$,

$$0 \leq f[x(r)] - t^* \leq r \left(\sum_{i=1}^m \ln(g_i[x(r)]) - \sum_{i=1}^m \ln[g_i(x^*)] \right),$$

and

$$\prod_{i=1}^m g_i[x(r)] \geq \prod_{i=1}^m g_i(x^*).$$

Let $\bar{x} \in X^*$ be an accumulation point of $x(r)$ as $r \rightarrow 0$. Then

$$\prod_{i=1}^m g_i[x(r)] - \prod_{i=1}^m g_i(\bar{x}) \geq \prod_{i=1}^m g_i(x^*) - \prod_{i=1}^m g_i(\bar{x}) \geq 0.$$

The left hand side of the above expression converges to zero by continuity. Thus

$$\prod_{i=1}^m g_i(\bar{x}) = \prod_{i=1}^m g_i(x^*),$$

which was to be shown. \square

The absence of universally binding constraints at optimality is, of course, equivalent to the existence of optimal solutions in interior of the feasible region. Also if, in this case, the analytic center x^* of the optimality region is unique, then (2.6) implies that it is the limit of the primal trajectory $x = x(r)$, – which is the behavior we wish to establish in general. In the presence of universally binding constraints, however, it may not hold, as some of the examples in the following section illustrate.

3. Examples and Counterexamples for the Primal Limit Conjecture

In this section we examine analytic centers of some convex programming problems in order to illustrate the issues raised previously. Consider the linear programming problem for $x = (x_1, x_2)$

$$(3.1) \quad \begin{array}{ll} \text{Minimize} & -2x_2 - x_3 \quad \text{subject to} \\ & g_1(x) = x_1 \geq 0, \\ & g_2(x) = x_2 \geq 0, \\ & g_3(x) = x_3 \geq 0, \\ & g_4(x) = 2 - x_1 - 2x_2 - x_3 \geq 0, \\ & g_5(x) = 4 + 3x_1 - 4x_2 - 2x_3 \geq 0. \end{array}$$

The feasible set X^F is a polytope, and the optimality region X^* is an edge of that polytope, characterized by $g_1(x) = g_4(x) = g_5(x) = 0$. Thus both kinds of degeneracies are present at optimality: optimal solutions are not unique, and neither are the KKT multipliers due to the linear dependence of the above three constraints $g_i(x), i \in Z = \{1, 4, 5\}$. Since this convex programming problem is differentiable, the optimality condition (1.9) can be applied to characterize primal trajectory points $x = x(r)$. The gradient

$$\nabla f(x) = \begin{pmatrix} 0 \\ -2 \\ -1 \end{pmatrix},$$

can be expressed in the form:

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \frac{r}{x_1} + \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \frac{r}{x_2} + \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \frac{r}{x_3} + \begin{pmatrix} -1 \\ -2 \\ -1 \end{pmatrix} \frac{r}{2 - x_1 - 2x_2 - x_3} + \begin{pmatrix} +3 \\ -4 \\ -2 \end{pmatrix} \frac{r}{4 + 3x_1 - 4x_2 - 2x_3}$$

This condition yields the following expressions for the primal trajectory:

$$x_1(r) = \frac{2 + 5r - \sqrt{4 - 4r + 25r^2}}{\sqrt{19} - 1} = \frac{4(\sqrt{19} + 1)r}{3(2 + 5r + \sqrt{4 - 4r + 25r^2})},$$

$$x_2(r) = \frac{2 - 5r + \sqrt{4 - 4r + 25r^2}}{8}, \quad x_3(r) = \frac{2 - 5r + \sqrt{4 - 4r + 25r^2}}{4}.$$

Clearly,

$$x_1(r) \rightarrow 0, \quad x_2(r) \rightarrow \frac{1}{2}, \quad x_3(r) \rightarrow 1,$$

as $r \rightarrow 0$. These limits also describe the solution of the optimization problem

$$(3.2) \quad \begin{array}{ll} \text{Maximize} & x_2 x_3 \\ & \text{subject to} \\ & g_1(x) = x_1 = 0, \\ & g_2(x) = x_2 \geq 0, \\ & g_3(x) = x_3 \geq 0, \\ & g_4(x) = 2 - x_1 - 2x_2 - x_3 = 0, \\ & g_5(x) = 4 + 3x_1 - 4x_2 - 2x_3 = 0, \end{array}$$

which characterizes the analytic center of the optimality region. To determine that maximum, the reader may choose to substitute $x_3 = 2 - 2x_2$ and to find the unconstrained maximum of $2x_2(1 - x_2)$.

In trying to extend this result to the general convex case, however, difficulties are encountered. Consider the functions

$$p(x_1, x_2) = \sqrt{(x_1 - 1)^2 + x_2^2}, \quad q(x_1, x_2) = \sqrt{(x_1 + 1)^2 + x_2^2},$$

and

$$(3.3) \quad f(x_1, x_2) = p(x_1, x_2) + q(x_1, x_2) - 2.$$

Both $p(x_1, x_2)$ and $q(x_1, x_2)$ are convex functions describing circular cones with apex at the points

$$\begin{pmatrix} +1 \\ 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} -1 \\ 0 \end{pmatrix},$$

respectively. The functions are linear on each ray emanating from their apex. They are strictly convex on every line not containing the apex. It follows that the function $f(x_1, x_2)$ is piecewise linear on the x -axis, and is strictly convex everywhere else. It assumes a minimum of zero on the interval spanned by the two apexes.

In the previous section, we promised to provide an example of analytic centers being not unique even if there are nonuniversally binding constraints, $N \neq \emptyset$, and if the SUMT minimizers $x(r)$ are unique. To this end consider the convex programming problem

$$\begin{aligned} \text{Minimize} \quad & f(x_1, x_2) && \text{subject to} \\ & g_1(x) = x_2 \geq 0 \\ & g_2(x) = 1 - x_2 \geq 0 \end{aligned}$$

thus has the bounded optimality region $X^* = \{(x_1, x_2)^T : -1 \leq x_1 \leq +1, x_2 = 0\}$ with $N = \{2\}$. The corresponding function $-\ln(1 - x_2)$, which is supposed to characterize the analytic center of the optimality region X^* , is constant there. Thus every optimal solution is an analytic center. On the other hand, the objective function $f(x_1, x_2)$ is strictly convex above the x -axis. The resulting barrier perturbations $L(x, r)$ are, therefore, strictly convex in the interior of the feasible region, the primal trajectory is well-defined, and has the unique limit $x^* = (0, 0)^T$. The objective function $f(x_1, x_2)$ is not differentiable at the apexes $(+1, 0)^T, (-1, 0)^T$, the focal points of the ellipses which describe its level sets. But the function $f(x_1, x_2)^2$ is differentiable and, subject to the above constraints, shows the same abnormal behavior.

The following examples show that, even if a unique analytic center of the optimality region exists, it may not be the limit of a well-defined primal trajectory. First,

consider the convex programming problem defined on $C = R_{\oplus}^2 = \{(x_1, x_2)^T : x_1, x_2 \geq 0\}$.

$$(3.4) \quad \begin{array}{ll} \text{Minimize } & x_2 \\ & \text{subject to} \\ & g_1(x) = x_1 \geq 0, \\ & g_2(x) = 2 - x_1 \geq 0, \\ & g_3(x) = x_2 \geq 0, \\ & g_4(x) = \sqrt{x_1 x_2} \geq 0. \end{array}$$

All barrier perturbations are differentiable in the interior $\text{int}X^F$ of the feasibility region. So the optimality condition (1.9) for minimizing a barrier perturbation $L(x, r)$ applies:

$$\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \frac{r}{x_1} + \begin{pmatrix} -1 \\ 0 \end{pmatrix} \frac{r}{2 - x_1} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} \frac{r}{x_2} + \frac{1}{2} \begin{pmatrix} \sqrt{x_2/x_1} \\ \sqrt{x_1/x_2} \end{pmatrix} \frac{r}{\sqrt{x_1 x_2}}.$$

For the point $(x_1(r), x_2(r))^T$ to lie on the primal trajectory then requires that

$$\frac{1}{x_1} - \frac{1}{2 - x_1} + \frac{1}{2x_1} = 0, \quad r \left(\frac{1}{x_2} + \frac{1}{2x_2} \right) = 1,$$

It follows that $x_1(r) = 6/5$, $x_2(r) = 3r/2$, and therefore $\lim_{r \rightarrow 0} x_1(r) = 6/5$. On the other hand, $N = \{1, 2\}$ and so the center $(x_1^*, x_2^*)^T$ of the optimality region is characterized by $x_2^* = 0$ and (see (2.4)) by the following maximization problem for x_1 :

$$\text{Maximize } x_1(2 - x_1) \quad \text{subject to } 0 < x_1 < 2.$$

The solution is clearly $x_1^* = 1 \neq 6/5$. The center of the optimality region thus differs from the limit of the primal trajectory for this example. Note, that the concave constraint functions have vertical tangents at boundary points of the feasible region, that is, for $x_1 = 0$ and $x_2 = 0$. They are therefore not differentiable at those points.

An even stronger counterexample to the conjecture that a unique analytic center is the limit of the primal trajectory $x = x(r)$, refers back to the function $f(x_1, x_2)$ defined above (3.3).

$$(3.5) \quad \begin{array}{ll} \text{Minimize } & f(x_1, x_2) \\ & \text{subject to} \\ & g_1(x) = x_1 \geq 0 \\ & g_2(x) = x_2 \geq 0. \end{array}$$

The feasibility region contains interior points, and the optimality region $X^* = \{(x_1, x_2)^T : -1 \leq x_1 \leq +1, x_2 = 0\}$ is bounded. Again, the barrier perturbation,

$$L(x_1, x_2, r) = \sqrt{(x_1 - 1)^2 + x_2^2} + \sqrt{(x_1 + 1)^2 + x_2^2} - r[\ln(x_1) + \ln(x_2)], \quad r > 0,$$

is differentiable in the positive orthant $R_{\oplus}^2 = \{(x_1, x_2)^T : x_1, x_2 > 0\}$. The necessary and sufficient optimality condition for the SUMT minimizer

$$x(r) = \begin{pmatrix} x_1 = x_1(r) \\ x_2 = x_2(r) \end{pmatrix}$$

thus is

$$(3.6) \quad \frac{x_1 - 1}{\sqrt{(x_1 - 1)^2 + x_2^2}} + \frac{x_1 + 1}{\sqrt{(x_1 + 1)^2 + x_2^2}} = r \left(\frac{1}{x_1} \right)$$

$$(3.7) \quad \frac{x_2}{\sqrt{(x_1 - 1)^2 + x_2^2}} + \frac{x_2}{\sqrt{(x_1 + 1)^2 + x_2^2}} = r \left(\frac{1}{x_2} \right).$$

For the following calculations, we abbreviate:

$$\begin{aligned} \sqrt{-} &:= \sqrt{(x_1 - 1)^2 + x_2^2} = \sqrt{x_1^2 + x_2^2 + 1 - 2x_1} \\ \sqrt{+} &:= \sqrt{(x_1 + 1)^2 + x_2^2} = \sqrt{x_1^2 + x_2^2 + 1 + 2x_1}. \end{aligned}$$

The optimality condition is a system of linear equations for $1/\sqrt{-}$ and $1/\sqrt{+}$ with solutions

$$\frac{1}{\sqrt{-}} = -\frac{r}{2} \left(\frac{1}{x_1} - \frac{x_1 + 1}{x_2^2} \right), \quad \frac{1}{\sqrt{+}} = -\frac{r}{2} \left(\frac{x_1 - 1}{x_2^2} - \frac{1}{x_1} \right),$$

or, since x_1 and x_2 are both positive,

$$\begin{aligned} 2x_1x_2^2 &= -r(x_2^2 - x_1^2 - x_1)\sqrt{-} \\ 2x_1x_2^2 &= -r(x_1^2 - x_2^2 - x_1)\sqrt{+}, \end{aligned}$$

and

$$(x_2^2 - x_1^2 - x_1)\sqrt{-} = (x_1^2 - x_2^2 - x_1)\sqrt{+}.$$

Squaring both sides of the above equation yields the necessary condition

$$(x_2^2 - x_1^2 - x_1)^2(x_1^2 + x_2^2 + 1 - 2x_1) = (x_1^2 - x_2^2 - x_1)^2(x_1^2 + x_2^2 + 1 + 2x_1).$$

We collect by common factors $(x_1^2 + x_2^2 + 1)$ and $2x_1$, finding, in view of

$$\begin{aligned}(x_2^2 - x_1^2 - x_1)^2 - (x_1^2 - x_2^2 - x_1)^2 &= 4x(x_1^2 - x_2^2) \\ (x_2^2 - x_1^2 - x_1)^2 + (x_1^2 - x_2^2 - x_1)^2 &= 2[(x_1^2 - x_2^2)^2 + x_1^2],\end{aligned}$$

that

$$4x_1(x_1^2 - x_2^2)(x_1^2 + x_2^2 + 1) = 2[(x_1^2 - x_2^2)^2 + x_1^2]2x_1.$$

Dividing by $4x_1 > 0$, yields

$$(x_1^2 - x_2^2)(x_1^2 + x_2^2 + 1) = (x_1^2 - x_2^2)^2 + x_1^2.$$

or, expanded,

$$2x_1^2x_2^2 - 2x_2^4 - x_2^2 = 0.$$

Dividing by $x_2^2 > 0$, finally leads to the following necessary optimality condition for the SUMT minimizers in this example:

$$2x_1(r)^2 - 2x_2(r)^2 - 1 = 0.$$

This is the equation of the arc of a hyperbola with asymptote $x_1 = x_2$. Also,

$$r = \sqrt{2} \left(\frac{4x_1^2 - 2}{4x_1^2 - 1} \right) x_1 = \sqrt{2} \left(\frac{4x_2^2}{4x_2^2 + 1} \right) x_1.$$

By continuity, and because $x_2(r) \rightarrow 0$ as $r \rightarrow 0$,

$$\begin{pmatrix} x_1(r) \\ x_2(r) \end{pmatrix} \rightarrow \begin{pmatrix} 1/\sqrt{2} \\ 0 \end{pmatrix} \text{ as } r \rightarrow 0.$$

This point, however, differs from the unique analytic center of the optimality region,

$$x^* = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

Indeed, $N = \{1\}$ and

$$\prod_{i \in N} g_i(x_1, x_2) = x_1,$$

the maximum of which over the interval $\{(x_1, x_2)^T : -1 \leq x_1 \leq +1, x_2 = 0\}$ is clearly assumed for $x_1^* = 1$.

Again we note, that the function $f(x_1, x_2)$ (3.3) is not differentiable at the points $(+1, 0)^T, (-1, 0)^T$, but its square, $f(x_1, x_2)^2$ is differentiable at these points and everywhere in X^F . After dividing by $2f(x_1, x_2) > 0$, the optimality condition for SUMT minimizers then takes the form (compare (3.6)):

$$\begin{aligned} \frac{x_1 - 1}{\sqrt{(x_1 - 1)^2 + x_2^2}} + \frac{x_1 + 1}{\sqrt{(x_1 + 1)^2 + x_2^2}} &= \frac{r}{2f(x_1, x_2)} \left(\frac{1}{x_1} \right) \\ \frac{x_2}{\sqrt{(x_1 - 1)^2 + x_2^2}} + \frac{x_2}{\sqrt{(x_1 + 1)^2 + x_2^2}} &= \frac{r}{2f(x_1, x_2)} \left(\frac{1}{x_2} \right). \end{aligned}$$

This shows that the squared function gives rise to same primal trajectory as the original function, only the dependence on the parameter r is affected. It follows that the discrepancy between trajectory limit and analytic center may obtain also for fully differentiable objective functions and linear constraints.

4. Rank-Integrity

As the examples in the previous section show, the characterization of the limit of the primal trajectory $x = x(r)$ as an analytic center requires a restriction of the class of programming problems (1.1) to be considered. We proceed to define such a restriction.

We say that a convex or concave function $c(x)$ on R^n has

$$(4.1) \quad \text{“rank-integrity”}$$

if for any line L and any line segment $S = [\underline{x}, \bar{x}] \subset L$ of positive length, the following holds:

$$(4.2) \quad \text{If the function } c(x) \text{ is constant on the segment } S, \text{ then it is constant on the entire line } L \text{ and on any line parallel to } L.$$

We proceed to characterize functions with rank-integrity. We call any function of a single real variable θ

$$(4.3) \quad \text{“constant-free”}$$

if there is no interval $\theta_1 < \theta < \theta_2$ on which the function is constant. We then call a function $Q(s), s \in R^n$, “constant-free” if its 1-dimensional restrictions are constant-free (4.3).

Constant functions as well as convex (concave) constant-free functions – such as strictly convex (concave) functions – have rank-integrity. In what follows, we will provide more detailed characterizations of rank-integrity without always mentioning these trivial cases. We will also restrict ourselves to discussing the convex case since the concave case is analogous.

(4.4) *A convex function $c(x)$ on R^n has rank-integrity if and only if it is direct sum of a constant function and a constant-free convex function. In other words, a convex function $c(x)$ has rank-integrity if and only if it is of the form*

$$c(x) = Q(Ax),$$

where A denotes an $l \times n$ -matrix, and $Q(s), s \in R^l$ is convex and constant-free.

Proof. Suppose the convex function $c(x)$ has rank-integrity and let M be the set of all directions w in which $c(x)$ is constant along a line. By the definition of rank-integrity, those directions indicate constant behavior regardless of the starting point. Thus it follows immediately that the set M is a linear subspace of R^n . This subspace may be described as the null-space of a matrix A of full rank: $M = \{w : Aw = 0\}$. Let M^\perp denote the orthogonal complement of M , and split any $x \in R^n$ into its two components in those orthogonal subspaces: $x = w + \hat{x}$. Then

$$c(x) = c(w + \hat{x}) = c(\hat{x}) = c[A^T(AA^T)^{-1}Ax] = Q(Ax).$$

We have to show that the convex function

$$Q(s) = c[A^T(AA^T)^{-1}s]$$

is constant-free. Assume to the contrary that $\underline{s} \neq \bar{s}$ define a constant line segment of $Q(s)$. Then $\underline{x} = A^T(AA^T)^{-1}\underline{s}$ and $\bar{x} = A^T(AA^T)^{-1}\bar{s}$ are points in R^n with $\underline{s} = A\underline{x}$ and $\bar{s} = A\bar{x}$. It follows that the points \underline{x}, \bar{x} bracket a constant line segment S for $c(x)$. By rank-integrity, $c(x)$ is then constant everywhere in direction $w = \bar{x} - \underline{x}$, which thus must lie in the null-space $Aw = A\bar{x} - A\underline{x} = \bar{s} - \underline{s} = 0$. This contradicts the assumption $\underline{s} \neq \bar{s}$. The convex function $Q(s)$ is therefore constant-free.

Conversely, suppose $c(x) = Q(Ax)$ with $Q(s)$ convex and constant-free. Suppose further that points $\underline{x} \neq \bar{x}$ are endpoints of a line segment S on which $c(x)$ has constant values. For $0 \leq \theta \leq 1$,

$$c[\underline{x} + \theta(\bar{x} - \underline{x})] = Q[A\underline{x} + \theta(A\bar{x} - A\underline{x})] = Q[\underline{s} + \theta(\bar{s} - \underline{s})],$$

where $\underline{s} = A\underline{x}$, $\bar{s} = A\bar{x}$. If $\underline{s} \neq \bar{s}$, the function $Q(s)$ would be constant on the line segment $AS \subset R^l$ in violation of the assumption that $Q(s)$ is constant-free. Thus $\underline{s} = \bar{s}$. This implies $Aw = 0$ for the direction $w = \bar{x} - \underline{x}$, and

$$c(x + \theta w) = Q(Ax + \theta Aw) = Q(Ax) = c(x)$$

for any x and θ . \square

Linear and convex quadratic functions $c(x)$ have rank-integrity. Indeed, those functions belong obviously to a more restricted class of functions with rank-integrity characterized by the property:

(4.5) *If the function $c(x)$ is linear on a segment S , then it is linear on any line L parallel to S .*

It follows again that each such function is the direct sum of a linear function and a strictly convex function, that is, it is of the form

$$c(x) = \alpha + a^T x + Q(Ax)$$

with A an $l \times n$ -matrix and $Q(s)$, $s \in R^l$, a strictly convex function. Note that functions of the form

$$c(x) = \alpha + a^T x + Q_1(A_1 x) + \dots + Q_k(A_k x),$$

with strictly convex functions Q_1, \dots, Q_k , are also of that form since additivity holds: $Q_1(A_1 x) + Q_1(A_2 x) = Q(Ax)$ with $Q = Q_1 \oplus Q_2$ the - strictly convex - direct sum of two strictly convex functions and

$$A = \begin{pmatrix} A_1 \\ A_2 \end{pmatrix}.$$

We will now examine implications of rank-integrity for the objective and constraint functions of convex programming problems. For convenience, we say the convex programming problem (1.1) has “rank-integrity” if the functions

$$f(x), g_i(x), i = 1, \dots, m,$$

have rank-integrity (4.1). For the remainder of this section, we will assume that this property holds. We first note that, in this case,

(4.6) *the barrier perturbations $L(x, r), r > 0$, have unique minimizers $x(r)$.*

Proof. Assume $L(x, r)$ has two minimizers $\underline{x}(r) \neq \bar{x}(r)$. Because of the convexity of $L(x, r)$, the two minimizers span a line segment S on which $L(x, r) = f(x) - r \sum_i \ln g_i(x), r > 0$ is constant. In general, if the sum of convex functions is linear and, in particular, constant, then each individual function must be linear. Thus $\ln g_i(x)$ will be linear on line segment S .

The following general property of the logarithm function is readily established by comparing derivatives.

(4.7) *If the logarithm of a positive function $\gamma(\theta)$ is linear in θ , then it is constant in θ .*

The linearity of $\ln g_i(x)$ on line segment S thus implies that $g_i(x)$ is constant on S . By rank-integrity, every constraint function must then be constant and positive on the entire line $L \supset S$. That line is therefore contained in the feasibility region X^F .

It follows that also the objective function $f(x) = L(x, r) + r \sum_i \ln g_i(x)$ is constant on line segment S , whence on the entire line L , again by rank-integrity. Thus $f(x)$ would have an unbounded level set in X^F , contradicting (1.3). (If the objective function $f(x)$ belongs to the special class (4.5) of functions with rank-integrity such as linear and convex quadratic functions, then the barrier perturbations $L(x, r)$ are actually strictly convex in x .) \square

Recall that there need not be any constraints $g_i(x)$ which are not universally binding at the optimality region; in other words, the index set N , on which the definition (2.4) of an analytic center of the optimality region is based, may be empty. In this case, all optimal solutions were defined to be analytic centers. Furthermore, as an example in Section 2 shows, the analytic center may be non-unique even if $N \neq \emptyset$ and the barrier perturbations $L(x, r)$ are strictly convex. The following proposition shows that these difficulties are avoided by the rank-integrity assumption.

(4.8) *The optimality region X^* has a unique analytic center x^* as defined by the maximization problem (2.4). Moreover x^* lies in the relative interior $\text{relint} X^*$ of the optimality region, and $g_i(x^*) > 0$ for $i \in N$.*

Proof. The existence of at least one analytic center was shown earlier (2.5).

The proposition is obviously true if X^* consists of a single point (even if $N = \emptyset$ as the relative interior of a single point consists of that point). We assume therefore that X^* has multiple points.

Next we show that X^* is essentially bounded by the constraints $g_i(x)$ with $i \in N$. Using the notation “affine” to indicate the affine hull or span of a set, and selecting a subset of the constraints by affixing the subscript N , we claim:

$$(4.9) \quad X^* = \{x \in \text{affine}X^* : g_N(x) \geq 0\} \quad (X^* = \text{affine}X^* \text{ if } N = \emptyset).$$

The inclusion $X^* \subseteq \{x \in \text{affine}X^* : g_N(x) \geq 0\}$ is trivial. If $X^* = \text{affine}X^*$ then the converse inclusion is also trivial. In order to prove that converse inclusion in general, select a point $\bar{x} \in \text{relint}X^*$, and let $x^\# \in \text{affine}X^*$ with $x^\# \notin X^*$. We have to show, that $x^\#$ violates some constraint: $g_i(x^\#) < 0$, $i \in N$.

Recall the definition (2.2) of the index set Z as characterizing the constraints in $g_Z(x)$ which are universally binding at optimality, that is, vanish for $x \in X^*$. Note that since X^* is convex, there exists $0 < \hat{\theta} < 1$ such that $\bar{x} + \theta(x^\# - \bar{x}) \in X^*$ for $0 \leq \theta \leq \hat{\theta}$ (see for instance Stoer and Witzgall [22], Lemma 3.2.9). For any $j \in Z$, the function $g_j[\bar{x} + \theta(x^\# - \bar{x})]$ is a constant, namely vanishing, function of θ on the interval $[0, \hat{\theta}]$. By rank-integrity, that function vanishes for all values θ . Thus $g_Z(x^\#) = 0$.

By the same token, the objective function $f(x)$ is seen to be constant at its optimal value, whence $f(x^\#) = t^*$. Thus only the violation of some constraint, $g_i(x^\#) < 0$, prevents $x^\#$ from being an optimal solution. Since $g_i(x)$ does not vanish on $\text{affine}X^*$, the index i is not in Z , and belongs therefore to its complement N . This proves (4.9).

The following characterization of the relative interior of the optimality region is simply a restatement of (4.9):

$$\text{relint}X^* = \{x \in \text{aff}X^* : g_N(x) > 0\} \neq \emptyset.$$

It follows again from (2.5) that any analytic center is contained in $\text{relint}X^*$.

We now proceed in analogy to the proof of (4.6), showing that the concave function $L^*(x) = \sum_{i \in N} \ln g_i(x)$ has a unique maximum on $\text{relint}X^*$. Assume to the contrary that there are two different maxima, that is, two different analytic centers $\underline{x}^* \neq \bar{x}^*$. Then these points define a line segment S of positive length on which $L^*(x)$ is constant. The individual summands $\ln g_i(x)$, $i \in N$, must therefore be linear. and

(4.7) implies that they are actually constant on the segment S . By rank-integrity, this constant behavior extends to the entire line $L \supset S$. In view of (4.9), this would imply $L \subseteq X^*$, violating the assumption that the optimality region is bounded. \square

5. Limits of Primal SUMT Trajectories

Trajectory points $x = x(r)$ may be characterized as optimal solutions to an auxiliary programming problem,

$$(5.1) \quad \text{Problem } \mathbf{B}_r : \quad \text{Maximize} \quad \sum_{i \in N} \ln[g_i(x)] \quad \text{subject to}$$

$$\begin{aligned} f(x) &\leq f[x(r)] \\ g_i(x) &> 0 \quad \text{for } i \in N \\ g_i(x) &\geq g_i[x(r)] \quad \text{for } i \in Z. \end{aligned}$$

Indeed, we will show that

$$(5.2) \quad \text{for any } r > 0, x(r) \text{ is the unique optimal solution of the maximization problem } \mathbf{B}_r.$$

Proof: Let x be an arbitrary feasible solution of problem \mathbf{B}_r above. Note first that, by the monotonicity of the logarithm function and the above conditions on x ,

$$\begin{aligned} \sum_{i \in N} \ln(g_i[x(r)]) - \sum_{i \in N} \ln[g_i(x)] \\ &= \sum_{i=1}^m \ln(g_i[x(r)]) - \sum_{i=1}^m \ln[g_i(x)] - \sum_{i \in Z} (\ln(g_i[x(r)]) - \ln[g_i(x)]) \\ &\geq \sum_{i=1}^m \ln(g_i[x(r)]) - \sum_{i=1}^m \ln[g_i(x)]. \end{aligned}$$

Next, because of the constraint $f(x) \leq f[x(r)]$ and since $r > 0$,

$$\begin{aligned} r \sum_{i=1}^m \ln(g_i[x(r)]) - r \sum_{i=1}^m \ln[g_i(x)] \\ &= f(x) - r \sum_{i=1}^m \ln[g_i(x)] - f[x(r)] - r \sum_{i=1}^m \ln(g_i[x(r)]) + f[x(r)] - f(x) \\ &\geq L(x, r) - L[x(r), r] \geq 0. \end{aligned}$$

This shows that, in order to maximize problem \mathbf{B}_r , it is necessary for x to minimize the barrier perturbation $L(x, r)$. Proposition 5.2) follows, as $x(r)$ is the only minimizer by (4.6) and also clearly a feasible solution of problem \mathbf{B}_r . \square

We are now ready to prove the first of two theorems.

(5.3) **Theorem 1:** *If the optimality region of a convex programming problem (1.1) is bounded, and if the Slater constraint qualification (1.4) holds, then the limit of the primal trajectory $x = x(r)$ (1.8) for $r \rightarrow 0$ is characterized as the unique analytic center x^* (2.4) of the optimality region X^* , if (i) the functions $f(x)$, $g_i(x)$, $i = 1 \dots m$, in the convex programming problem have rank-integrity (4.1); or if (ii) the optimality region has interior points and the analytic center is unique.*

Proof: The second part of the theorem is a corollary of the proposition (2.6).

As to the first part, there exists by (4.8) a unique analytic center $x^* \in X^*$, and $g_N(x^*) > 0$ holds. Consider a sequence $\{r_j\}$, $j = 1, \dots$, of positive barrier parameters decreasing to 0. According to the basic fact (1.7), requiring only continuity assumptions, the sequence of minimizers $x(r_j)$ has an accumulation point \bar{x} with

$$f[x(r_j)] \rightarrow t^* = f(\bar{x}), \quad \bar{x} \in X^*, \quad g_Z(\bar{x}) = 0.$$

Without loss of generality, we may assume that the accumulation point is actually the limit

$$x(r_j) \rightarrow \bar{x} \quad \text{as } j \rightarrow \infty,$$

of that sequence, as otherwise we would select a suitable subsequence. We now define

$$(5.4) \quad \tilde{x}(r_j) = x(r_j) + x^* - \bar{x},$$

where x^* is the unique analytic center. The idea of the proof is to utilize (5.2) by establishing $\tilde{x}(r_j)$, for sufficiently large j , as a feasible solution of problem \mathbf{B}_{r_j} , thus gaining an upper bound on corresponding values of its objective function.

First, note that $\tilde{x}(r_j) \rightarrow x^*$ follows by definition (5.4). In view of $g_N(x^*) > 0$, $g_N[\tilde{x}(r_j)] > 0$ will hold for suitably large indices j .

Next, we observe that the functions $f(x)$, $g_i(x)$, $i \in Z$, are constant on the line segment $S = \{\bar{x} + \theta(x^* - \bar{x}) : 0 \leq \theta \leq 1\}$. Indeed, the ends of S are in the convex

optimality region X^* , whence $f(x)$ is constant on S , and $g_Z(x)$ vanishes there. By rank-integrity, the above functions are constant everywhere in the direction $x^* - \bar{x}$. Since $\tilde{x}(r) - x(r_j) = x^* - \bar{x}$ by definition, it follows that there is constant behavior in that direction, too:

$$(5.5) \quad f[\tilde{x}(r_j)] = f[x(r_j)], \quad g_Z[\tilde{x}(r_j)] = g_Z[x(r_j)].$$

These relations, together with the previous observation, ensure that $\tilde{x}(r_j)$ is a feasible solution for problem \mathbf{B}_{r_j} . Since $x(r_j)$ is the optimal solution of that problem,

$$\sum_{i \in N} \ln g_i[\tilde{x}(r_j)] \leq \sum_{i \in N} \ln g_i[x(r_j)],$$

or, in terms of products of constraints,

$$\prod_{i \in N} g_i[\tilde{x}(r_j)] \leq \prod_{i \in N} g_i[x(r_j)].$$

In the limit $j \rightarrow \infty$, the above relation becomes

$$\prod_{i \in N} g_i(x^*) \leq \prod_{i \in N} g_i(\bar{x}).$$

Thus \bar{x} is a maximizer of problem (2.4). By (4.8), $\bar{x} = x^*$. Hence every accumulation point of the trajectory $x = x(r)$ coincides with x^* , proving Theorem 1. \square

6. Analytic Centers of the Dual Optimality Region

For differentiable convex programming problems, we will also investigate limits of the dual trajectory (1.10) defined in the Introduction, and again we wish to generalize results known to hold for linear programming problems.

In what follows, we will always assume that

$$(6.1) \quad \text{the primal trajectory } x = x(r), \quad r > 0 \text{ has a unique limit } x^* \text{ as } r \rightarrow 0.$$

Note that the limit x^* is not required to be the analytic center of the primal optimality region. We then revisit the set (1.12)

$$U^* = \{u \in R^m : (x^*, u) \text{ Wolfe dual optimal (1.11)}\}$$

which plays the role of a dual optimality region. U^* is convex. To see this, recall that by definition of the Wolfe dualization, the vectors $u \in U^*$ are the KKT multipliers confirming the primal optimal solution x^* . It is readily seen that if u^1 and u^2 are both vectors of KKT multipliers for the same primal optimal solution x^* , then so is any convex combination $\lambda u_i^1 + \mu u_i^2$, $\lambda + \mu = 1$, $\lambda, \mu \geq 0$.

In order to dualize the concept of analytic center, we introduce the index sets

$$(6.2) \quad V = \{1 \leq i \leq m : u_i = 0 \text{ for all } u \in U^*\}$$

and its complement

$$(6.3) \quad M = \{1 \leq i \leq m : i \notin V\},$$

where u_i denotes the i -th component of vector u .

In view of the complementary slackness condition $g_i(x^*)u_i = 0$ for KKT multipliers, we can express

$$(6.4) \quad U^* = \{u \geq 0 : \nabla f(x^*) - \sum_{i \in M} \nabla g_i(x^*)u_i = 0, u_i = 0 \text{ for } i \in V\}.$$

We now define the “*analytic center of the dual optimality region*” by the following optimization problem:

$$(6.5) \quad \text{Maximize } \prod_{i \in M} u_i \quad \text{subject to } u \in U^*.$$

For $M = \emptyset$, we admit every point $u^* \in U^*$ as an analytic center. Again it will be convenient to reformulate this definition in terms of logarithms. To this end we have to ensure that the dual optimality region contains points for which those logarithms are defined.

(6.6) *There exists a point $\hat{u} \in U^*$ such that $\hat{u}_i > 0$ for all $i \in M$. Consequently*

$$i \in M \Rightarrow g_i(x^*) = 0.$$

Proof. If $M = \emptyset$, the above is trivial. Let then j be some index in M . If $u_j = 0$ for every $u \in U^*$, then j would by definition be in V . Thus there exists a point $u^{(j)} \in U^*$ with $u_j^{(j)} > 0$. Define

$$\hat{u} = \frac{1}{|M|} \left(\sum_{j \in M} u^{(j)} \right),$$

where $|M|$ is the cardinality of M . Since U^* is convex, $\hat{u} \in U^*$ and $\hat{u}_i > 0$ for all $i \in M$. Also, the complementary slackness condition inherent in the definition of U^* holds and implies $g_i(x^*)\hat{u}_i = 0$ for all i . Since $\hat{u}_i > 0$ for $i \in M$, the second part of the proposition follows. (Note that if x^* is a "strict minimizer", that is, $g_i(x^*) = 0$ if and only if $i \in Z$, then (6.6) implies $M \subseteq Z$). \square

Because of (6.6), problem (6.5) can be reformulated as

(6.7) Maximize $\sum_{i \in M} \ln u_i$ subject to

$$u_i > 0 \text{ for } i \in M \quad \text{and} \quad u_i = 0 \text{ for } i \in V.$$

Indeed, the objective function of problem (6.5) has a positive value for any \hat{u} as defined above. Hence its maximizer u^* , - if it exists -, satisfies $u_i^* > 0$ for all $i \in M$. The constraints $u_i \geq 0$ in (6.5) may therefore be replaced by $u_i > 0$ without affecting the solution of the problem. The same is true if the objective function is replaced by its logarithm.

Note that, since the sum of the logarithms $\ln u_i$ is a strictly concave function of the variables $u_i, i \in M$, and since the remaining variables are fixed at zero,

(6.8) *if an analytic center u^* of the dual optimality region U^* exists, then it is unique.*

Again we conjecture that, for convex programming problems, dual analytic centers characterize the limits of dual trajectories in much the same way as primal analytic centers do for primal trajectories, and we are interested in finding conditions under which that conjecture can be decided.

7. Examples and Counterexamples for the Dual Limit Conjecture

Returning to our first example (3.1), we recall that for the limit of the primal trajectory $x = x(r)$,

$$x(0) = \begin{pmatrix} 0 \\ 1/2 \\ 1 \end{pmatrix}.$$

and that, for $r > 0$,

$$\frac{x_1(r)}{2 - x_1(r) - 2x_2(r) - x_3(r)} = \frac{3 + \sqrt{19}}{5}, \quad \frac{x_1(r)}{4 + 3x_1(r) - 4x_2(r) - 2x_3(r)} = \frac{-2 + \sqrt{19}}{15}.$$

These expressions define the dual trajectory $u = u(r)$ and yield, in view of the primal trajectory limit $x(0)$, the following dual limits as the barrier parameter r tends to zero:

$$\begin{aligned} u_1(r) &= \frac{r}{x_1(r)} \rightarrow u_1(0) = \frac{-1 + \sqrt{19}}{6}, \\ u_2(r) &= \frac{r}{x_2(r)} \rightarrow u_2(0) = 0, \\ u_3(r) &= \frac{r}{x_3(r)} \rightarrow u_3(0) = 0, \\ u_4(r) &= \frac{r}{2 - x_1(r) - 2x_2(r) - x_3(r)} \rightarrow u_4(0) = \frac{8 + \sqrt{19}}{15}, \\ u_5(r) &= \frac{r}{4 + 3x_1(r) - 4x_2(r) - 2x_3(r)} \rightarrow u_5(0) = \frac{7 - \sqrt{19}}{30}. \end{aligned}$$

Because the primal programming problem is a linear programming problem, Wolfe dualization (1.11) produces, independently of x , the familiar dual linear programming problem:

$$(7.1) \quad \begin{aligned} \text{Maximize} \quad & -2u_4 - 4u_5 && \text{subject to} \\ & u_1 - u_4 + 3u_5 = 0, \\ & u_2 - 2u_4 - 4u_5 = -2, \\ & u_3 - u_4 - 2u_5 = -1, \\ & u \geq 0. \end{aligned}$$

The optimality region of this linear programming problem is an edge characterized by the additional constraints $u_2 = u_3 = 0$. Thus $V = \{2, 3\}$, and since the equations in the above problem are equivalent to the gradient expression

$$\nabla f(x^*) = \begin{pmatrix} 0 \\ -2 \\ -1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} u_1 + \begin{pmatrix} -1 \\ -2 \\ -1 \end{pmatrix} u_4 + \begin{pmatrix} 3 \\ -4 \\ -2 \end{pmatrix} u_5, \quad u_1, u_4, u_5 \geq 0,$$

it is verified that the optimal region in question is indeed U^* as defined in (6.4). The analytic center u^* of that dual optimality region is then defined by the fol-

lowing maximization program (compare (6.5))

$$(7.2) \quad \begin{array}{ll} \text{Maximize} & u_1 u_4 u_5 \\ \text{subject to} & \\ & u_1 - u_4 + 3u_5 = 0, \\ & u_2 - 2u_4 - 4u_5 = -2, \\ & u_3 - u_4 - 2u_5 = -1, \\ & u_2 = u_3 = 0, \end{array}$$

which can be treated as an unconstrained problem. For instance, expressing u_1 and u_4 by u_5 , the reader finds $u_1 u_4 u_5 = 10u_5^3 - 7u_5^2 + u_5$. Differentiating yields a quadratic equation solved by $u_5^* = (7 - \sqrt{19})/30$ since $u_1 = 1 - 5u_5 \geq 0$ requires $u_5 \leq 1/5$. Thus $u(0) = u^*$ follows.

The limits u_i^* admit several interpretations. First, they can serve as KKT multipliers in the optimality criterion for the given programming problem (3.1). Second, the subset $u_i^*, i \in Z = \{1, 4, 5\}$, of that full complement of KKT multipliers furnishes KKT multipliers for the optimization problem (3.2).

For general convex programming problems, there are again exceptions. Consider the problem

$$(7.3) \quad \begin{array}{ll} \text{Minimize} & x_2 \\ \text{subject to} & \\ & g_1(x) = x_1 \geq 0, \\ & g_2(x) = x_2 \geq 0, \\ & g_3(x) = x_2 - x_1^2 \geq 0. \end{array}$$

The optimality condition for the minimizer $x = x(r)$ of a barrier perturbation $L(x, r)$ takes the form

$$\begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \frac{r}{x_1} + \begin{pmatrix} 0 \\ 1 \end{pmatrix} \frac{r}{x_2} + \begin{pmatrix} -2x_1 \\ 1 \end{pmatrix} \frac{r}{x_2 - x_1^2}.$$

The primal trajectory is therefore characterized by the equations

$$\frac{1}{x_1} - \frac{2x_1}{x_2 - x_1^2} = 0, \quad \frac{r}{x_2} + \frac{r}{x_2 - x_1^2} = 1.$$

Note that because of $x_2(r) = 3x_1(r)^2$ the primal trajectory is tangential to the constraint plane $x_2 = 0$ at the minimizer $x^* = (0, 0)$. It is easy to express the trajectory coordinates in terms of $r > 0$:

$$r = \frac{6x_1^2}{5} \Rightarrow x_1(r) = \sqrt{\frac{5r}{6}}, \quad x_2(r) = \frac{5r}{2}.$$

As r tends to zero, we thus find the limits

$$\frac{r}{x_1(r)} \rightarrow u_1(0) = 0, \quad \frac{r}{x_2(r)} \rightarrow u_2(0) = \frac{2}{5}, \quad \frac{r}{x_2(r) - x_1(r)^2} \rightarrow u_3(0) = \frac{3}{5}.$$

On the other hand, $M = \{2, 3\}$, $V = \{1\}$, and by (6.4)

$$U^* = \{u : u_2 + u_3 = 1, u_1 = 0\}.$$

Thus the optimization problem that characterizes the dual analytic center (6.5) becomes

$$\begin{aligned} \text{Maximize } & u_2 u_3 \quad \text{subject to} \\ & u_2 + u_3 = 1, \quad u_2, u_3 \geq 0. \end{aligned}$$

The unique solution to this problem is, however, $u_2^* = 1/2 \neq u_2(0)$ and $u_3^* = 1/2 \neq u_3(0)$.

8. Characterizing Dual Trajectories

Portions of the dual trajectory $u = u(r)$ can again be characterized as optimal solutions of a family of auxiliary problems to be formulated below.

We will use the notation

$$\nabla g(x) = [\nabla g_1(x), \dots, \nabla g_m(x)]$$

to indicate the $n \times m$ -matrix whose columns are the gradients of the constraint functions. Index sets are again used in a self explanatory fashion for extracting subvectors, such as

$$g_Z(x^*), \quad u_M(r),$$

or submatrices such as

$$\nabla g_B(x).$$

Index sets are sometimes used as subscripts just to indicate the index range of a vector as, for example, in the definition (8.1) below of the vector $d_C(r)$. Also, reciprocals and logarithms of vectors such as in

$$1/u(r), \quad \ln u(r)$$

are to be understood as vectors of reciprocals and logarithms, respectively, of components.

Let again x^* denote the limit of the primal trajectory $x = x(r)$, which is assumed to exist (6.1), and recall the index set M and its complement V (6.2). Split index set M into two portions B and C , $B \cup C = M$, $B \cap C = \emptyset$, such that $\nabla g_B(x^*)$ is a basis of the matrix ∇g_M .

Let the superscript “+” indicate formation of the Penrose-Moore generalized inverse of a matrix. Because $\nabla g_B(x^*)$ has full column rank,

$$\nabla g_B(x^*)^+ = [\nabla g_B(x^*)^T \nabla g_B(x^*)]^{-1} \nabla g_B(x^*)^T.$$

Since $\nabla g_B(x^*)$ is the limit of $\nabla g_B[x(r)]$ as $r \rightarrow 0$, the latter has full rank for sufficiently small values of $r > 0$. For such values of r we define the vector

$$(8.1) \quad d_C(r) = \frac{1}{u_C(r)} - \nabla g_C[x(r)]^T \nabla g_B[x(r)]^+ \frac{1}{u_B(r)},$$

which will set up a correction term for the problem below, where $u(r)$ is the dual trajectory as defined by (1.10) in the Introduction.

(8.2) Problem Q_r : Maximize $\ln u_B + \ln u_C - d_C(r)^T u_C$ subject to

$$[u_B, u_C] > 0,$$

$$\nabla f[x(r)] - \nabla g_B[x(r)]u_B - \nabla g_C[x(r)]u_C - \nabla g_V[x(r)]u_V(r) = 0,$$

where the index sets B, C partition the index set M .

(8.3) The vector $[u_B(r), u_C(r)]$ solves, for all sufficiently small $r > 0$, the convex maximization problem Q_r .

Proof. By its definition, the concatenated vector $[u_B(r), u_C(r)]$ is positive and satisfies the equation constraints of problem Q_r . It is therefore feasible. For it to be also optimal, it suffices that there exist a vector of Lagrange multipliers $w(r)$ such that

$$(8.4) \quad \begin{bmatrix} -1/u_B(r) \\ -1/u_C(r) + d_C(r) \end{bmatrix} = \begin{bmatrix} \nabla g_B[x(r)]^T \\ \nabla g_C[x(r)]^T \end{bmatrix} w(r).$$

Such a vector $w(r)$ indeed exists:

$$w(r) = -\nabla g_B[x(r)]^+ \frac{1}{u_B(r)}. \quad \square$$

Several other important facts about the dual trajectory $u = u(r)$ can be ascertained under general conditions following McCormick [14].

(8.5) *the dual trajectory $u = u(r)$ remains bounded, as $r > 0$ tends to zero.*

Proof: Let x^* be the limit of the primal trajectory $x(r)$ with respect to which the dual trajectory is defined. Assume that $u(r)$ is not bounded. Then there exists a sequence $r_j > 0$, $j = 1, \dots$, which converges to zero such that $\|u(r_j)\| \rightarrow \infty$. The sequence of unit vectors

$$\frac{u(r_j)}{\|u(r_j)\|}$$

has a nonzero accumulation point $v \geq 0$, $\|v\| = 1$, which we may assume to be a limit. (Otherwise we would select a suitable subsequence.) Passing to the limit of the relation

$$\frac{1}{\|u(r_j)\|} \nabla f[x(r_j)] = \sum_{i=1}^m \nabla g_i[x(r_j)] \frac{u_i(r_j)}{\|u(r_j)\|},$$

we find by continuity

$$0 = \sum_{i=1}^m \nabla g_i(x^*) v_i.$$

At least one component v_h is positive. Thus

$$-\nabla g_h(x^*) = \sum_{i \neq h} \nabla g_i(x^*) \frac{v_i}{v_h}.$$

The ratios $v_i/v_h \geq 0$ can be interpreted as KKT multipliers for the problem of maximizing the function $g_h(x)$ subject to the remaining constraints $g_i(x) \geq 0$, $i \neq h$. In addition, complementary slackness holds since $g_i(x^*) > 0$ implies for $r \rightarrow 0$

$$u_i(r) = \frac{r}{g_i[x(r)]} \rightarrow 0$$

and therefore $v_i = 0$. We also note that for the same reason $g_h(x^*) = 0$. Since all functions involved are concave, the KKT conditions suffice to ascertain that $g_h(x^*) = 0$ is the maximum value of $g_h(x) \geq 0$ subject to the remaining constraints and, particularly, in the feasible region X^F . This contradicts the Slater constraint

qualification (1.4). \square

Proposition (8.5) assures the existence of accumulation points \bar{u} of $u(r)$ as $r > 0$ tends to zero. Moreover, convergence holds if all these accumulation points can be shown to be equal. The next proposition is a step in this direction.

(8.6) *If \bar{u} is an accumulation point of $u(r)$ as $r > 0$ tends to zero, then*

$$\bar{u}_i = 0 \quad \text{for } i \in V,$$

where x^ is the limit of $x(r)$.*

Proof: In the previous proof, $g_i(x^*) > 0$ was seen to imply $u_i(r) \rightarrow 0$. Hence $\bar{u}_i = 0$ if $g_i(x^*) > 0$, again establishing complementary slackness. Thus (x^*, \bar{u}) is an optimal solution to the Wolfe dualization (1.11). The proposition then follows by the definition (6.2) of the index set V . \square

If the vectors $d_C(r)$ and $u_V(r)$ were to converge to zero as $x(r)$ converges to x^* , then the maximization problems Q_r (8.2) would tend to the maximization problem (6.5), which defines the analytic center of the dual optimality region. This observation defines our approach to proving this analytic center to be the limit of the dual trajectory. We also note, that

(8.7) *the analytic center of the dual optimality region exists and is unique.*

Proof: For existence, it has to be shown that the feasible region of the the maximization problem (6.5) is bounded, which follows by an argument analogous to the proof of proposition (8.5). Uniqueness follows from the fact that the logarithmic objective function is concave as stated earlier in (6.8). \square

9. Limits of Dual Trajectories

The characterization of the dual trajectory limit as an analytic center requires additional assumptions such as the following.

- (9.1) (i) *The constraint functions $g_i(x)$, $i = 1, \dots, m$, are linear.*
(ii) *If x^* is the primal trajectory limit, then*

$$\frac{x(r) - x^*}{r} \leq \text{constant},$$

that is, the above vector remains bounded as $r > 0$ tends to zero.

Note that example (7.3) violates those assumptions. We are now able to state our second theorem.

- (9.2) **Theorem 2:** *If the primal trajectory $x = x(r)$ converges to a unique optimal solution $x^* \in X^*$, and if one of the assumptions (9.1.i,ii) holds then the dual trajectory $u = u(r)$ converges to the analytic center u^* of the dual optimality region U^* .*

The proof of this theorem will be based on the properties of the auxiliary maximization problem Q_r (8.2), which by (8.3) characterizes trajectory points $u = u(r)$. Indeed, we will show that

- (9.3) *Theorem 2 holds if*
(i) *the vector*

$$d_C(r) = \frac{1}{u_C(r)} - \nabla g_C[x(r)]^T \nabla g_B[x(r)]^{+T} \frac{1}{u_B(r)},$$

in problem Q_r vanishes as $r > 0$ tends to zero;
and (ii) for each accumulation point \bar{u} of the dual trajectory $u(r)$, the components \bar{u}_i with $i \in M$ are positive.

Proof: We first note that u_V converges to zero by (8.6). Therefore we need to concern ourselves only with

$$u_M(r) = [u_B(r), u_C(r)].$$

By (8.5), the trajectory $u(r)$ possesses an accumulation point \bar{u} , whence there exists a sequence $r_j > 0$, $j = 1, \dots$, such that

$$u(r_j) \rightarrow \bar{u} \quad \text{for } j \rightarrow \infty.$$

By its definition $u(r_j) = [u_B(r_j), u_C(r_j), u_V(r_j)]$ satisfies

$$\nabla f[x(r_j)] - \nabla g_M[x(r_j)]u_M(r_j) - \nabla g_V[x(r_j)]u_V(r_j) = 0.$$

By the continuity of the gradients and again by (8.3), we find in the limit $r_j \rightarrow 0$ that

$$\nabla f(x^*) - \nabla g_M(x^*)\bar{u}_M = 0.$$

Thus \bar{u}_M is a feasible point of problem (6.5) and therefore, a candidate for analytic center. We proceed to show that \bar{u}_M is an optimal point. This requires that there exist a vector of Lagrange multipliers $w^* \in R^n$ such that

$$(9.4) \quad -\frac{1}{\bar{u}_M} = \nabla g_M(x^*)^T w^*.$$

Note that by (8.3) $[u_B(r), u_C(r)] = u_M(r)$ is the maximizer for problem Q_r : its optimality condition (8.4) is satisfied because for the vector of Lagrange multipliers

$$w(r) = -\nabla g_B[x(r)]^T \frac{1}{u_C(r)}$$

the relation

$$\begin{bmatrix} -1/u_B(r) \\ -1/u_C(r) + d_C(r) \end{bmatrix} = \begin{bmatrix} \nabla g_B[x(r)]^T \\ \nabla g_C[x(r)]^T \end{bmatrix} w(r)$$

holds. By hypotheses, $w(r_j)$ has the limit

$$w^* = -\frac{1}{\bar{u}_C}$$

and by continuity $\bar{u}_M = (\bar{u}_B, \bar{u}_C)$ must satisfy (9.4). Thus \bar{u}_M is an analytic center of the dual optimality region U^* . Since this holds for every accumulation point, and since the analytic center of U^* is unique by (8.7), the proposition follows. \square

We will now proceed with the

Proof of Theorem 2: Suppose (9.1.i) holds. Then, denoting by $g(x)$ the m vector of constraints $g_i(x)$,

$$g(x) = A^T x - b, \quad \nabla g(x) = A,$$

for some m vector b and some $n \times m$ matrix A . In what follows, index sets used as subscripts again denote corresponding subvectors and submatrices.

In particular, the submatrix A_B has full rank so that $A_B^+ A_B = I$ for its Moore-Penrose pseudo inverse $A_B^+ = (A_B^T A_B)^{-1} A_B^T$. By the choice of submatrix A_B , the

columns of submatrix A_C are linear combinations of the columns of A_B : $A_C = A_B A_B^+ A_C$. By (6.6) $g_M(x^*) = 0$. Thus $A_M^T x^* = b_M$ and, in particular,

$$(9.5) \quad b_C = A_C^T x^* = A_C^T A_B^{+T} A_B^T x^* = A_C^T A_B^{+T} b_B.$$

Now

$$\begin{aligned} d_C(r) &= \frac{1}{u_C(r)} - \nabla g_C[x(r)]^T \nabla g_B[r]^{+T} \frac{1}{u_B(r)} \\ &= \frac{1}{r} \left(g_C[x(r)] - \nabla g_C[x(r)]^T \nabla g_B[x(r)]^{+T} g_B[x(r)] \right) \\ &= \frac{1}{r} \left([A_C^T x(r) - b_C] - A_C^T A_B^{+T} [A_B^T x(r) - b_B] \right) \\ &= -\frac{1}{r} \left(b_C - A_C^T A_B^{+T} b_B \right) \\ &= 0 \end{aligned}$$

for all $r > 0$ by (9.5). This establishes condition (9.3.i). To verify condition (9.3.ii), choose an accumulation point \bar{u} and a sequence $r_j > 0$, $j = 1, \dots$, which converges to zero such that the resulting sequence $u(r_j)$ converges to \bar{u} . The limit x^* of the primal trajectory $x(r)$ is an optimal solution $x^* \in X^*$. Hence there exists, by (6.6), a point $\hat{u} \in U^*$ such that $\hat{u}_M > 0$ and $\hat{u}_V = 0$ hold in conjunction with $g_M(x^*) = 0$. Since

$$\nabla f(x^*) - \nabla g(x^*) \hat{u} = \nabla f(x^*) - \nabla g_M \hat{u}_M = \nabla f(x^*) - A_M \hat{u}_M = 0, \quad \hat{u}_M > 0,$$

the point \hat{u}_M is feasible for problem (6.5) and is thus a potential analytic center. Define

$$\tilde{u}_M(r_j) = u_M(r_j) + \hat{u}_M - \bar{u}_M.$$

Since $\nabla f(x^*) = \nabla g_M(x^*)^T u_M = A_M^T u_M$ holds for both \bar{u}_M and \hat{u}_M , the vector $u_M = \tilde{u}_M(r_j)$ satisfies

$$\begin{aligned} \nabla f[x(r_j)] - \nabla g_M[x(r_j)] u_M - \nabla g_V[x(r_j)] u_V(r_j) \\ = \nabla f[x(r_j)] - A_M u_M - A_V u_V(r_j) = 0. \end{aligned}$$

$\tilde{u}_M(r_j)$ converges to $\hat{u}_M > 0$. Hence we may assume $\tilde{u}_M(r_j) \geq 0$ (or otherwise select a suitable subsequence). $\tilde{u}_M(r_j)$ is therefore a feasible solution to problem \mathbf{Q}_{r_j} . If the limit \bar{u}_i of any component of $u_M(r_j)$ equals zero, the objective function values of $u_M(r_j)$ for problems \mathbf{Q}_{r_j} will go to infinity, and $\tilde{u}_M(r)$ will eventually

produce smaller values than $u_M(r_j)$, contradicting the assumption that the latter is a minimizer. Thus condition (9.3.ii) thus holds also, and Theorem 2 follows under assumption (9.1.i). We remark that the above argument could have been based on a general result of Witzgall, Boggs, and Domich [23]: the analytic centers of linear constraint systems are continuous with respect to changes of their right hand sides.

Suppose assumption (9.1.ii) holds, and recall that $g_M(x^*) = 0$ by (6.6). For $r > 0$ by Taylor's Theorem:

$$(9.6) \quad \frac{1}{u_M(r)} = \frac{g_M[x(r)]}{r} = \nabla g_M[\hat{x}(r)]^T \frac{x(r) - y^*}{r}$$

where $\hat{x}(r)$ is a point on the line segment $[x^*, x(r)]$. We will first show that the vector $d_C(r)$ in that problem converges to zero as required by condition (9.3.ii). By (9.6), and since $B \cup C$ is a partition of M , vector $d_C(r)$ is of the form

$$(9.7) \quad d_C(r) = D_C(r) \frac{x(r) - x^*}{r}$$

with the matrix $D_C(r)$ given by

$$D_C(r) = \nabla g_C[\hat{x}(r)]^T - \nabla g_C[x(r)]^T \nabla g_B[x(r)]^{+T} g_B[\hat{x}(r)]^T$$

Here $r > 0$ must be sufficiently small for the matrix $\nabla g_B[x(r)]$ to have full rank. It then follows that its pseudoinverse converges to $g_B(x^*)^+$. Clearly, $\hat{x}(r)$ converges to x^* - together with $x(r)$ - as $r > 0$ tends to zero. It follows by continuity that

$$D_C(r) \rightarrow D_C^* = \nabla g_C(x^*)^T [I - \nabla g_B(x^*)^{+T} \nabla g_B(x^*)].$$

Recall that the columns of $\nabla g_C(x^*)$ lie in the column space of $\nabla g_B(x^*)$:

$$\nabla g_C(x^*) = \nabla g_B(x^*) \nabla g_B(x^*)^+ \nabla g_C(x^*).$$

Hence

$$D_C^* = \nabla g_C(x^*)^T \nabla g_B(x^*)^{+T} \nabla g_B(x^*)^T [I - \nabla g_B(x^*)^{+T} \nabla g_B(x^*)^T] = 0.$$

Thus assumption (9.1.ii) implies $d_C(r) \rightarrow 0$ in view of (9.7).

To prove the remaining condition (9.3.ii), we assume again that accumulation point \bar{u} is the limit of $u(r_j)$, $r_j > 0$, $j = 1, \dots$, as $r_j \rightarrow 0$. Since

$$\frac{x(r_j) - x^*}{r_j}$$

remains bounded for $r_j \rightarrow 0$ by (9.1.ii), we may assume that it, too, converges (or else select a suitable subsequence of the above). By (9.6),

$$\frac{1}{u_M(r_j)} = \frac{g_M[x(r_j)]}{r_j} = \nabla g_M[\hat{x}(r_j)]^T \frac{x(r_j) - x^*}{r_j}.$$

The reciprocal of $u_M(r)$ converges since $\nabla g_M[\hat{x}(r_j)]$ converges by continuity for $j \rightarrow \infty$. But this is possible only if the limit \bar{u}_M is positive. The conditions of (9.3) are thus met, and Theorem 2 is proved. (Note that (9.1.ii) implies $M \supseteq Z$. If, furthermore, the primal trajectory limit x^* is a strict minimizer, that is, if $g_i(x^*) = 0$ only if $i \in Z$, then $M = Z$ by (6.6).) \square

Acknowledgement

The authors are indebted to A. J. Goldman and A. J. Kearsley for suggestions and helpful comments. J. F. Lawrence carefully read the latest manuscript and suggested some extensions of an earlier rank-integrity concept. We also thank the anonymous referees who commented on the first submission of this paper six years ago.

References

- [1] ADLER, I., AND MONTEIRO, R. D. C. The limiting behavior of the affine scaling trajectories for linear programming problems. Technical Report ESRC 88-9, Engineering Systems Research Center, University of California at Berkeley, April 1988.
- [2] ANSTREICHER, K. M. Linear programming and the Newton barrier flow. *Math. Programming* 41 (1988), 367–373.
- [3] BAYER, D. A., AND LAGARIAS, J. C. The nonlinear geometry of linear programming: I. Affine and projective scaling trajectories. *Trans. Amer. Math. Soc.* 314 (1986), 499–526.
- [4] BAYER, D. A., AND LAGARIAS, J. C. The nonlinear geometry of linear programming: II. Legendre Transform coordinates and central trajectories. *Trans. Amer. Math. Soc.* 314 (1989), 527–581.
- [5] FIACCO, A. V., AND MCCORMICK, G. P. Programming under nonlinear constraints by unconstrained minimization: A primal-dual method. Research Paper RAC-TP-96, The Research Analysis Corporation, 1963.

- [6] FIACCO, A. V., AND MCCORMICK, G. P. Computational algorithm for the Sequential Unconstrained Minimization Technique for nonlinear programming. *Management Sci.* 10 (1964), 601–617.
- [7] FIACCO, A. V., AND MCCORMICK, G. P. The Sequential Unconstrained Minimization Technique for nonlinear programming: A primal-dual method. *Management Sci.* 10 (1964), 360–366.
- [8] FIACCO, A. V., AND MCCORMICK, G. P. *Nonlinear Programming: Sequential Unconstrained Minimization Techniques*, vol. 4 of *Classics in Applied Mathematics*. Society for Industrial and Applied Mathematics, Philadelphia, 1990.
- [9] FRISCH, K. R. The logarithmic potential method of convex programming. Technical report, University Institute of Economics, Oslo, Norway, 1955.
- [10] HUARD, P. Résolution des p. m. à contraintes non linéaires par la méthode des centres. Tech. Rep. Note E. D. F. no. HR 5.690, Électricité De France, May 1964.
- [11] HUARD, P. Resolution of mathematical programming with nonlinear constraints by the method of centres. In *Nonlinear Programming*, J. Abadie, Ed. North Holland, Amsterdam, 1967, pp. 209–219.
- [12] KARMAKAR, N. A new polynomial-time algorithm for linear programming. *Combinatorica* 4 (1984), 373–395.
- [13] MCCORMICK, G. P. *Nonlinear Programming: Theory, Algorithms and Applications*. John Wiley and Sons, Inc., New York, 1983.
- [14] MCCORMICK, G. P. The projective SUMT method for convex programming. *Math. of Operations Research* 14 (1989), 203–223.
- [15] MEGIDDO, N. Pathways to the optimal set in linear programming. In *Progress in Mathematical Programming: Interior-point Algorithms and Related Methods* (N. Megiddo, ed.); *Papers from the Conference held in Pacific Grove, California, March 1-4, 1987*. Springer-Verlag, New York, 1989, pp. 131–158.
- [16] NESTEROV, YU. E., AND NEMIROVSKII, A. S. Self-concordant functions and polynomial time methods in convex programming. Monograph, Central Economical and Mathematical Institute, USSR Academy of Science, Moscow, 1990.

- [17] NESTEROV, YU. E., AND NEMIROVSKII, A. S. *Interior-point polynomial algorithms in convex programming*, vol. 13 of *SIAM Studies in Applied Mathematics*. Society for industrial and Applied Mathematics (SIAM), 1994.
- [18] RENEGAR, J. A polynomial-time algorithm, based on Newton's method, for linear programming. Report 07118-86, Mathematical Sciences Research Institute, University of California at Berkeley, June 1986.
- [19] RENEGAR, J. A polynomial-time algorithm, based on Newton's method, for linear programming. *Math. Programming* 40 (1988), 59–93.
- [20] SLATER, M. L. Lagrange multipliers revisited: A contribution to nonlinear programming. Discussion Paper Math 403, Cowles Commission, November 1950.
- [21] SONNEVEND, G. An “analytic center” for polyhedrons and new classes of global algorithms for linear (smooth, convex) programming. In *Proc. 12th IFIP Conference on System Modeling and Optimization* (Budapest, 1985).
- [22] STOER, J., AND WITZGALL, C. *Convexity and Optimization in Finite Dimensions*, vol. 163 of *Die Grundlehren der Mathematischen Wissenschaften*. Springer Verlag, Berlin, 1970.
- [23] WITZGALL, C., BOGGS, P. T., AND DOMICH, P. D. On the convergence behavior of trajectories for linear programming. *Contemporary Mathematics* 114 (1990), 161–187. *Proceedings of the 1988 AMS-IME-SIAM Research Conference on “Mathematical Developments Arising from Linear Programming Algorithms”*.
- [24] WOLFE, P. A duality theorem for nonlinear programming. *Quart. Appl. Math.* 19 (1961), 239–244.

