MATHEMATICS OF OPERATIONS RESEARCH
Vol 17, No. 1, February 1992
Printed in U S A

ON THE CONTINUOUS TRAJECTORIES FOR A POTENTIAL REDUCTION ALGORITHM FOR LINEAR PROGRAMMING*

RENATO D. C. MONTEIRO

For a possibly degenerate linear program $\min_x\{c^Tx; Ax = b, x \geqslant 0\}$ (A is an $m \times n$ real matrix, $b \in \mathbf{R}^m$ and $c \in \mathbf{R}^n$), whose optimal value is 0, we study the limiting behavior of the trajectories of the family of vector fields

$$\Phi_q(x) \equiv -XP_{AX}\left[Xc - \frac{c^Tx}{q}\mathbf{1}\right],$$

for all values of $q \geqslant n$, where X is the diagonal matrix associated with x and P_{AX} is the projection operator onto the null space of AX. A polynomial algorithm based on the directions $\Phi_q(x)$ has been presented by Gonzaga [6] when $q = n$ or $q = n + \sqrt{n}$. We show that all trajectories of $\Phi_q(x)$ converge to the unique "center" of the optimal face of the given linear program. When this face consists of a unique vertex, it is shown that any trajectory of $\Phi_q(x)$ approaches this vertex along the same direction. When the optimal face consists of more than one point, we show that there is a threshold value $\tau > 0$ such that: for $q > \tau$, "most" of the trajectories of $\Phi_q(x)$ converge to the "center" tangentially to the optimal face and that the direction of approach of a trajectory of $\Phi_q(x)$ depends on the initial condition; for $q < \tau$ ($q = \tau$), the trajectories of $\Phi_q(x)$ converge to the "center" along a unique direction (along several directions which depend on the initial condition) forming a positive angle with the optimal face.

1. Introduction. Continuous trajectories related to the area of optimization have been discussed long ago by the nonlinear programming community (see, for example, Fiacco and McCormick [5]). Recently, with the introduction by Karmarkar of a new interior point algorithm for solving linear programming problems in his seminal paper [8], a new effort appeared toward studying the continuous trajectories related to this algorithm, usually called the projective scaling algorithm. Another algorithm, namely, the affine scaling algorithm, was introduced by Dikin [4] long before Karmarkar's work but Dikin's result remained unknown in the west until recently (1988). Before then, the affine scaling algorithm was (independently) rediscovered by Barnes [2] and Vanderbei, Meketon and Freedman [15]. Continuous trajectories related to these interior point algorithms have been analyzed by several authors. Bayer and Lagarias [3] present a systematic study of some properties and mathematical structures of these trajectories. Megiddo and Shub [10] and Shub [14] analyze the limiting behavior of these trajectories near the optimal vertex under primal and dual nondegeneracy assumptions. Using a different approach, Adler and Monteiro [1], Monteiro [11], Witzgall, Boggs and Domich [16] further analyze the convergence and limiting behavior of these continuous trajectories without imposing any nondegeneracy assumptions. Another effort toward analyzing continuous trajectories for linear pro-

*Received September 26, 1989; revised May 20, 1990.
AMS 1980 subject classification. Primary: 90C05. Secondary 90C25.
IAOR 1973 subject classification. Main: Programming: Linear.
OR/MS Index 1978 subject classification. Primary: 643 Programming/Linear/Algorithms.
Key words Interior point algorithms, linear programming, Karmarkar's algorithm, continuous trajectories for linear programming, potential reduction algorithms, potential function.

0364-765X/92/1701/0225/$01.25

grams was the work of Megiddo [10] which studies the weighted barrier trajectories in the framework of primal-dual complementarity relationship for the special case of linear and convex quadratic programming problems. A polynomial algorithm, namely, the primal-dual affine scaling algorithm, which has the weighted barrier trajectories as its continuous trajectories, was presented and analyzed by Monteiro, Adler and Resende [12].

In Gonzaga [6], a polynomial potential reduction (PR) algorithm has been developed for solving linear programming problems in the form

$$\text{(P)} \qquad r^* = \min_x \{c^T x; Ax = b, x \geqslant 0\},$$

where A is a real $m \times n$-matrix and b, c are real vectors of length m and n, respectively, and for which the following assumptions are imposed: $r^* = 0$ and the set $S_I = \{x; Ax = b, x > 0\}$ is nonempty and bounded. To each point $x \in S_I$, the PR algorithm assigns the direction given as

$$\Phi_q(x) \equiv -XP_{AX}\left[Xc - \frac{c^T x}{q}\mathbf{1}\right],$$

where P_{AX} denotes the projection operator onto the subspace $\{h; AXh = 0\}$ and where the constant q is either considered to be equal to n or $n + \sqrt{n}$. It has been shown in [6] that, when $q = n$ ($q = n + \sqrt{n}$), the PR algorithm solves (P) in at most $O(n^2L)$ ($O(nL)$) iterations where L denotes the number of bits necessary to represent the data of problem (P).

The objective of this paper is to analyze the continuous trajectories associated with the PR algorithm, that is, those trajectories induced by the vector field $\Phi_q(\cdot)$ under the assumptions above and for any value of $q \geqslant n$. The term "PR trajectories" will be used throughout this paper to refer to these trajectories. We show that all PR trajectories converge to the unique "center" of the optimal face of problem (P). When this face consists of a unique vertex, it is shown that any PR trajectory approaches this vertex along the same asymptotic direction of convergence (which is, by definition, the limit of the unit vectors tangent to the trajectory, as the trajectory approaches the optimal face). In [1], the affine scaling trajectories are characterized as the set of points traced by the optimal solutions of certain parametrized convex optimization problems. This same characterization is proved for the PR trajectories and is then used to derive the results above. However, such characterization for the PR trajectories turns out to be much more involved than the one for the affine scaling trajectories. Moreover, when the optimal face has dimension greater than zero, we obtain surprising results, at least to the author, about the way these trajectories may approach the optimal face. In this case, the asymptotic direction of convergence may vary according to whether $q > 2|N|$, $q < 2|N|$ or $q = 2|N|$, where the set $N \subseteq \{1, \ldots, n\}$ is such that $(x_j \geqslant 0, j \in N)$ is the set of always-active constraints for the optimal face of problem (P). In the first case, we show that "most" of the PR trajectories approach the center of the optimal face tangentially to this face and that their asymptotic direction of convergence depend on the initial condition. In the second case, the trajectories behave like in the case of a unique optimal vertex, that is, they approach this center along the same asymptotic direction of convergence. (This unique direction is the asymptotic direction of convergence corresponding to the primal "central trajectory".) In the third case, the PR trajectories approach the center of the optimal face along many possible asymptotic directions of convergence, which depend on the initial condition, but all these directions form a positive angle with the optimal face.

Our paper is organized as follows. In §2, we present some basic facts of ordinary differential equations and then use them to prove some preliminary results about the PR trajectories. In §3, we analyze the limiting behavior of the PR trajectories. In this section, dual estimates, analogous to the dual estimates used in the context of the affine scaling algorithm, are defined and then used, in §4, to analyze the asymptotic direction of convergence of the PR trajectories.

2. Problem description and preliminary results. In this section, we present some basic facts of the theory of ordinary differential equations and use them to prove some preliminary results about the PR trajectories. We also describe the PR algorithm in more detail, set the assumptions that will be used throughout the paper (unless otherwise stated) and introduce the terminology necessary for our development.

2.1. *Some facts about ordinary differential equations.* In this subsection, we start by discussing some basic results from the theory of ordinary differential equations (ODE) that will be useful later.

Let E be an affine space of $\mathbf{R}^n$ and let $T(E)$ denote the tangent space to E, that is, the subspace obtained by translating E to the origin of $\mathbf{R}^n$. Let $g\colon (\alpha, \beta) \to \mathbf{R}$ and $\Phi\colon U \to T(E)$ be C^∞-functions where α and β are numbers in $\mathbf{R} \cup \{-\infty, \infty\}$ such that $\alpha < \beta$ and $U \subseteq E$ is an open subset of E (with the topology induced as a subset of $\mathbf{R}^n$). In this paper, we consider only differential equations of the form

$$\dot{x} = g(t)\Phi(x). \tag{2.1}$$

The set $U \times (\alpha, \beta)$ is called the *domain of definition* of (2.1). The ODE (2.1) is called an *autonomous* ODE if its domain of definition is the set $U \times \mathbf{R}$ and $g(t) = 1$ for all $t \in \mathbf{R}$. When the ODE (2.1) is autonomous, we say that $\Phi\colon U \to T(E)$ is the *vector field* associated with (2.1). A *solution* of (2.1) is a differentiable path $x\colon I \to U$ such that $\dot{x}(t) = g(t)\Phi(x(t))$ for all t in the open interval $I \subseteq (\alpha, \beta)$. The solution $x\colon I \to U$ of (2.1) is called a *maximal solution* if it cannot be further extended, that is, any other solution $y\colon J \to U$ of (2.1) with the property that $J \supseteq I$ and $y(t) = x(t)$ for all $t \in I$ satisfies $I = J$. The set $x(I)$ is then referred to as the *trajectory* determined by the maximal solution $x\colon I \to U$ of (2.1). Given $t^0 \in (\alpha, \beta)$ and $x^0 \in U$, the ODE (2.1) together with the initial condition $x(t^0) = x^0$ defines the problem

$$\dot{x} = g(t)\Phi(x), \qquad x(t^0) = x^0, \tag{2.2}$$

which is called a *Cauchy problem*. It is well known that the Cauchy problem (2.2) has a unique maximal solution. Also, the following result is well-known in the context of ODE theory (see, for example, Hirsch and Smale [7, p. 171]).

PROPOSITION 2.1. *Let $x\colon (w^-, w^+) \to U$ be the maximal solution of the Cauchy problem*

$$\dot{x} = \Phi(x), \qquad x(t^0) = x^0. \tag{2.3}$$

If $w^+ < \infty$ (respectively, $w^- > -\infty$) then given any compact set $K \subseteq U$, there exists an $\epsilon > 0$ such that $x(t) \notin K$ for all $t \in [w^+ - \epsilon, w^+)$ (respectively, $t \in (w^-, w^- + \epsilon]$).

There is also a version of Proposition 2.1 that holds for the more general Cauchy problem (2.2) but the above result is all we need. The assumption that w^+ be finite in Proposition 2.1 is quite unpleasant for our purposes. The next result shows that we can drop this assumption at the cost of adding extra conditions on the ODE $\dot{x} = \Phi(x)$.

PROPOSITION 2.2. *Assume that there exists a continuous function $W: U \to \mathbf{R}$ such that for any solution $\psi(t)$ of $\dot{x} = \Phi(x)$, $W(\psi(t))$ is a (strictly) decreasing function of t. Then any maximal solution $x: (w^-, w^+) \to U$ of $\dot{x} = \Phi(x)$ satisfies: given any compact set $K \subseteq U$, there exists a number $\beta \in (w^-, w^+)$ such that $x(t) \notin K$ for all $t \in [\beta, w^+)$.*

A proof of Proposition 2.2 is given in the Appendix (§5). The following results will be useful later.

PROPOSITION 2.3. *Let $\psi: (\alpha^-, \alpha^+) \to U$ and $x: (w^-, w^+) \to U$ denote the maximal solutions of the Cauchy problems* (2.2) *and* (2.3), *respectively. Assume that $g(t) > 0$ for all $t \in (\alpha, \beta)$ and let $\gamma: (\alpha, \beta) \to \mathbf{R}$ be the function defined by $\gamma(t) \equiv t^0 + \int_{t^0}^{t} g(s)\, ds$ for all $t \in (\alpha, \beta)$. Then we have*

(a) $(\alpha^-, \alpha^+) = \{t \in (\alpha, \beta); w^- < \gamma(t) < w^+\}$,

(b) $\psi(t) = x(\gamma(t))$ *for all* $t \in (\alpha^-, \alpha^+)$.

We observe that both α^+ and w^+ (respectively, α^- and w^-) in the statement of Proposition 2.3 are allowed to be $+\infty$ (respectively, $-\infty$). A more general result than Proposition 2.3, which does not assume that $g(t) > 0$ over (α, β), is stated and proved in the Appendix (§5). Note that statement (a) implies that the trajectory determined by the maximal solution of (2.2) is always a subset of the trajectory determined by the maximal solution of (2.3) and that the first trajectory is a proper subset of the latter one if either

$$\inf_{t \in (\alpha, \beta)} \gamma(t) > w^- \quad \text{or} \quad \sup_{t \in (\alpha, \beta)} \gamma(t) < w^+.$$

PROPOSITION 2.4. *Assume that u and v are continuous functions defined in the interval $[\alpha, \beta] \subseteq \mathbf{R}$. If, for some $d \in \mathbf{R}$, the functions u and v satisfy the relation*

$$u(t) = d + \int_{\alpha}^{t} u(s) v(s)\, ds \quad \text{for all } t \in [\alpha, \beta], \tag{2.4}$$

then the following relation holds:

$$u(t) = d \exp\left[\int_{\alpha}^{t} v(s)\, ds\right] \quad \text{for all } t \in [\alpha, \beta]. \tag{2.5}$$

PROOF. Just observe that the functions (over the interval $[\alpha, \beta]$) defined by the right-hand sides of relations (2.4) and (2.5) are solutions of the Cauchy problem $\dot{u} = v(t)u$, $u(\alpha) = d$. Hence, they must be equal. □

2.2. *Problem description.* In this subsection, we describe the problem which will be the subject of our study in this paper. We also state the assumptions that will be used throughout this paper (unless otherwise stated). In Gonzaga [6], a potential reduction algorithm for solving linear programming problems was described which converges in polynomial time. We are interested in studying the behavior of the continuous trajectories for the vector field which is naturally associated with this algorithm. The potential reduction algorithm is briefly reviewed in this subsection.

We start by introducing some notation and terminology. Consider the linear programming problem in standard form

$$\text{(P)} \qquad \min_x \{c^T x;\ Ax = b,\ x \geqslant 0\},$$

and its dual problem

$$\text{(D)} \qquad \max_{y,z} \{b^T y;\ A^T y + z = c,\ z \geqslant 0\},$$

where A is an $m \times n$-matrix and b, c are vectors of length m and n, respectively. The following notation will be used throughout this paper. Let r^* denote the optimal value of problems (P) and (D). Let

$$S_A = \{x \in \mathbf{R}^n;\ Ax = b\},$$

$$S_F = \{x;\ x \in S_A,\ x \geqslant 0\},$$

$$S_I = \{x;\ x \in S_A,\ x > 0\},$$

$$S_O = \{x;\ x \in S_F,\ c^T x = r^*\},$$

$$T_A = \{(y, z) \in \mathbf{R}^m \times \mathbf{R}^n;\ A^T y + z = c\},$$

$$T_F = \{(y, z);\ (y, z) \in T_A,\ z \geqslant 0\},$$

$$T_I = \{(y, z);\ (y, z) \in T_A,\ z > 0\},$$

$$T_O = \{(y, z);\ (y, z) \in T_F,\ b^T y = r^*\}.$$

The sets S_F and T_F are the feasible sets of problems (P) and (D), respectively, S_A and T_A are the affine hulls of S_F and T_F, respectively, S_I and T_I are the relative interior of S_F and T_F, respectively, and S_0 and T_0 are the optimal faces of problems (P) and (D), respectively. Also $\mathbf{1}$ will denote the vector of all ones whose dimension is dictated by the appropriate context. If x is a lower case letter that denotes a vector $x = (x_1, \ldots, x_n)^T$, then a capital letter will denote the diagonal matrix with the components of the vector on the diagonal, i.e., $X = \mathrm{diag}(x_1, \ldots, x_n)$. Also $\mathbf{R}^n_+$ will denote the set of n-vectors with all components strictly positive. If x and z are two n-vectors, we define their product xz to be the vector $XZ\mathbf{1} = (x_1 z_1, \ldots, x_n z_n)^T$. The inverse of x under this operation is denoted by x^{-1} and is given by $x^{-1} \equiv X^{-1}\mathbf{1}$. In this way, expressions like x^{-1} and $x^{-2}z$ are defined if all the components of x are nonzero. No confusion should arise between the expressions xz and $x^T z$ where the latter just denotes the inner product of x and z. Given an $m \times n$ matrix H and a subset B of the index set $\{1, \ldots, n\}$, we denote by H_B the submatrix of H associated with the index set B. Moreover, the subspace generated by the columns of H will be denoted by range(H). The projection operator onto the subspace $\{z \in \mathbf{R}^n;\ Hz = 0\}$ will be denoted by P_H. If H has full row rank then P_H is given by

$$P_H \equiv \left[I - H^T (HH^T)^{-1} H\right].$$

Another terminology that will be used throughout the paper is the following. For $x \in \mathbf{R}^n$, let $\text{supp}(x) = \{j \in \{1,\dots,n\};\ x_j \neq 0\}$. We define two index sets B and N as follows:

$$B = \bigcup_{x \in S_O} \text{supp}(x), \qquad N = \bigcup_{(y,z) \in T_O} \text{supp}(z). \tag{2.6}$$

It is well known (see, for example, Schrijver [13]) that $B \cup N = \{1,\dots,n\}$ and $B \cap N = \varnothing$. Moreover, the optimal faces S_O and T_O can be expressed in terms of the index sets B and N as follows:

$$S_O = \{x \in S_F;\ x_N = 0\}, \tag{2.7}$$

$$T_O = \{(y,z) \in T_F;\ z_B = 0\}. \tag{2.8}$$

Recall that r^* denotes the optimal value of problems (P) and (D). We make the following assumptions about problem (P):

ASSUMPTION 2.1. (a) $S_I \neq \varnothing$;
(b) $r^* = 0$ *and* $c^T x > 0$ *for all* $x \in S_I$;
(c) *the set* S_O *is nonempty and bounded*;
(d) $\text{rank}(A) = m$.

Under the conditions that $S_O \neq \varnothing$ (hence, that $T_O \neq \varnothing$) and $\text{rank}(A) = m$, we note that Assumption 2.1(a) is equivalent to the boundedness of the optimal set T_O and that Assumption 2.1(c) is equivalent to the condition $T_I \neq \varnothing$ (see [9], for example).

In the introduction, we have briefly outlined the potential reduction (PR) algorithm presented in [6]. We next describe the PR algorithm in more detail. Consider the following potential function defined on the set S_I:

$$\Pi_q(x) \equiv q \ln c^T x - \sum_{j=1}^{n} \ln x_j, \tag{2.9}$$

where the constant q will be specified later. The algorithm can be described as follows.

Procedure PR algorithm $(A, b, c, x^0, \epsilon, x)$

Input: The input consists of the data A, b, c of problem (P), an initial point $x^0 \in S_I$ and a tolerance $\epsilon > 0$.

Output: A point $x \in S_I$ such that $c^T x < \epsilon$. Recall that the optimal value $r^* = 0$.

begin

  Set $x := x^0$;

  while $c^T x > \epsilon$ **do**

    $p = -XP_{AX}[Xc - \frac{c^T x}{q}\mathbf{1}]$;

    $\bar{\alpha} = \min\{-x_i/p_i;\ p_i < 0\}$;

    $\alpha^* = \text{argmin}\{\Pi_q(x + \alpha p);\ \alpha \in [0, \bar{\alpha}]\}$;

    $x = x + \alpha^* p$;

  od;

end

Under Assumption 2.1, with condition (c) replaced by the stronger condition that the set S_F is bounded, it has been shown in [6] that, if $q = n$ or $q = n + \sqrt{n}$, then the

PR algorithm can be used to find an optimal solution for (P) in at most $O(n^2L)$ iterations if $q = n$ or $O(nL)$ iterations if $q = n + \sqrt{n}$, where L denotes the number of bits necessary to represent the data of problem (P).

We are interested in studying the continuous trajectories of the vector field Φ_q: $S_I \to T(S_A)$ defined as

$$\Phi_q(x) \equiv -XP_{AX}\left[Xc - \frac{c^Tx}{q}\mathbf{1}\right], \tag{2.10}$$

that is, the trajectories determined by the maximal solutions of $\dot{x} = \Phi_q(x)$, where we assume that the constant q is fixed and satisfies the following conditions.

ASSUMPTION 2.2. (a) $q \geqslant n$;
(b) *If* $q = n$ *we also assume that the set* S_F *is bounded.*

The following relation is a consequence of Assumptions 2.1 and 2.2:

$$0 < |N| < q. \tag{2.11}$$

To see that, note that Assumption 2.1 (b) implies that $|N| > 0$, where the set N is as defined in (2.6). Clearly, $|N| \leqslant n \leqslant q$. However, if $q = n$ then, by (b) of Assumption 2.2, the set S_F is bounded and this implies that $|N| < n$. To see this last implication, assume that S_F is bounded. This implies that $0 \notin S_F$, for otherwise, it would follow that $S_F = \{0\}$, which is impossible due to Assumption 2.1 (a). If $|N| = n$, or equivalently, $N = \{1, \ldots, n\}$ were true then relation (2.7) and the fact that $0 \notin S_F$ would imply that $S_O = \varnothing$ which is a contradiction in view of Assumption 2.1. Hence, the set S_F being bounded implies $|N| < n$.

Before ending this section, we make some straightforward remarks that will be used frequently later. If $x \in S_A$ and $(y, z) \in T_A$ then $x^Tz = x^T(c - A^Ty) = c^Tx - b^Ty$. In particular, if $x \in S_A$ and $(\bar{y}, \bar{z}) \in T_0$ then $c^Tx = c^Tx - b^T\bar{y} = x^T\bar{z} = x_N^T\bar{z}_N$ since $\bar{z}_B = 0$, by relation (2.6) or (2.8), and $b^T\bar{y} = r^* = 0$, by Assumption 2.1(b). Similarly, if $\bar{x} \in S_O$ and $(y, z) \in T_A$ then $b^Ty = -\bar{x}_B^Tz_B$. For $(y, z) \in T_F$, $b^Ty \leqslant 0$ by the previous statement or by general duality theory.

2.3. *Preliminary results.* Using the basic results of the theory of ODE presented in §2.1, in this subsection we prove the following result.

THEOREM 2.1. *Let* $x\colon (w^-, w^+) \to S_I$ *be a maximal solution of the ODE* $\dot{x} = \Phi_q(x)$, *where the vector field* Φ_q: $S_I \to T(S_A)$ *is defined in* (2.10). *Then* $\lim_{t \to w^+} c^Tx(t) = 0$.

The proof of this theorem requires several preliminary lemmas. The next lemma shows that a zero, if there is any, of the vector field $\Phi_q(\cdot)$ must lie on the primal central trajectory (see, for example, [9]), that is, the set of points given by

$$\{x \in S_I;\, xz = \mu\mathbf{1} \text{ for some } (y, z) \in T_I \text{ and } \mu \in \mathbf{R}_+\}.$$

LEMMA 2.1. *For a point* $x \in S_I$, $\Phi_q(x) = 0$ *if and only if* $xz = (c^Tx/q)\mathbf{1}$ *for some* $(y, z) \in T_I$.

PROOF. The following equivalences are straightforward. We have

$$\Phi_q(x) = 0 \leftrightarrow P_{AX}\left[Xc - (c^Tx/q)\mathbf{1}\right] = 0 \leftrightarrow Xc - (c^Tx/q)\mathbf{1} \in \text{range}(XA^T)$$
$$\leftrightarrow c - (c^Tx/q)x^{-1} \in \text{range}(A^T) \leftrightarrow xz = (c^Tx/q)\mathbf{1} \quad \text{for some } (y, z) \in T_I.$$

□

The next lemma guarantees that, for any $q \geqslant n$, the vector field $\Phi_q(\cdot)$ does not vanish, and hence that the two equivalent statements of Lemma 2.1 are indeed empty.

LEMMA 2.2. *Under Assumption* 2.1 *and Assumption* 2.2, $\Phi_q(x) \neq 0$ *for all* $x \in S_I$.

PROOF. Assume, by contradiction, that $\Phi_q(x) = 0$, for some $x \in S_I$. Hence, by Lemma 2.1, it follows that $xz = (c^T x/q)\mathbf{1}$, for some $(y, z) \in T_I$. Multiplying this last relation on the left by $\mathbf{1}^T$ and using the identity $c^T x - b^T y = x^T z$, we obtain

$$(1 - n/q)c^T x = b^T y. \tag{2.12}$$

By Assumption 2.1 (b), we have that $c^T x > 0$ since $x \in S_I$. Hence, for $q > n$, the right side of (2.12) is positive while the left side is nonnegative since $(y, z) \in T_I \subseteq T_F$. Hence, we have a contradiction. For $q = n$, we obtain a contradiction as follows. If $q = n$ then $b^T y < 0$, since otherwise, we would have $(y, z) \in T_O \cap T_I$, and relation (2.8) would imply that $B = \varnothing$, or equivalently, that $N = \{1, \dots, n\}$, which is impossible due to relation (2.11). Hence, if $q = n$, relation (2.12) also gives a contradiction. □

LEMMA 2.3. *Consider the potential function* $\Pi_q\colon S_I \to \mathbf{R}$ *as defined in* (2.9). *Then, for any solution* $x(t)$ *of the ODE* $\dot{x} = \Phi_q(x)$, $t \to \Pi_q(x(t))$ *is a decreasing function of* t.

PROOF. For all $t \in \mathbf{R}$ for which $x(t)$ is defined, we have

$$\begin{aligned} \frac{d}{dt}\Pi_q(x(t)) &= \nabla \Pi_q(x(t))^T \dot{x}(t) \\ &= \left[(q/c^T x(t))c - x(t)^{-1}\right]^T \dot{x}(t) \\ &= -\frac{q}{c^T x(t)} \|P_{AX(t)}[X(t)c - (c^T x(t)/q)\mathbf{1}]\|^2 \\ &= -\frac{q}{c^T x(t)} \|x(t)^{-1}\dot{x}(t)\|^2 < 0 \end{aligned}$$

since $\dot{x}(t) = \Phi_q(x(t)) \neq 0$ by Lemma 2.2. Hence, $\Pi_q(x(t))$ is a decreasing function of t. □

LEMMA 2.4. *Assume that the set of optimal solutions of the LP problem* (P) *is bounded. Then there exist constants* $\gamma > 0$ *and* $\eta > 0$ *such that* $c^T x/\mathbf{1}^T x \geqslant \gamma$ *for all* $x \in \{x \in S_F; \mathbf{1}^T x \geqslant \eta\}$.

PROOF. It is not hard to verify, using LP duality theory, that the assumption implies that, for some $\epsilon > 0$,

$$r(\lambda) \equiv \min\left\{(c - \lambda\mathbf{1})^T x; x \in S_F\right\}$$

exists if $\lambda \in \mathbf{R}$ and $|\lambda| \leqslant \epsilon$. In particular, letting $r \equiv r(\epsilon)$, we obtain that $c^T x \geqslant \epsilon\mathbf{1}^T x + r$ for all $x \in S_F$. Hence, if we let $\gamma \equiv \epsilon/2$ and $\eta \equiv 2|r|/\epsilon$, we obtain

$$\frac{c^T x}{\mathbf{1}^T x} \geqslant \frac{\epsilon\mathbf{1}^T x + r}{\mathbf{1}^T x} \geqslant \epsilon - \frac{|r|}{\mathbf{1}^T x} \geqslant \frac{\epsilon}{2} = \gamma$$

for all $x \in \{x \in S_F; \mathbf{1}^T x \geqslant \eta\}$. □

LEMMA 2.5. *If $q > n$ then, for any $\delta \in \mathbf{R}$, the set $\{x \in S_I;\ \Pi_q(x) \leqslant \delta\}$ is bounded.*

PROOF. One can easily show, using the concavity of the logarithmic function, that for all $x \in S_I$,

$$\Pi_q(x) \geqslant q \ln c^T x - n \ln \mathbf{1}^T x + n \ln n$$

$$\geqslant q \ln c^T x - n \ln \mathbf{1}^T x.$$

By Lemma 2.4, it follows that there exist constants $\gamma > 0$ and $\eta > 0$ such that $c^T x / \mathbf{1}^T x \geqslant \gamma$ for all $x \in \{x \in S_F;\ \mathbf{1}^T x \geqslant \eta\}$. To prove the lemma, it is sufficient to show that the set $F \equiv \{x \in S_I; \Pi_q(x) \leqslant \delta, \mathbf{1}^T x \geqslant \eta\}$ is bounded. Indeed, if $x \in F$, we have

$$\delta \geqslant \Pi_q(x) \geqslant q \ln c^T x - n \ln \mathbf{1}^T x$$

$$\geqslant q \ln(\gamma \mathbf{1}^T x) - n \ln \mathbf{1}^T x$$

$$\geqslant q \ln \gamma + (q - n) \ln \mathbf{1}^T x$$

which implies that

$$\mathbf{1}^T x \leqslant \exp\left[\frac{\delta - q \ln \gamma}{q - n}\right].$$

Hence, the set F is bounded and the lemma follows. □

LEMMA 2.6. *For any $\beta > 0$ and $M \in \mathbf{R}$, the set $K = \{x \in S_I; c^T x \geqslant \beta, \Pi_q(x) \leqslant M\}$ is compact.*

PROOF. For $l \in \mathbf{R}^n$ and $u \in \mathbf{R}^n$, let $B(l) \equiv \{x \in S_I; x_j \geqslant l_j,\ j = 1, \ldots, n\}$ and $C(u) \equiv \{x \in S_I; x_j \leqslant u_j,\ j = 1, \ldots, n\}$. Noting that the set K is obviously closed with respect to S_I, it is sufficient to show that, for some $l \in \mathbf{R}^n_+$ and $u \in \mathbf{R}^n_+$, $K \subseteq B(l) \cap C(u)$, since the set $B(l) \cap C(u)$ is clearly a compact subset of S_I if $l \in \mathbf{R}^n_+$. Indeed, since for $q = n$, the set S_F is bounded, by Assumption 2.2, and for $q > n$, the set $\{x \in S_I;\ \Pi_q(x) \leqslant M\}$ is bounded, by Lemma 2.5, it follows that $K \subseteq C(u)$ for some $u \in \mathbf{R}^n_+$. We next show that $K \subseteq B(l)$ for some $l \in \mathbf{R}^n_+$. So, assume that $x \in K$ and let $k \in \{1, \ldots, n\}$. Then

$$\ln x_k = q \ln c^T x - \Pi_q(x) - \sum_{j \neq k} \ln x_j$$

$$\geqslant q \ln \beta - M - \sum_{j \neq k} \ln u_j$$

which shows that $x_k \geqslant l_k$ where $l_k \equiv \exp[M - q \ln \beta - \Sigma_{j \neq k} \ln u_j] > 0$. We have thus shown that $K \subseteq B(l)$ and the lemma follows. □

We are now in a position to prove Theorem 2.1.

PROOF OF THEOREM 2.1. Choose an arbitrary $t^0 \in (w^-, w^+)$ and let $x^0 \equiv x(t^0)$. It follows by Lemma 2.3 that

$$x([t^0, w^+)) \subseteq \{x \in S_I; \Pi_q(x) \leqslant \Pi_q(x^0)\}. \tag{2.13}$$

To prove the theorem, it is sufficient to show that $\limsup_{t \to w^+} c^T x(t) = 0$ in view of

Assumption 2.1(b). Assume by contradiction that $\limsup_{t\to w^+} c^T x(t) > 0$. It follows that, for some $\epsilon > 0$, there exists a sequence (t^k) with $\lim_{k\to\infty} t^k = w^+$ and such that $c^T x(t^k) \geqslant \epsilon$ for all k. In view of (2.13), for all k sufficiently large, $x(t^k) \in \{x \in S_I; \Pi_q(x) \leqslant \Pi_q(x^0), c^T x \geqslant \epsilon\}$ which is a compact set by Lemma 2.6. This contradicts the conclusion of Proposition 2.2 which clearly holds, in view of Lemma 2.3. □

We collect in the following result some immediate consequences of Theorem 2.1 that will be used later.

COROLLARY 2.1. *Let* x: $(w^-, w^+) \to S_I$ *be a maximal solution of the ODE* $\dot{x} = \Phi_q(x)$. *Then*

(a) *for every* $t^0 \in (w^-, w^+)$, *the set* $x([t^0, w^+))$ *is bounded*;

(b) $\lim_{t\to w^+} x_N(t) = 0$;

(c) *if* $\bar{x}$ *is an accumulation point of* $x(t)$, *as* t *tends to* w^+, *that is, if* $\bar{x} = \lim_{t\to w^+} x(t^k)$ *for some sequence* $(t^k) \subseteq (w^-, w^+)$ *converging to* w^+, *then* $\bar{x} \in S_O$.

The reader can easily verify the validity of Corollary 2.1 using (b) and (c) of Assumption 2.1, relation (2.7) and Theorem 2.1. We just observe that (c) of Assumption 2.1 implies that the set $\{x \in S_F | c^T x \leqslant M\}$ is bounded for any $M \in \mathbf{R}$. Note also that an alternative way to show (a) of Corollary 2.1 is by means of relation (2.13) and Lemma 2.5.

Corollary 2.1 does not guarantee that $\lim_{t\to w^+} x(t)$ exists. Clearly, if S_O consists of a unique point x^* then $\lim_{t\to w^+} x(t)$ exists and is equal to x^*. For the case in which S_O consists of more than one point, more appropriate techniques will be developed in the next sections with the purpose of showing that $\lim_{t\to w^+} x(t)$ indeed exists. Furthermore, we will show that all maximal solutions of $\dot{x} = \Phi_q(x)$ converge to the same point $x^* \in S_O$ and that x^* is the unique "center" (to be defined later) of the face S_O.

3. Limiting behavior of the maximal solutions of $\dot{x} = \Phi_q(x)$. In this section, we study in more detail the limiting behavior of the maximal solutions of the ODE $\dot{x} = \Phi_q(x)$. We show that all these maximal solutions converge to the "center" of the optimal face S_O. We will also introduce the concepts of dual solution curves and dual estimates which will play a fundamental role in this and the next section.

Throughout this section, we will assume that a fixed $x^0 \in S_I$ is given and, without loss of generality, we focus our attention on the Cauchy problem as follows:

$$\dot{x} = \Phi_q(x), \qquad x(0) = x^0. \tag{3.1}$$

It turns out that the analysis of the maximal solution of (3.1) is easier if we consider a more suitable Cauchy problem to be described next. This new Cauchy problem is essentially equivalent to the Cauchy problem (3.1) and, as will be shown, it can be easily "integrated" with the aid of some dual solution curves associated with its maximal solution. So, consider the following Cauchy problem:

$$\dot{x} = \frac{1}{h}\Phi_q(x), \tag{3.2a}$$

$$\dot{h} = \frac{c^T x}{q} - h, \tag{3.2b}$$

$$x(0) = x^0, \qquad h(0) = 1. \tag{3.2c}$$

Here, the vector field associated with the ODE (3.2a), (3.2b) is the mapping

$$\Psi_q(x,h) = \left(h^{-1}\Phi_q(x), (c^T x/q) - h\right)$$

whose domain of definition is the set $S_I \times \mathbf{R}_+$. Note that if $(x(\mu), h(\mu))$ is any solution of the ODE (3.2a) and (3.2b) then $h(\mu) > 0$ for all μ in the domain of definition of this solution since we are restricting the domain of definition of its associated vector field to the set $S_I \times \mathbf{R}_+$.

The following remark follows immediately from the results of §2.3.

REMARK 3.1. (a) The vector field $\Psi(x,h)$ does not vanish in the set $S_I \times \mathbf{R}_+$.

(b) If $(x(\mu), h(\mu))$ is a solution of the ODE $(\dot{x}, \dot{h}) = \Psi_q(x,h)$ then the potential function $\overline{\Pi}_q$: $S_I \times \mathbf{R}_+ \to \mathbf{R}$ defined as $\overline{\Pi}_q(x,h) = \Pi_q(x)$ is such that $\mu \to \overline{\Pi}_q(x(\mu), h(\mu))$ is a decreasing function of μ.

Throughout this section, we let ξ: $(w^-, w^+) \to S_I$ and (x,h): $(\alpha^-, \alpha^+) \to S_I \times \mathbf{R}_+$ denote the maximal solutions of the Cauchy problems (3.1) and (3.2), respectively. We will refer to the curve x: $(\alpha^-, \alpha^+) \to S_I$ as the x-component of the maximal solution of (3.2). The following result will be useful later.

PROPOSITION 3.1. *Let* ξ: $(w^-, w^+) \to S_I$ *and* (x,h): $(\alpha^-, \alpha^+) \to S_I \times \mathbf{R}_+$ *denote the maximal solutions of* (3.1) *and* (3.2), *respectively. Then*

(a) $h(\mu) = e^{-\mu}\int_0^\mu e^s g(s)\,ds + e^{-\mu}$ *for all* $\mu \in (\alpha^-, \alpha^+)$, *where* $g(\mu) \equiv c^T x(\mu)/q$ *for* $\mu \in (\alpha^-, \alpha^+)$.

(b) *Let* $\eta(\mu) \equiv \int_0^\mu h(s)^{-1}\,ds$ *for all* $\mu \in (\alpha^-, \alpha^+)$. *Then* $\eta((\alpha^-, \alpha^+)) \subseteq (w^-, w^+)$ *and* $x(\mu) = \xi(\eta(\mu))$ *for all* $\mu \in (\alpha^-, \alpha^+)$.

(c) *The sets* $x([0, \alpha^+)) \subseteq S_I$ *and* $h([0, \alpha^+)) \subseteq \mathbf{R}$ *are bounded.*

(d) $\eta([0, \alpha^+)) = [0, w^+)$ *and, as a consequence,* $x([0, \alpha^+)) = \xi([0, w^+))$.

PROOF. Statement (a) follows by integrating the relation $\dot{h}(\mu) = (c^T x(\mu)/q) - h(\mu)$ over the interval $[0, \mu]$ and using the initial condition $h(0) = 1$. Statement (b) follows from Proposition 2.3 and the fact that x: $(\alpha^-, \alpha^+) \to S_I$ is clearly the maximal solution of the Cauchy problem

$$\dot{x} = \frac{1}{h(\mu)}\Phi_q(x), \qquad x(0) = x^0,$$

whose domain of definition is the set $S_I \times (\alpha^-, \alpha^+)$. To prove statement (c), note that $x([0, \alpha^+))$ is bounded since, by statement (b), $x([0, \alpha^+)) \subseteq \xi([0, w^+))$ and the latter set is bounded due to (a) of Corollary 2.1. It then follows that there exists $M > 0$ such that $g(\mu) \leqslant M$ for all $\mu \in [0, \alpha^+)$. Hence,

$$h(\mu) = e^{-\mu}\left[\int_0^\mu e^s g(s)\,ds + 1\right] \leqslant e^{-\mu}\left[M(e^\mu - 1) + 1\right],$$

which shows that $h([0, \alpha^+))$ is bounded from above. Since, by statement (a) or by the observation preceding Remark 3.1, $h(\cdot)$ is a positive function in $[0, \alpha^+)$, it follows that $h([0, \alpha^+))$ is bounded. We next show statement (d). Assume by contradiction that $\eta([0, \alpha^+)) \neq [0, w^+)$. It follows from statement (b) that, for some $\beta \in [0, w^+)$, $\eta([0, \alpha^+)) \subseteq [0, \beta]$ and hence that $x([0, \alpha^+)) \subseteq \xi([0, \beta])$. In other words, $x([0, \alpha^+))$ is contained in a compact subset of S_I. Noting that the function $\overline{\Pi}_q(x,h)$ defined in (b) of Remark 3.1 fulfills the assumption of Proposition 2.2 with respect to any solution of the ODE (3.2a) and (3.2b), it follows from the conclusions of this proposition and the previous observation that $h(\mu)$ must lie outside any compact subset of $\mathbf{R}_+$, for all $\mu \in (\alpha^-, \alpha^+)$ sufficiently close to α^+. But this implies that $\lim_{\mu \to \alpha^+} h(\mu) = 0$ since, by statement (c), the set $h([0, \alpha^+))$ is bounded. On the other hand, since $x([0, \alpha^+))$ is contained in a compact subset of S_I, $g(\mu) \equiv c^T x(\mu)/q$ is bounded away from 0, that is, for some $\gamma > 0$, $g(\mu) \geqslant \gamma$ for all $\mu \in [0, \alpha^+)$. As a consequence, statement (a) implies

$$h(\mu) \geqslant \gamma e^{-\mu} \int_0^{\mu} e^s \, ds = \gamma[1 - e^{-\mu}],$$

which shows that $\liminf_{\mu \to \alpha^+} h(\mu) > 0$. This contradicts the fact that $\lim_{\mu \to \alpha^+} h(\mu) = 0$ and therefore (d) follows. □

Let ξ: $(w^-, w^+) \to S_I$ and x: $(\alpha^-, \alpha^+) \to S_I$ denote, respectively, the maximal solution of (3.1) and the x-component of the maximal solution of (3.2). Statements (b) and (d) of Proposition 3.1 show that the functions $\xi|[0, w^+)$ and $x|[0, \alpha^+)$ are parametrizations of the same "piece of trajectory" $x([0, \alpha^+)) = \xi([0, w^+))$. Hence, the limiting behavior of the PR trajectory through the point x^0 can be studied by analyzing the behavior of the curve $x(\mu)$, as μ tends to α^+. This observation also establishes our earlier assertion that the two Cauchy problems (3.1) and (3.2) are essentially equivalent.

As a consequence of Proposition 3.1, Theorem 2.1 and Corollary 2.1 we obtain the following result whose proof is immediate.

THEOREM 3.1. *Let* (x, h): $(\alpha^-, \alpha^+) \to S_I \times \mathbf{R}_+$ *denote the maximal solution of* (3.2). *Then* $\lim_{\mu \to \alpha^+} c^T x(\mu) = 0$ *and* $\lim_{\mu \to \alpha^+} x_N(\mu) = 0$.

At this point, the reader may wonder why we are considering two different parametrizations for the PR trajectory through the point x^0. The curve $x(\cdot)$ is in fact the parametrization we are interested in, since, as we will see later, it has an easier "description". The parametrization $\xi(\cdot)$ will not be used anymore from this point on, but we found it useful to use $\xi(\cdot)$ as a tool in proving Theorem 3.1 (with the aid of Theorem 2.1 and Corollary 2.1). In fact, we were not able to prove Theorem 3.1 in a more direct way by just considering the parametrization $x(\cdot)$.

Up to this point, we have said nothing about the value of the right endpoint α^+ of the interval of definition of the maximal solution of the Cauchy problem (3.2). The next result gives a necessary and sufficient condition for $\alpha = +\infty$. Later, it is shown that this condition is always satisfied (Proposition 3.5).

PROPOSITION 3.2. *Let* (x, h): $(\alpha^-, \alpha^+) \to S_I \times \mathbf{R}_+$ *denote the maximal solution of* (3.2). *Then* $\alpha^+ = \infty$ *if and only if* $\lim_{\mu \to \alpha^+} h(\mu) = 0$.

PROOF. The 'if' part of the equivalence follows easily from statement (a) of Proposition 3.1. To show the 'only if' part, assume that $\alpha^+ = \infty$. Let $\epsilon > 0$ be given. Let μ^0 be such that $g(\mu) \equiv c^T x(\mu)/q \leqslant \epsilon/2$ for all $\mu \geqslant \mu^0$. Let $\mu^1 \geqslant \mu^0$ be such that

$$e^{-\mu}\left[\int_0^{\mu^0} e^s g(s) \, ds + 1\right] \leqslant \frac{\epsilon}{2}$$

for all $\mu \geqslant \mu^1$. Thus, if $\mu \geqslant \mu^1$, it follows from (a) of Proposition 3.1 that

$$\begin{aligned} h(\mu) &= e^{-\mu}\left[\int_0^{\mu} e^s g(s) \, ds + 1\right] \\ &= e^{-\mu}\left[\frac{\epsilon}{2}\int_{\mu^0}^{\mu} e^s \, ds + \int_0^{\mu^0} e^s g(s) \, ds + 1\right] \\ &\leqslant \frac{\epsilon}{2} e^{-\mu}\left[e^{\mu} - e^{\mu^0}\right] + \frac{\epsilon}{2} \leqslant \epsilon. \end{aligned}$$

We have thus shown that $\alpha^+ = +\infty$ implies $\lim_{\mu \to \alpha^+} h(\mu) = 0$. □

Before we introduce the notion of dual solution curves associated with the maximal solution of the Cauchy problem (3.2), we define the concept of dual estimates that naturally arise in the context of the PR algorithm.

DEFINITION 3.1. Given a point $x \in S_I$, the dual estimate $(y^E(x), z^E(x)) \in T_A$ at the point x is defined as follows:

$$y^E(x) = (AX^2A^T)^{-1}AX\left(Xc - \frac{c^Tx}{q}\mathbf{1}\right),$$

$$z^E(x) = c - A^Ty^E(x).$$

From the algorithmic point of view, the usefulness of these dual estimates is that, for a near optimal point $x \in S_I$, $(y^E(x), z^E(x))$ can be used in an attempt to obtain a near optimal (possibly infeasible) solution for the dual problem (D). From the theoretical point of view, we will see in §4 that these dual estimates play an important role in the analysis of the asymptotic direction of convergence of the PR trajectories. The reader familiar with the primal affine scaling algorithm (see [2], [4] and [15]) has perhaps realized that the definition of the dual estimates above closely parallels the ones for the dual estimates associated with the primal affine scaling algorithm.

We next introduce dual solution curves which are naturally associated with the maximal solution of the Cauchy problem (3.2) and then show how to "integrate" (3.2) using these dual solution curves. Let (x, h): $(\alpha^-, \alpha^+) \to S_I \times \mathbf{R}_+$ denote the maximal solution of (3.2). For a given $(y^0, z^0) \in T_A$, we define the *(associated) dual solution curve through* (y^0, z^0) to be the maximal solution (y, z): $(\alpha^-, \alpha^+) \to T_A$ of the following Cauchy problem:

$$\dot{y} = y^E(x(\mu)) - y,$$

$$\dot{z} = z^E(x(\mu)) - z,$$

$$y(0) = y^0, \qquad z(0) = z^0,$$

whose domain of definition is the set $T_A \times (\alpha^-, \alpha^+)$. The following remark gives some relations involving the maximal solution of (3.2) and its associated dual solution curves.

REMARK 3.2. The maximal solution (x, h): $(\alpha^-, \alpha^+) \to S_I \times \mathbf{R}_+$ of (3.2) and its associated dual solution curve (y, z): $(\alpha^-, \alpha^+) \to T_A$ through $(y^0, z^0) \in T_A$ satisfy the following relations:

$$\text{(3.3a)} \qquad \dot{z}(\mu) + h(\mu)x(\mu)^{-2}\dot{x}(\mu) = \frac{c^Tx(\mu)}{q}x(\mu)^{-1} - z(\mu),$$

$$\text{(3.3b)} \qquad \dot{h}(\mu) = \frac{c^Tx(\mu)}{q} - h(\mu),$$

$$\text{(3.3c)} \qquad A\dot{x}(\mu) = 0,$$

$$\text{(3.3d)} \qquad A^T\dot{y}(\mu) + \dot{z}(\mu) = 0.$$

Remark 3.2 follows by direct verification that the relations above are equivalent to the relations $(\dot{x}(\mu), \dot{h}(\mu)) = \Psi_q(x(\mu), h(\mu))$, $\dot{y}(\mu) = y^E(x(\mu)) - y(\mu)$ and $\dot{z}(\mu) = z^E(x(\mu)) - z(\mu)$, where $\Psi_q(\cdot)$ is the vector field defined after the relations in (3.2).

The motivation for introducing the Cauchy problem (3.2) as an aid to study the PR trajectories is not clear up to this point. However, the main motivation for introducing (3.2) is given by the following result which shows that the maximal solutions of the Cauchy problem (3.2) can be easily "integrated" with the aid of the dual solution curves introduced before.

PROPOSITION 3.4. *Let* $(x, h)\colon (\alpha^-, \alpha^+) \to S_I \times \mathbf{R}_+$ *denote the maximal solution of* (3.2) *and let* $(y, z)\colon (\alpha^-, \alpha^+) \to T_A$ *denote its associated dual solution curve through* $(y^0, z^0) \in T_A$. *Then, for all* $\mu \in (\alpha^-, \alpha^+)$,

$$z(\mu) - h(\mu)x(\mu)^{-1} = pe^{-\mu}, \tag{3.4}$$

where $p \equiv z^0 - (x^0)^{-1}$.

PROOF. It is sufficient to show that $w(\mu) \equiv z(\mu) - h(\mu)x(\mu)^{-1}$, $\mu \in (\alpha^-, \alpha^+)$, is a solution of

$$\dot{w}(\mu) = -w(\mu), \qquad w(0) = p,$$

since the unique solution of this Cauchy problem in the interval (α^-, α^+) is equal to $pe^{-\mu}$. That $w(0) = p$ is immediate. To show that $\dot{w}(\mu) = -w(\mu)$, note that relations (3.3a) and (3.4) imply

$$\begin{aligned}\dot{w}(\mu) &= \frac{d}{d\mu}\left(z(\mu) - h(\mu)x(\mu)^{-1}\right)\\ &= \dot{z}(\mu) + h(\mu)x(\mu)^{-2}\dot{x}(\mu) - \dot{h}(\mu)x(\mu)^{-1}\\ &= \frac{c^T x(\mu)}{q}x(\mu)^{-1} - z(\mu) - \dot{h}(\mu)x(\mu)^{-1}\\ &= h(\mu)x(\mu)^{-1} - z(\mu) = -w(\mu)\end{aligned}$$

from which the result follows. □

The relations in (3.3) and relation (3.4) will play a fundamental role in the later results of this paper. These relations allow us to express the vectors $x(\mu), (y(\mu), z(\mu))$, $\dot{x}(\mu)$ and $(\dot{y}(\mu), \dot{z}(\mu))$ (or, subvectors of these vectors) as optimal solutions of certain convex optimization problems. The following corollary of Proposition 3.4, although not needed in our discussion, illustrates our previous comment and also points out a strong similarity between the approach undertaken in this work and the one presented in [1].

COROLLARY 3.1. *Let* $(x, h)\colon (\alpha^-, \alpha^+) \to S_I \times \mathbf{R}_+$ *denote the maximal solution of* (3.2) *and let* $(y, z)\colon (\alpha^-, \alpha^+) \to T_A$ *denote its associated dual solution curve through* $(y^0, z^0) \in T_A$. *Then, for all* $\mu \in (\alpha^-, \alpha^+)$,

(a) $x(\mu)$ *is the (unique) optimal solution of the problem*

$$\min_x \left\{c^T x - e^{-\mu}p^T x - h(\mu)\sum_{j=1}^n \ln x_j;\ Ax = b,\ x > 0\right\}. \tag{3.5}$$

(b) $(y(\mu), z(\mu))$ *is the (unique) optimal solution of the problem*

$$(3.6)\qquad \max_{y,z}\left\{b^T y + h(\mu)\sum_{j=1}^{n}\ln(z_j - e^{-\mu}p_j);\ A^T y + z = c,\ z - e^{-\mu}p > 0\right\}.$$

PROOF. The result follows immediately by noting that the optimality conditions of problems (3.5) and (3.6) both reduce to $z(\mu) - h(\mu)x^{-1} = pe^{-\mu}$, $x(\mu) \in S_I$ and $(y(\mu), z(\mu)) \in T_A$. □

PROPOSITION 3.5. *Let* $(x, h)\colon (\alpha^-, \alpha^+) \to S_I \times \mathbf{R}_+$ *denote the maximal solution of* (3.2) *and let* $(y, z)\colon (\alpha^-, \alpha^+) \to T_A$ *denote its associated dual solution curve through* $(y^0, z^0) \in T_A$. *Then*

(a) *the set* $\{(y(\mu), z(\mu));\ \mu \in [0, \alpha^+)\}$ *is bounded,*

(b) $\lim_{\mu\to\alpha^+} h(\mu) = 0$ *and hence* $\alpha^+ = \infty$,

(c) $\lim_{\mu\to\infty} b^T y(\mu) = 0$.

PROOF. Note that statement (a) follows once we show that $\{z(\mu);\ \mu \in [0, \alpha^+)\}$ is bounded, since this fact and the condition $\operatorname{rank}(A) = m$ (see (d) of Assumption 2.1) implies that $\{y(\mu);\ \mu \in [0, \alpha^+)\}$ is bounded. To show that, it is sufficient to show that the set $\{z(\mu) - e^{-\mu}p;\ \mu \in [0, \alpha^+)\}$ is bounded, where $p \equiv z^0 - (x^0)^{-1}$. From relation (3.4), it follows that

$$(3.7)\qquad nh(\mu) + e^{-\mu}p^T x(\mu) = x(\mu)^T z(\mu) = c^T x(\mu) - b^T y(\mu)$$

which implies that

$$\begin{aligned}(x^0)^T(z(\mu) - e^{-\mu}p) &= (x^0)^T z(\mu) - e^{-\mu}p^T x^0\\ &= c^T x^0 - b^T y(\mu) - e^{-\mu}p^T x^0\\ &= c^T x^0 + nh(\mu) + e^{-\mu}p^T x(\mu) - c^T x(\mu) - e^{-\mu}p^T x^0\end{aligned}$$

for all $\mu \in (\alpha^-, \alpha^+)$. By (c) of Proposition 3.1, we know that the sets $x([0, \alpha^+))$ and $h([0, \alpha^+))$ are bounded. This implies that each of the terms in the last expression above is bounded in $[0, \alpha^+)$. Since $x^0 > 0$ and $(z(\mu) - e^{-\mu}p) > 0$ for all $\mu \in [0, \alpha^+)$, it follows that $\{z(\mu) - e^{-\mu}p;\ \mu \in [0, \alpha^+)\}$ is bounded. To show (b), note that by (3.4)

$$(3.8)\qquad x_j(\mu) = \frac{h(\mu)}{z_j(\mu) - e^{-\mu}p_j}.$$

In particular, consider relation (3.8) for an index $j \in N$. By Theorem 3.1, $\lim_{\mu\to\alpha^+} x_j(\mu) = 0$. But this implies that if $\limsup_{\mu\to\alpha^+} h(\mu) > 0$ then $\limsup_{\mu\to\alpha^+} z_j(\mu) - e^{-\mu}p = +\infty$ which is impossible due to statement (a) above. Hence, $\lim_{\mu\to\alpha^+} h(\mu) = \limsup_{\mu\to\alpha^+} h(\mu) = 0$. As a consequence of Proposition 3.2, it follows that $\alpha^+ = +\infty$.

Statement (c) is an immediate consequence of statement (b), Theorem 3.1 and relation (3.7). □

The last proposition establishes the fact that any interval of definition of a maximal solution of the ODE $(\dot{x}, \dot{h}) = \Psi(x, h)$ is of the form $(\alpha^-, +\infty)$. So, from now on, we write $+\infty$ instead of α^+ for the right endpoint of these maximal intervals. If ξ:

$(w^-, w^+) \to S_I$ is the maximal solution of (3.1), we can easily show, using statement (b) of Proposition 3.1 and Proposition 3.5, that also $w^+ = +\infty$.

Before we analyze the limiting behavior of the maximal solution of (3.2), we examine the limiting behavior of its associated dual solution curves in the following result.

THEOREM 3.2. *Let (y, z): $(\alpha^-, \infty) \to T_A$ be the dual solution curve through any $(y^0, z^0) \in T_A$ associated with the maximal solution (x, h): $(\alpha^-, \infty) \to S_I \times \mathbf{R}_+$ of the Cauchy problem* (3.2). *Then* $\lim_{\mu\to\infty}(y(\mu), z(\mu)) = (y^*, z^*)$ *where (y^*, z^*) is the optimal solution of problem (D) (and hence $z_B^* = 0$) such that (y^*, z_N^*) is the (unique) optimal solution of the problem*

$$\max_{y, z_N}\left\{\sum_{j\in N} \ln z_j;\ A_B^T y = c_B,\ A_N^T y + z_N = c_N,\ z_N > 0\right\}. \tag{3.9}$$

PROOF. By (a) of Proposition 3.5, it is enough to show that any accumulation point $(\bar{y}, \bar{z})$ of $(y(\mu), z(\mu))$, as μ tends to ∞, satisfies $(\bar{y}, \bar{z}) = (y^*, z^*)$. So, assume that $(\bar{y}, \bar{z}) = \lim_{k\to\infty}(y(\mu^k), z(\mu^k))$ for some sequence (μ^k) tending to ∞. We break the proof that $(\bar{y}, \bar{z}) = (y^*, z^*)$ up into three claims.

Claim 1. $(\bar{y}, \bar{z}) \in T_O$ and hence $\bar{z}_B = z_B^* = 0$.

Indeed, first note that $(\bar{y}, \bar{z}) \in T_A$ and $b^T\bar{y} = 0$ due to statement (c) of Proposition 3.5. Moreover,

$$\bar{z} = \lim_{k\to\infty} z(\mu^k) = \lim_{k\to\infty}\left(z(\mu^k) - e^{-\mu^k}p\right) \geqslant 0$$

where $p = z^0 - (x^0)^{-1}$, since $z(\mu) - e^{-\mu}p = h(\mu)x(\mu)^{-1} > 0$ for all $\mu \in (\alpha^-, \infty)$, by Proposition 3.4. Hence, $(\bar{y}, \bar{z}) \in T_O$ and the claim follows.

Claim 2. $(y(\mu), z_N(\mu))$ is the (unique) optimal solution of the problem

$$\max_{y, z_N}\left\{\sum_{j\in N} \ln(z_j - e^{-\mu}p_j);\ A_B^T y = c_B - z_B(\mu),\ A_N^T y + z_N = c_N,\ z_N - e^{-\mu}p_N > 0\right\}. \tag{3.10}$$

Indeed, first note that

$$A_N x_N(\mu) = b - A_B x_B(\mu) \in \operatorname{range}(A_B) \tag{3.11}$$

since $b \in \operatorname{range}(A_B)$. Since, by relation (3.4), $x_N(\mu) = h(\mu)(z_N(\mu) - e^{-\mu}p_N)^{-1}$, relation (3.11) implies that $A_N(z_N(\mu) - e^{-\mu}p_N)^{-1} \in \operatorname{range}(A_B)$ which together with the relations $A_B^T y(\mu) = c_B - z_B(\mu)$ and $A_N^T y(\mu) + z_N(\mu) = c_N$ show that $(y(\mu), z_N(\mu))$ satisfy the optimality conditions for problem (3.10).

Claim 3. $(\bar{y}, \bar{z}_N) = (y^*, z_N^*)$.

Indeed, assume by contradiction that $(\bar{y}, \bar{z}_N) \neq (y^*, z_N^*)$. Let $\Delta y \equiv y^* - \bar{y}$ and $\Delta z_N \equiv z_N^* - \bar{z}_N$ and let $(\tilde{y}^k, \tilde{z}_N^k) \equiv (y(\mu^k) + \Delta y, z_N(\mu^k) + \Delta z_N)$ for all k. Observe that Claim 1 and the definition of (y^*, z^*) imply that $A_B^T\Delta y = 0$ and $A_N^T\Delta y + \Delta z_N = 0$. This implies that $A_B^T\tilde{y}^k = c_B - z_B(\mu^k)$ and $A_N^T\tilde{y}^k + \tilde{z}_N^k = c_N$. Moreover,

$$\lim_{k\to\infty}\left(\tilde{y}^k, \tilde{z}_N^k\right) = \lim_{k\to\infty}\left(y(\mu^k), z_N(\mu^k)\right) + (\Delta y, \Delta z_N)$$

$$= \left(\bar{y}, \bar{z}_N\right) + \left(y^* - \bar{y}, z_N^* - \bar{z}_N\right) = \left(y^*, z_N^*\right).$$

It follows from the fact that $z_N^* > 0$ and the two last observations that, for all k sufficiently large, $(\bar{y}^k, \bar{z}_N^k)$ is feasible for problem (3.10) with $\mu = \mu^k$. This implies that, for all k sufficiently large,

$$\sum_{j \in N} \ln\left(\bar{z}_j^k - e^{-\mu^k} p_j\right) \leqslant \sum_{j \in N} \ln\left(z_j(\mu^k) - e^{-\mu^k} p_j\right)$$

which in turn implies, upon taking the limit as $k \to \infty$, that either $\sum_{j \in N} \ln z_j^* \leqslant \sum_{j \in N} \ln \bar{z}_j$ if $\bar{z}_N > 0$ or that $\sum_{j \in N} \ln z_j^* \leqslant -\infty$ if $\bar{z}_j = 0$ for some $j \in N$. In view of the definition of (y^*, z^*), it follows that either one of the cases above results in a contradiction and therefore, we must have $(\bar{y}, \bar{z}_N) = (y^*, z_N^*)$. □

The optimal solution (y^*, z^*), as in the statement of Theorem 3.2, will be referred to as the *center* of the optimal face T_O.

As our main aim of this section, we next turn out attention to the study of the limiting behavior of the maximal solution of the Cauchy problem (3.2). We first prove two preliminary lemmas.

LEMMA 3.1. *Let (x, h): $(\alpha^-, \infty) \to S_I \times \mathbf{R}_+$ denote the maximal solution of* (3.2). *Then* $\lim_{\mu \to \infty} c^T x(\mu)/h(\mu) = |N|$.

PROOF. First note that, by the remarks at the end of §2.2 and relation (3.4), we have

$$c^T x(\mu) = (z_N^*)^T x_N(\mu) = h(\mu)(z_N^*)^T \left(z_N(\mu) - e^{-\mu} p_N\right)^{-1}$$

and since, by Theorem 3.2, $\lim_{\mu \to \infty}(z_N(\mu) - e^{-\mu} p_N)^{-1} = (z_N^*)^{-1}$, the result follows. □

LEMMA 3.2. *Let (x, h): $(\alpha^-, \infty) \to S_I \times \mathbf{R}_+$ denote the maximal solution of* (3.2). *Then* $\lim_{\mu \to \infty} h(\mu) e^{\mu} = +\infty$.

PROOF. By relation (2.11), $|N| > 0$ and hence, by the previous lemma, $\lim_{\mu \to \infty} g(\mu)/h(\mu) > 0$ where $g(\mu) \equiv c^T x(\mu)/q$ for $\mu \in (\alpha^-, \infty)$. Thus, there exist constants $\gamma > 0$ and $\mu^0 \geqslant 0$ such that $g(\mu) \geqslant \gamma h(\mu)$ for all $\mu \geqslant \mu^0$. Then, by (a) of Proposition 3.1,

(3.12)

$$h(\mu) e^{\mu} = \int_0^{\mu} e^s g(s)\, ds + 1 = h(\mu^0) e^{\mu^0} + \int_{\mu^0}^{\mu} e^s g(s)\, ds$$

$$\geqslant h(\mu^0) e^{\mu^0} + \gamma \int_{\mu^0}^{\mu} e^s h(s)\, ds \geqslant h(\mu^0) e^{\mu^0} + \gamma h(\mu^0) e^{\mu^0} (\mu - \mu^0)$$

where the last inequality follows from the fact that $h(\mu) e^{\mu}$ is an increasing function of μ and hence $h(\mu) e^{\mu} \geqslant h(\mu^0) e^{\mu^0}$ for all $\mu \geqslant \mu^0$. Inequality (3.12) shows that $\lim_{\mu \to \infty} h(\mu) e^{\mu} = +\infty$. □

We are now in a position to analyze the limiting behavior of the maximal solution of (3.2). The next result refines the content of Theorem 3.1 and is the main result of this section.

THEOREM 3.3. *Let (x, h): $(\alpha^-, \infty) \to S_I \times \mathbf{R}_+$ denote the maximal solution of* (3.2). *Then* $\lim_{\mu \to \infty} x(\mu) = x^*$ *where x^* is the optimal solution of problem* (P) (*and* *hence* $x_N^* = 0$) *such that* x_B^* *is the (unique) optimal solution of the problem*

$$\max_{x_B}\left\{\sum_{j\in B}\ln x_j;\ A_Bx_B = b,\ x_B > 0\right\}.$$

PROOF. By (c) of Proposition 3.5, the set $x([0,\alpha^+))$ is bounded. Hence, it is enough to show that any accumulation point $\bar{x}$ of $x(\mu)$, as μ tends to ∞, satisfies $\bar{x} = x^*$. So, assume that $\bar{x} = \lim_{k\to\infty} x(\mu^k)$ for some sequence (μ^k) tending to ∞. Clearly, from Theorem 3.1, it follows that $\bar{x} \in S_O$ and $\bar{x}_N = 0 = x_N^*$. So, it is enough to show that $\bar{x}_B = x_B^*$.

Claim. $x_B(\mu)$ is the (unique) optimal solution of the problem

(3.13)

$$\max_{x_B}\left\{(e^\mu h(\mu))^{-1} p_B^T x_B + \sum_{j\in B}\ln x_j;\ A_Bx_B = b - A_Nx_N(\mu),\ x_B > 0\right\}.$$

Indeed, by the relation (3.4) and the fact that $c_B \in \text{range}(A_B^T)$, it follows that

$$\begin{aligned}(e^\mu h(\mu))^{-1}p_B + x_B(\mu)^{-1} &= h(\mu)^{-1}z_B(\mu)\\ &= h(\mu)^{-1}\left(c_B - A_B^T y(\mu)\right) \in \text{range}\left(A_B^T\right)\end{aligned}$$

which together with the relation $A_Bx_B(\mu) = b - A_Nx_N(\mu)$ shows that $x_B(\mu)$ satisfies the optimality conditions for problem (3.13). The proof that $\bar{x}_B = x_B^*$ now follows by using the same arguments as used in Claim 3 of the proof of Theorem 3.2 and using the fact that, by Lemma 3.2, $\lim_{\mu\to\infty} e^\mu h(\mu) = \infty$. □

The optimal solution x^*, as in the statement of Theorem 3.3, will be referred to as the *center* of the optimal face S_O. Note that the characterization of the center x^* in Theorem 3.3 implies that $(x_B^*)^{-1} \in \text{range}(A_B^T)$.

4. The asymptotic direction of convergence of the PR trajectories. In this section, we are interested in analyzing the way the PR trajectories approach the optimal face of problem (P). More specifically, we want to study the asymptotic direction of convergence of the PR trajectories near the optimal face of (P). If $(x,h)\colon (\alpha^-,\infty) \to S_I \times \mathbf{R}_+$ denotes the maximal solution of (3.2), this is achieved by analyzing the limit of the unit vector $\dot{x}(\mu)/\|\dot{x}(\mu)\|$ as μ approaches $+\infty$.

Similar to the approach undertaken in [1], we first study the limit of the dual estimates $(y^E(x(\mu)), z^E(x(\mu)))$ at points along the maximal solution $x(\mu)$ of (3.2), as μ approaches $+\infty$. The following remark will be used in the proof of the next result.

REMARK 4.1. For a point $x \in S_I$, the dual estimate $(y^E(x), z^E(x))$ is the unique optimal solution of the following convex quadratic problem:

$$\min_{y,z}\left\{\|Xz - (c^Tx/q)\mathbf{1}\|^2;\ A^Ty + z = c\right\}. \tag{4.1}$$

In fact, the remark above can be used as an alternative definition of the dual estimate $(y^E(x), z^E(x))$. Note that this alternative definition holds even in the case we do not assume that the constraint matrix A has full row rank.

We can now state the following crucial theorem.

THEOREM 4.1. *Let* $(x, h): (\alpha^-, \infty) \to S_I \times \mathbf{R}_+$ *denote the maximal solution of* (3.2). *Let* $(y^E(\mu), z^E(\mu))$ *denote the point* $(y^E(x(\mu)), z^E(x(\mu)))$ *for all* $\mu \in (\alpha^-, +\infty)$. *Then*

(a) $\lim_{\mu \to \infty}(y^E(\mu), z^E(\mu)) = (y^*, z^*)$ *where* (y^*, z^*) *is the center of the optimal face of the dual problem* (D).

(b) $\lim_{\mu \to \infty} z_B^E(\mu)/h(\mu) = (|N|/q)(x_B^*)^{-1}$ *where* x^* *is the center of the optimal face of problem* (P).

PROOF. To make the proof clearer, we break it up into several steps.

Step 1. By Remark 4.1, it follows that $(y^E(\mu), z^E(\mu))$ is the unique optimal solution of the problem

$$\min_{r,s}\left\{\|x(\mu)s - g(\mu)\mathbf{1}\|^2; A^T r + s = c\right\}, \tag{4.2}$$

where $g(\mu) \equiv c^t x(\mu)/q$, for $\mu \in (\alpha^-, +\infty)$. Since, by relation (3.4), $x(\mu) = h(\mu)(z(\mu) - e^{-\mu}p)^{-1}$ for some $p \in \mathbf{R}^n$, problem (4.2) can be rewritten in terms of the index sets B and N as follows:

(4.3)

$$\min_{r,s} \|x_B(\mu)s_B - g(\mu)\mathbf{1}_B\|^2 + h(\mu)^2\|(z_N(\mu) - e^{-\mu}p_N)^{-1}s_N - (g(\mu)/h(\mu))\mathbf{1}_N\|^2$$

$$\text{s.t. } A_B^T r + s_B = c_B, \qquad A_N^T r + s_N = c_N.$$

Step 2. Let $w_B(\mu)$, for $\mu \in (\alpha^-, \infty)$, denote the (unique) optimal solution of the problem

$$\theta_B(\mu) \equiv \min_{w_B}\left\{\|x_B(\mu)w_B - (g(\mu)/h(\mu))\mathbf{1}_B\|^2; w_B \in \text{range}(A_B^T)\right\}. \tag{4.4}$$

We claim that the optimal value $\theta_B(\mu)$ of (4.4) satisfies $\lim_{\mu \to \infty} \theta_B(\mu) = 0$. Indeed, first observe that, by Theorem 3.3 and Lemma 3.1, we have $\lim_{\mu \to \infty} x_B(\mu) = x_B^*$ and $\lim_{\mu \to \infty}(g(\mu)/h(\mu)) = |N|/q$. Using the definition of $w_B(\mu)$ as the optimal solution of problem (4.4) and the previous observations, one can easily show that the set $w_B([0, \infty))$ is bounded. Moreover, $w_B(\mu)$ satisfies the optimality conditions for problem (4.4), that is, for all $\mu \in (\alpha^-, \infty)$ there holds

$$x_B(\mu)^2 w_B(\mu) - (g(\mu)/h(\mu))x_B(\mu) \in \text{range}(A_B), \qquad w_B(\mu) \in \text{range}(A_B^T).$$

Now, let $\overline{w}_B$ be an accumulation point of $w_B(\mu)$, as μ tends to ∞. Upon letting μ tend to ∞ in the relations above, we obtain

$$(x_B^*)^2 \overline{w}_B - (|N|/q)x_B^* \in \text{range}(A_B), \qquad \overline{w}_B \in \text{range}(A_B^T).$$

But this implies that $\overline{w}_B$ is the (unique) optimal solution of the problem

$$\min_{w_B}\left\{\|x_B^* w_B - (|N|/q)\mathbf{1}_B\|^2; w_B \in \text{range}(A_B^T)\right\} \tag{4.5}$$

since the above relations are exactly the optimality conditions for problem (4.5). Moreover, using the fact that $(x_B^*)^{-1} \in \text{range}(A_B^T)$, one can easily see that the unique optimal solution of (4.5) is $(|N|/q)(x_B^*)^{-1}$. Hence, $\overline{w}_B = (|N|/q)(x_B^*)^{-1}$. Since this holds for any accumulation point of $w_B(\mu)$, as μ tends to ∞, it follows that $\lim_{\mu\to\infty} w_B(\mu) = (|N|/q)(x_B^*)^{-1}$. But this clearly implies our claim that $\lim_{\mu\to\infty}\theta_B(\mu) = 0$.

Step 3. Let $(r(\mu), w_N(\mu))$, for $\mu \in (\alpha^-, \infty)$, denote the (unique) optimal solution of the problem

$$\theta_N(\mu) = \min_{r, w_N} \|(z_N(\mu) - e^{-\mu}p_N)^{-1}w_N - (g(\mu)/h(\mu))\mathbf{1}_N\|^2$$

$$\text{s.t. } A_B^T r = c_B - h(\mu)w_B(\mu), \qquad A_N^T r + w_N = c_N.$$

Using Theorem 3.2, Lemma 3.1 and similar arguments as in the proof of Step 2, one can show that, as μ tends to ∞, $(r(\mu), w_N(\mu))$ and $\theta_N(\mu)$ converge, respectively, to the (unique) optimal solution (r^*, w_N^*) and optimal value θ_N^* of the problem

$$\min_{r, w_N}\left\{\|(z_N^*)^{-1}w_N - (|N|/q)\mathbf{1}_N\|^2;\ A_B^T r = c_B,\ A_N^T r + w_N = c_N\right\}. \tag{4.6}$$

From the optimality conditions of (4.6), one can verify that $(r^*, w_N^*) = (y^*, z_N^*)$. Hence, $\lim_{\mu\to\infty}\theta_N(\mu) = \theta_N^* = \|1 - (|N|/q)\mathbf{1}_N\|^2$.

Step 4. Let $s(\mu) \in \mathbf{R}^n$, for $\mu \in (\alpha^-, \infty)$, denote the point such that $s_B(\mu) = h(\mu)w_B(\mu)$ and $s_N(\mu) = w_N(\mu)$, where $w_B(\mu)$ and $w_N(\mu)$ are as defined in Steps 2 and 3, respectively. Then $(r(\mu), s(\mu))$ is feasible to (4.3). Hence,

$$\begin{aligned}
&h(\mu)^2\theta_B(\mu) + h(\mu)^2\theta_N(\mu) \\
&\quad = h(\mu)^2\|x_B(\mu)w_B(\mu) - (g(\mu)/h(\mu))\mathbf{1}_B\|^2 \\
&\qquad + h(\mu)^2\|(z_N(\mu) - e^{-\mu}p_N)^{-1}w_N(\mu) - (g(\mu)/h(\mu))\mathbf{1}_N\|^2 \\
&\quad = \|x_B(\mu)s_B(\mu) - g(\mu)\mathbf{1}_B\|^2 \\
&\qquad + h(\mu)^2\left\|(z_N(\mu) - e^{-\mu}p_N)^{-1}s_N(\mu) - \frac{g(\mu)}{h(\mu)}\mathbf{1}_N\right\|^2 \\
&\quad \geqslant \|x_B(\mu)z_B^E(\mu) - g(\mu)\mathbf{1}_B\|^2 \\
&\qquad + h(\mu)^2\left\|(z_N(\mu) - e^{-\mu}p_N)^{-1}z_N^E(\mu) - \frac{g(\mu)}{h(\mu)}\mathbf{1}_N\right\|^2.
\end{aligned} \tag{4.7}$$

Moreover, $z_B^E(\mu)/h(\mu)$ is clearly feasible to (4.4). Hence,

$$\begin{aligned}
h(\mu)^2\theta_B(\mu) &= h(\mu)^2\|x_B(\mu)w_B(\mu) - (g(\mu)/h(\mu))\mathbf{1}_B\|^2 \\
&\leqslant h(\mu)^2\|x_B(\mu)(z_B^E(\mu)/h(\mu)) - (g(\mu)/h(\mu))\mathbf{1}_B\|^2 \\
&= \|x_B(\mu)z_B^E(\mu) - g(\mu)\mathbf{1}_B\|^2.
\end{aligned} \tag{4.8}$$

Combining (4.7) and (4.8), we obtain that

$$\theta_N(\mu) \geqslant \|(z_N(\mu) - e^{-\mu}p_N)^{-1}z_N^E(\mu) - (g(\mu)/h(\mu))\mathbf{1}_N\|^2. \tag{4.9}$$

Step 5. In this step, we show that $\lim_{\mu\to\infty} z_B^E(\mu) = z_B^* = 0$. First, note that inequality (4.7) implies that

$$h(\mu)^2[\theta_B(\mu) + \theta_N(\mu)] \geqslant \|x_B(\mu)z_B^E(\mu) - g(\mu)\mathbf{1}_B\|. \tag{4.10}$$

Since, by Proposition 3.5 (b), $\lim_{\mu\to\infty} h(\mu) = 0$ and, by Steps 2 and 3 above, $\lim_{\mu\to\infty} \theta_B(\mu)$ and $\lim_{\mu\to\infty} \theta_N(\mu)$ are finite, it follows from (4.10) that $\lim_{\mu\to\infty} z_B^E(\mu)x_B(\mu) - g(\mu)\mathbf{1}_B = 0$. Hence, $\lim_{\mu\to\infty} z_B^E(\mu) = 0$ since, by Theorem 3.3, $\lim_{\mu\to\infty} x_B(\mu) = x_B^* > 0$ and, by Theorem 3.1, $\lim_{\mu\to\infty} g(\mu) = 0$.

Step 6. In this step, we show that $\lim_{\mu\to\infty}(y^E(\mu), z_N^E(\mu)) = (y^*, z_N^*)$. First, note that relation (4.9) implies that $z_N^E([0,\infty))$ is bounded, since, by Theorem 3.2,

$$\lim_{\mu\to\infty} (z_N(\mu) - e^{-\mu}p_N)^{-1} = (z_N^*)^{-1} > 0.$$

Since A has full row rank, it also follows that $y^E([0,\infty))$ is bounded. Let $(\bar{y}, \bar{z}_N)$ be an accumulation point of $(y^E(\mu), z_N^E(\mu))$ as μ tends to ∞. Obviously, in view of Step 5, $(\bar{y}, \bar{z}_N)$ is feasible to problem (4.6). Hence,

$$\theta_N^* \leqslant \|(z_N^*)^{-1}\bar{z}_N - (|N|/q)\mathbf{1}_N\|.$$

Moreover, upon letting μ tend to ∞ in relation (4.9), it follows that

$$\theta_N^* \geqslant \|(z_N^*)^{-1}\bar{z}_N - (|N|/q)\mathbf{1}_N\|.$$

Hence, $(\bar{y}, \bar{z}_N)$ is an optimal solution for problem (4.6) and therefore, $\bar{y} = y^*$ and $\bar{z}_N = z_N^*$. Since this holds for any accumulation point of $(y^E(\mu), z_N^E(\mu))$, Step 6 follows. Note that statement (a) follows from Steps 5 and 6.

Step 7. In this step, we show statement (b). First, note that inequality (4.7) implies that

$$\theta_B(\mu) + \theta_N(\mu) \geqslant \|x_B(\mu)(z_B^E(\mu)/h(\mu)) - (g(\mu)/h(\mu))\mathbf{1}_B\|^2$$
$$+\|(z_N(\mu) - e^{-\mu}p_N)^{-1}z_N^E(\mu) - (g(\mu)/h(\mu))\mathbf{1}_N\|^2.$$

This relation implies that $\{z_B^E(\mu)/h(\mu); \mu \in [0,\infty)\}$ is bounded. If $\tilde{z}_B$ is an accumulation point of $z_B^E(\mu)/h(\mu)$, as μ tends to ∞, the relation above also implies, upon letting μ tend to ∞, that

$$\theta_B^* + \theta_N^* \geqslant \|x_B^*\tilde{z}_B - (|N|/q)\mathbf{1}_B\|^2 + \theta_N^*,$$

or equivalently, that $\tilde{z}_B = (|N|/q)(x_B^*)^{-1}$, since $\theta_B^* = 0$ by Step 2. Since this holds for any accumulation point of $z_B^E(\mu)/h(\mu)$, Step 7 follows. □

Observe that Theorem 4.1 implies that $(y^E(x), z^E(x)) \in T_I$, for all points x along a PR trajectory which are sufficiently close to the center x^* of the face S_O. A similar conclusion for the case of the primal affine scaling trajectories does not seem to be true (see [1]).

Let (x,h): $(\alpha^-, +\infty) \to S_I \times \mathbf{R}_+$ denote the maximal solution of (3.2) and let (y,z): $(\alpha^-, +\infty) \to T_A$ denote its associated dual solution curve through $(y^0, z^0) \in T_A$. Recall, from the way the dual solution curve $(y(\cdot), z(\cdot))$ was defined, that for all

$\mu \in (\alpha^-, +\infty)$,

$$z^L(x(\mu)) = z(\mu) + \dot{z}(\mu). \tag{4.11}$$

When the optimal face S_O (T_O) consists of a unique vertex, the next theorem completely characterizes the asymptotic direction of convergence of the (associated dual solution curve of the) maximal solution of (3.2). We will see that relation (4.11) and Theorem 4.1 play a fundamental role in the proof of the next result.

THEOREM 4.2. *Let (x, h): $(\alpha^-, +\infty) \to S_I \times \mathbf{R}_+$ denote the maximal solution of (3.2) and let (y, z): $(\alpha^-, +\infty) \to T_A$ denote its associated dual solution curve through any point $(y^0, z^0) \in T_A$. Then*

$$\text{(a)} \qquad \lim_{\mu \to \infty} \frac{\dot{x}_N(\mu)}{h(\mu)} = \left(\frac{|N|}{q} - 1\right)(z_N^*)^{-1},$$

$$\text{(b)} \qquad \lim_{\mu \to \infty} \frac{\dot{z}_B(\mu)}{h(\mu)} = \left(\frac{|N|}{q} - 1\right)(x_B^*)^{-1}.$$

PROOF. We first show (a). Using relations (3.3a), (3.4) and (4.11), one can easily verify that

$$\frac{\dot{x}_N(\mu)}{h(\mu)} = \frac{g(\mu)}{h(\mu)}(z_N(\mu) - e^{-\mu}p_N)^{-1} - z_N^E(x(\mu))(z_N(\mu) - e^{-\mu}p_N)^{-2}.$$

Letting μ tend to ∞ in the expression above, then Lemma 3.1, Theorem 3.2 and Theorem 4.1 imply

$$\lim_{\mu \to \infty} \frac{\dot{x}_N(\mu)}{h(\mu)} = \frac{|N|}{q}(z_N^*)^{-1} - (z_N^*)(z_N^*)^{-2} = \left(\frac{|N|}{q} - 1\right)(z_N^*)^{-1}$$

from which (a) follows. To show (b), first note that by relation (3.4), Lemma 3.2 and Theorem 3.3

$$\lim_{\mu \to \infty} \frac{z_B(\mu)}{h(\mu)} = \lim_{\mu \to \infty} x_B(\mu)^{-1} + (e^{\mu}h(\mu))^{-1}p_B = (x_B^*)^{-1}$$

which together with relation (4.11) and statement (b) of Theorem 4.1 implies

$$\lim_{\mu \to \infty} \frac{\dot{z}_B(\mu)}{h(\mu)} = \lim_{\mu \to \infty} \frac{z_B^E(x(\mu)) - z_B(\mu)}{h(\mu)} = \left(\frac{|N|}{q} - 1\right)(x_B^*)^{-1}. \qquad \square$$

The following observations, whose verifications are left to the reader, are consequences of Theorem 4.2. When the optimal face S_O has dimension 0 (and, hence, A_B has full column rank) then $\lim_{\mu \to \infty} \dot{x}(\mu)/\|\dot{x}(\mu)\| = d^*/\|d^*\|$ where $d_N^* = -(z_N^*)^{-1}$ and $d_B^* = (A_B^T A_B)^{-1} A_B^T A_N (z_N^*)^{-1}$. Similarly, if the dimension of the optimal face T_O is 0 (and, hence, A_B has full row rank) then $\lim_{\mu \to \infty} \dot{z}(\mu)/\|\dot{z}(\mu)\| = \bar{d}/\|\bar{d}\|$ where $\bar{d}_B = -(x_B^*)^{-1}$ and $\bar{d}_N = -A_N^T(A_B A_B^T)^{-1} A_B (x_B^*)^{-1}$.

When the dimension of the optimal face T_O is greater that 0, the dual solution curves associated with the maximal solutions of the vector field $\Psi(x, h)$ still have a unique well-characterized asymptotic direction of convergence as the next theorem shows.

THEOREM 4.3. *Let* (x, h): $(\alpha^-, +\infty) \to S_I \times \mathbf{R}_+$ *denote the maximal solution of* (3.2) *and let* (y, z): $(\alpha^-, +\infty) \to T_A$ *denote its associated dual solution curve through any point* $(y^0, z^0) \in T_A$. *Then the limit* $\lim_{\mu\to\infty} \dot{z}_N(\mu)/h(\mu)$ *exists and is equal to the* (*unique*) *optimal solution of the problem*

$$\begin{aligned}(4.12)\qquad &\min_{r, s_N} \|(z_N^*)^{-1} s_N\|^2\\ &\text{s.t. } A_B^T r = (1 - |N|/q)(x_B^*)^{-1}, \quad A_N^T r + s_N = 0.\end{aligned}$$

PROOF. Let $(r(\mu), s(\mu)) \equiv h(\mu)^{-1}(\dot{y}(\mu), \dot{z}(\mu))$ for $\mu \in (\alpha^-, +\infty)$. We first show that $(r(\mu), s_N(\mu))$ is the (unique) optimal solution of

$$\begin{aligned}(4.13)\qquad &\min_{r, s_N} \left\|(z_N(\mu) - e^{-\mu} p_N)^{-1}\left[s_N + (e^{\mu} h(\mu))^{-1} p_N\right]\right\|^2\\ &\text{s.t. } A_B^T r = -s_B(\mu), \qquad A_N^T r + s_N = 0.\end{aligned}$$

Indeed, it is enough to verify that $(r(\mu), s_N(\mu))$ satisfy the optimality conditions for (4.13), that is, we have to show that

$$(4.14a)\quad A_N(z_N(\mu) - e^{-\mu} p_N)^{-2}\left[s_N(\mu) + (e^{\mu} h(\mu))^{-1} p_N\right] \in \text{range}(A_B),$$

$$(4.14b)\qquad A_B^T r(\mu) = -s_B(\mu),$$

$$(4.14c)\qquad A_N^T r(\mu) + s_N(\mu) = 0.$$

The equalities (4.14b) and (4.14c) are equivalent to relation (3.3d). Using (3.3a) and (3.4), one can easily verify that

$$\begin{aligned}(4.15)\qquad &(z_N(\mu) - e^{-\mu} p_N)^{-2}\left[s_N(\mu) + (e^{\mu} h(\mu))^{-1} p_N\right]\\ &= \frac{1}{h(\mu)}(z_N(\mu) - e^{-\mu} p_N)^{-2}(\dot{z}_N(\mu) + e^{-\mu} p_N)\\ &= \frac{1}{h(\mu)^3}\left[(g(\mu) - h(\mu)) x_N(\mu) - h(\mu) \dot{x}_N(\mu)\right].\end{aligned}$$

Moreover, $A_N \dot{x}_N(\mu) = -A_B \dot{x}_B(\mu) \in \text{range}(A_B)$ and $A x_N(\mu) = b - A_B x_B(\mu) \in \text{range}(A_B)$ since $b \in \text{range}(A_B)$. Relation (4.15) and this last observation shows that (4.14a) holds. One can now easily verify that, since $(r(\mu), s_N(\mu))$ is optimal for (4.13), the set $\{(r(\mu), s_N(\mu));\ \mu \in (\alpha^-, +\infty)\}$ is bounded. Moreover, from Lemma 3.2, Theorem 3.2, statement (b) of Theorem 4.2 and the relations in (4.14), it follows that, if $(\bar{r}, \bar{s}_N)$ is an accumulation point of $(r(\mu), s_N(\mu))$, as μ tends to ∞, then

$$A_N(z_N^*)^{-2} \bar{s}_N \in \text{range}(A_B), \quad A_B^T \bar{r} = (1 - |N|/q)(x_B^*)^{-1}, \quad A_N^T \bar{r} + \bar{s}_N = 0.$$

Since these relations are the optimality conditions for the unique optimal solution of problem (4.12), the result follows. □

For convenience, we have stated Theorem 4.2 and Theorem 4.3 in terms of the limit $\lim_{\mu\to\infty} \dot{z}(\mu)/h(\mu)$ but one can easily modify the statements of these theorems in terms of the limit $\lim_{\mu\to\infty} \dot{z}(\mu)/\|\dot{z}(\mu)\|$.

The analysis of the asymptotic direction of convergence of the maximal solution of (3.2) for the case in which the dimension of S_O is greater than 0 is more involved. We first state and prove three preliminary lemmas.

LEMMA 4.1. *Let* $(x,h)\colon (\alpha^-, +\infty) \to S_I \times \mathbf{R}_+$ *denote the maximal solution of* (3.2). *Then the limit* $\lim_{\mu\to\infty}[c^Tx(\mu) - |N|h(\mu)]/h(\mu)^2$ *exists and is finite.*

PROOF. Let $(y,z)\colon(\alpha^-,\infty)\to T_A$ denote the associated dual solution curve through any point $(y^0, z^0) \in T_A$. Then, using relation (3.4) and the remarks at the end of subsection 2.2, we obtain

$$\frac{c^Tx(\mu) - |N|h(\mu)}{h(\mu)^2} = \frac{(z_N^*)^T x_N(\mu) - |N|h(\mu)}{h(\mu)^2}$$

$$= \frac{(z_N^*)^T(z_N(\mu) - p_N e^{-\mu})^{-1} - |N|}{h(\mu)}$$

$$= \frac{[z_N^* - (z_N(\mu) - p_N e^{-\mu})]^T(z_N - p_N e^{-\mu})^{-1}}{h(\mu)}$$

$$= \left[\frac{z_N^* - z_N(\mu)}{h(\mu)}\right]^T (z_N(\mu) - p_N e^{-\mu})^{-1}$$

$$+ [e^{\mu}h(\mu)]^{-1} p_N^T (z_N - p_N e^{-\mu})^{-1}$$

$$= -(z_N^*)^{-1} \lim_{\mu\to\infty} \frac{\dot z_N(\mu)}{\dot h(\mu)},$$

where the last equality follows from the L'Hopital rule, Theorem 3.2 and Lemma 3.2. Also, note that the limit in the last expression above exists due to Theorem 4.3, Lemma 3.1 and relation (3.3b). □

LEMMA 4.2. *Let* $(x,h)\colon (\alpha^-, +\infty) \to S_I \times \mathbf{R}_+$ *denote the maximal solution of* (3.2). *Then*

(a) *if* $q > 2|N|$ *then* $\lim_{\mu\to\infty} h(\mu)^2 e^{\mu} = 0$,
(b) *if* $q < 2|N|$ *then* $\lim_{\mu\to\infty} h(\mu)^2 e^{\mu} = \infty$,
(c) *if* $q = 2|N|$ *then* $\lim_{\mu\to\infty} h(\mu)^2 e^{\mu} = l$ *for some* $l \in \mathbf{R}_+$.

The proof of Lemma 4.2 is an immediate consequence of the following stronger result.

LEMMA 4.3. *Let* $(x,h)\colon (\alpha^-, +\infty) \to S_I \times \mathbf{R}_+$ *denote the maximal solution of* (3.2). *Let* $\gamma = |N|/q$. *Then* $\lim_{\mu\to\infty} h(\mu)e^{(1-\gamma)\mu}$ *exists and is positive.*

PROOF. Let $g(\mu) = c^Tx(\mu)/q$ and let $\rho(\mu) \equiv [g(\mu) - \gamma h(\mu)]/h(\mu)^2$ for all $\mu \in (\alpha^-, +\infty)$. It follows from Lemma 4.1 that $\lim_{\mu\to\infty}\rho(\mu)$ exists and is finite. In particular, this implies that there exists $M > 0$ such that

(4.16) $$|\rho(\mu)| \leqslant M \quad \text{for all } \mu \geqslant 0.$$

We make the following claim.

Claim. $e^{(1-\gamma)\mu}h(\mu) = e^{\int_0^\mu \rho(s)h(s)\,ds}$ for all $\mu \geq 0$.

Indeed, from Proposition 3.1 (a) and the definition of $\rho(\cdot)$ above, it follows that

$$e^{\mu}h(\mu) = 1 + \int_0^{\mu} g(s)e^{s}\,ds = 1 + \int_0^{\mu}[\gamma + \rho(s)h(s)]e^{s}h(s)\,ds.$$

Hence, by Proposition 2.4, it follows that

$$e^{\mu}h(\mu) = e^{\int_0^{\mu}[\gamma+\rho(s)h(s)]\,ds} = e^{\gamma\mu}e^{\int_0^{\mu}\rho(s)h(s)\,ds}$$

which implies our claim. In view of this claim, we only need to show that $\lim_{\mu\to\infty}\int_0^{\mu}\rho(s)h(s)\,ds$ exists and is finite. We first show that

$$\int_0^{\infty} h(s)\,ds \equiv \lim_{\mu\to\infty}\int_0^{\mu} h(s)\,ds < \infty. \tag{4.17}$$

Indeed, let ϵ be such that $\epsilon \in (0, 1-\gamma)$. (Note that $0 < \gamma < 1$, by relation (2.11).) Since, by Proposition 3.5(b), $\lim_{\mu\to\infty} h(\mu) = 0$, it follows that there exists $\mu^0 > 0$ such that $|\rho(\mu)|h(\mu) \leq \epsilon$ for all $\mu \geq \mu^0$. Hence, using the claim, it is easy to see that $e^{(1-\gamma)\mu}h(\mu) \leq C(\mu^0)e^{\epsilon\mu}$ for all $\mu \geq \mu^0$, where

$$C(\mu^0) \equiv \exp\left[\int_0^{\mu^0}\rho(s)h(s)\,ds - \epsilon\mu^0\right].$$

This implies that $h(\mu) \leq C(\mu^0)e^{(\gamma+\epsilon-1)\mu}$ and since $\gamma + \epsilon - 1 < 0$, we obtain that $\int_0^{\infty}h(s)\,ds < \infty$. We next show that

$$\int_0^{\infty}\rho(s)h(s)\,ds \equiv \lim_{\mu\to\infty}\int_0^{\mu}\rho(s)h(s)\,ds$$

exists and is finite. Indeed, the function $\mu \to \rho(\mu)h(\mu)$ is absolutely integrable since (4.16) and (4.17) imply that

$$\int_0^{\infty}|\rho(s)h(s)|ds \leq M\int_0^{\infty}h(s)\,ds < \infty.$$

Hence, $\int_0^{\infty}\rho(s)h(s)\,ds$ exists and is finite. □

We are now in a position to state the main result of this section, which shows the behavior of the asymptotic direction of convergence of the PR trajectories for three possible ranges of values of the parameter q.

THEOREM 4.4. *Let* (x, h): $(\alpha^-, +\infty) \to S_I \times \mathbf{R}_+$ *denote the maximal solution of* (3.2). *Then*

(1) *if* $q < 2|N|$ *then* $\lim_{\mu\to\infty}\dot{x}(\mu)/\|\dot{x}(\mu)\| = d^*/\|d^*\|$, *where* $d_N^* = -(z_N^*)^{-1}$ *and* d_B^* *is the unique optimal solution of the problem*

$$\min_{d_B}\left\{\|(x_B^*)^{-1}d_B\|^2;\ A_B d_B = A_N(z_N^*)^{-1}\right\},$$

(2) *if* $q > 2|N|$ *then we have one of the following possibilities*:

(a) *if* $(x_B^0)^{-1} \in \operatorname{range}(A_B^T)$ *then* $\lim_{\mu\to\infty}\dot{x}(\mu)/\|\dot{x}(\mu)\| = d^*/\|d^*\|$, *where* d^* *is as in* (1) *above*;

(b) *if* $(x_B^0)^{-1} \notin \text{range}(A_B^T)$ *then* $\lim_{\mu\to\infty} \dot{x}(\mu)/\|\dot{x}(\mu)\| = \hat{d}/\|\hat{d}\|$, *where* $\hat{d}_N = 0$ *and* $\hat{d}_B$ *is the unique optimal solution of the problem*

$$(4.18) \qquad \min_{d_B}\left\{-\left[\left(x_B^0\right)^{-1}\right]^T d_B + \tfrac{1}{2}\|(x_B^*)^{-1}d_B\|^2;\ A_B d_B = 0\right\}$$

(*note that* $\hat{d}_B \neq 0$ *since* $(x_B^0)^{-1} \notin \text{range}(A_B^T)$),

(3) *if* $q = 2|N|$ *then* $\lim_{\mu\to\infty} \dot{x}(\mu)/\|\dot{x}(\mu)\| = \bar{d}/\|\bar{d}\|$, *where* $\bar{d}_N = -(z_N^*)^{-1}$ *and* $\bar{d}_B$ *is the unique optimal solution of the problem*

$$\min_{d_B}\left\{-l(x^0)^{-1}\frac{|N|}{q-|N|}\left[\left(x_B^0\right)^{-1}\right]^T d_B + \frac{1}{2}\|(x_B^*)^{-1}d_B\|^2;\ A_B d_B = A_N(z_N^*)^{-1}\right\},$$

where $l(x^0) \equiv \lim_{\mu\to\infty} h(\mu)^2 e^{\mu}$.

PROOF. Let $(y,z)\colon(\alpha^-,\infty)\to T_A$ denote the associated dual solution curve through any point $(y^0,z^0)\in T_A$. Using relations (3.3a), (3.4) and the fact that $p_B = z_B^0 - (x_B^0)^{-1}$, one can easily see that

$$h(\mu)x_B(\mu)^{-2}\dot{x}_B(\mu) - \frac{g(\mu)}{h(\mu)}e^{-\mu}\left(x_B^0\right)^{-1}$$
$$= \left(\frac{g(\mu)}{h(\mu)} - 1\right)z_B(\mu) - \dot{z}_B(\mu) - \frac{g(\mu)}{h(\mu)}e^{-\mu}z_B^0,$$

where $g(\mu) \equiv c^T x(\mu)/q$. Since the vectors $z_B(\mu)$, $\dot{z}_B(\mu)$ and z_B^0 all lie in the subspace $\text{range}(A_B^T)$, it follows that

$$(4.19) \qquad h(\mu)x_B(\mu)^{-2}\dot{x}_B(\mu) - \frac{g(\mu)}{h(\mu)}e^{-\mu}\left(x_B^0\right)^{-1} \in \text{range}\left(A_B^T\right).$$

We now show statement (1). Let $d(\mu) \equiv \dot{x}(\mu)/h(\mu)$ for $\mu \in (\alpha^-,\infty)$. Then (4.19) implies that

$$(4.20) \qquad x_B(\mu)^{-2}d_B(\mu) - j(\mu)\left(x_B^0\right)^{-1} \in \text{range}\left(A_B^T\right)$$

where $j(\mu) \equiv g(\mu)[h(\mu)^2e^{\mu}]^{-1}/h(\mu)$. Expression (4.20) implies that $d_B(\mu)$ is the unique optimal solution of the problem

$$\min_{d_B}\left\{-j(\mu)\left[\left(x_B^0\right)^{-1}\right]^T d_B + \tfrac{1}{2}\|x_B(\mu)^{-1}d_B\|^2;\ A_B d_B = -A_N d_N(\mu)\right\}.$$

Moreover, from Lemma 3.1, Lemma 4.2(b), the definition of $g(\mu)$ above and Theorem 4.2(a), it follows that $\lim_{\mu\to\infty} j(\mu) = 0$ and $\lim_{\mu\to\infty} d_N(\mu) = [(|N|/q) - 1](z_N^*)^{-1}$. Now, arguing like in the proof of Theorem 4.3 and using the statements above, it follows that, as μ tends to ∞, $d_B(\mu)$ converges to the optimal solution of the problem

$$\min_{d_B}\left\{\left\|(x_B^*)^{-1}d_B\right\|^2;\ A_B d_B = (1-|N|/q)A_N(z_N^*)^{-1}\right\}.$$

Since $\dot{x}(\mu)/\|\dot{x}(\mu)\| = d(\mu)/\|d(\mu)\|$, it is easy to verify that statement (1) holds. The proof of statement (2)(a) is essentially the same as the proof of (1) above if one lets

$j(\mu) \equiv 0$ in all the steps of the proof of (1). We just observe that, since $(x_B^0)^{-1} \in \text{range}(A_B^T)$ then (4.19) implies that $x_B(\mu)^{-2}d_B(\mu) \in \text{range}(A_B^T)$. We next show statement (2)(b). Let $r(\mu) \equiv [h(\mu)^2 e^\mu / g(\mu)]\dot{x}(\mu)$ for all $\mu \in (\alpha^-, +\infty)$. Relation (4.19) implies that

$$-\left(x_B^0\right)^{-1} + x_B(\mu)^{-2} r_B(\mu) \in \text{range}\left(A_B^T\right)$$

from which it follows that $r_B(\mu)$ is the unique optimal solution of the problem

$$\min_{r_B}\left\{-\left[\left(x_B^0\right)^{-1}\right]^T r_B + \tfrac{1}{2}\|x_B(\mu)^{-1} r_B\|^2;\ A_B r_B = -A_N r_N(\mu)\right\}.$$

Moreover, using Theorem 4.2(a), Lemma 4.2(a) and Lemma 3.1 and the definition of $g(\mu)$ above, one can easily verify that $\lim_{\mu\to\infty} r_N(\mu) = 0$. Using this fact, one can show that, as μ tends to ∞, $r_B(\mu)$ converges to the unique optimal solution of problem (4.18). Since $\dot{x}(\mu)/\|\dot{x}(\mu)\| = r(\mu)/\|r(\mu)\|$, the statement follows. The proof of statement (3) follows by using Lemma 4.2(c) and the same type of argument as above. We leave the details to the reader. □

Theorem 4.4 can be described in words as follows. For $q < 2|N|$, all PR trajectories approach the center of the primal optimal face along the same asymptotic direction of convergence. Since the primal central trajectory is a special PR trajectory, we can alternatively say that, for $q < 2|N|$, all PR trajectories become asymptotically tangent to the central trajectory. For $q = 2|N|$, the asymptotic direction of convergence varies according to the initial point x^0, but none of these directions are tangent to the optimal face. (Note that the direction d is tangent to the optimal face if and only if $d_N = 0$ and $A_B d_B = 0$.) For $q > 2|N|$, there are two possible cases depending on the initial point x^0. If $(x_B^0)^{-1} \in \text{range}(A_B^T)$ then the PR trajectory through x^0 becomes asymptotically tangent to the central trajectory. If $(x_B^0)^{-1} \notin \text{range}(A_B^T)$ then the PR trajectory through x^0 approaches the center of the primal optimal face tangentially to this face and the direction of approach still depends on the initial point x^0.

Finally, we mention that Theorem 4.4 generalizes our previous result for the case in which the optimal face consists of a unique vertex. Note that in this case, we always have $(x_B^0)^{-1} \in \text{range}(A_B^T)$, and therefore, all three cases of Theorem 4.4 reduce to the same case, that is, all PR trajectories become asymptotically tangent to the central trajectory.

5. Appendix. In this appendix, we give the proofs of Propositions 2.2 and 2.3. We should mention that some of the notation used in here is as defined in §2.1. So, the reader is referred to that subsection for those terms which are not defined explicitly in here.

We start by stating Proposition 2.3 in a more general version which does not assume that the function $g: (\alpha, \beta) \to \mathbf{R}$ is strictly positive. Proposition 2.3 can be easily seen to be a consequence of the following result.

PROPOSITION A.1. *Let $x: (w^-, w^+) \to U$ denote the maximal solution of the Cauchy problem (2.3). Let $\gamma: (\alpha, \beta) \to \mathbf{R}$ be the function defined by $\gamma(t) \equiv t^0 + \int_{t_0}^t g(s)\,ds$ for all $t \in (\alpha, \beta)$. Consider the interval (α^-, α^+) where*

$$\alpha^- = \inf\left\{t \in (\alpha, \beta) | t \leqslant t^0; \gamma\left((t, t^0]\right) \subseteq (w^-, w^+)\right\},$$

$$\alpha^+ = \sup\left\{t \in (\alpha, \beta) | t \geqslant t^0; \gamma\left([t^0, t)\right) \subseteq (w^-, w^+)\right\}.$$

Then the maximal solution of the Cauchy problem (2.3) *is the function* $\psi: (\alpha^-, \alpha^+) \to U$ *defined by* $\psi(t) = x(\gamma(t))$ *for all* $t \in (\alpha^-, \alpha^+)$.

PROOF. One can easily verify that $\psi(\cdot)$ is well defined and that it is a solution (2.3). We have only to show that $\psi(\cdot)$ is maximal. Assume by contradiction that $\psi(\cdot)$ is not a maximal solution of (2.3) and, without loss of generality, assume that $\psi(\cdot)$ could be extended beyond the right endpoint α^+ of its interval of definition. This clearly implies that $\alpha^+ < \beta$ and that $\psi([t^0, \alpha^+)) \subseteq K$ where K is a compact subset of U. The definition of α^+ and the fact that $\alpha^+ < \beta$ imply that either

(i) $[t^0, w^+) \subseteq \gamma([t^0, \alpha^+))$ and $w^+ < \infty$, or

(ii) $(w^-, t^0] \subseteq \gamma([t^0, \alpha^+))$ and $w^- > -\infty$.

Assume without loss of generality that (i) holds. Then it follows from the definition of $\psi(\cdot)$ and the above arguments that $x([t^0, w^+)) \subseteq \psi([t^0, \alpha^+)) \subseteq K$. But then Proposition 2.1 applied to the maximal solution $x(\cdot)$ implies that $w^+ = \infty$ which contradicts the second statement of case (i) above. This completes the proof of the proposition. □

We now turn our efforts toward showing Proposition 2.2. Given $\bar{x} \in U$, let $\phi(t, \bar{x})$ denote the maximal solution of the Cauchy problem $\dot{x} = \Phi(x)$, $x(0) = \bar{x}$ and let $J(\bar{x})$ denote the open interval over which this solution is defined. Also, let

$$Q = \{(t, \bar{x}) \in \mathbf{R} \times U | \bar{x} \in U; t \in J(\bar{x})\}.$$

The *flow* associated with $\dot{x} = \Phi(x)$ is the mapping $\phi: Q \to U$ as defined above but now viewed as a function of the pair $(t, \bar{x}) \in Q$. We also write $\phi(t, \bar{x}) = \phi_t(\bar{x})$. The following lemma, whose proof may be found in Hirsch and Smale [7, p. 175, Theorems 1 and 2], is useful in the proof of Proposition 2.2.

LEMMA A.1. *The following properties of the flow* $\phi: Q \to U$ *defined above holds*:

(a) *if* $(t, \bar{x}) \in Q$ *then* $J(\phi_t(\bar{x})) = \{r - t | r \in J(\bar{x})\}$ *and* $\phi_{t+s}(\bar{x}) = \phi_s(\phi_t(\bar{x}))$ *for all* $s \in J(\phi_t(\bar{x}))$,

(b) *the domain of definition* Q *of the flow* $\phi(\cdot)$ *is an open subset of* $\mathbf{R} \times U$. *Moreover, the flow* $\phi(\cdot)$ *is continuous.*

We are now ready to prove Proposition 2.2.

PROOF OF PROPOSITION 2.2. Let a maximal solution $x: (w^-, w^+) \to U$ of $\dot{x} = \Phi(x)$ and a compact subset K of U be given. Assume, without loss of generality, that $0 \in (w^-, w^+)$ and let $\bar{x} = x(0)$. By definition of the flow $\phi(\cdot)$, we have $\phi_t(\bar{x}) = x(t)$ for all $t \in (w^-, w^+) = J(\bar{x})$. Assume by contradiction that there does not exist $\beta \in (w^-, w^+)$ such that $\phi_t(\bar{x}) \notin K$ for all $t \in [\beta, w^+)$. Then there exists a subsequence $(t^k) \subseteq (w^-, w^+)$ such that $t^k \to w^+$, as k tends to ∞, and $\phi_{t^k}(\bar{x}) \in K$ for all k (note that, by Proposition 2.1, this implies that $w^+ = +\infty$). Since K is compact, we may assume, without loss of generality, that $\lim_{k \to \infty} \phi_{t^k}(\bar{x}) = \bar{v}$ for some $\bar{v} \in K$. Since, by the assumptions, $t \to W(\phi_t(\bar{x}))$ is a decreasing function and W is continuous, it follows that $W(\bar{v}) \leqslant W(\phi_t(\bar{x}))$ for all $t \in J(\bar{x})$. On the other hand, consider the maximal solution $t \to \phi_t(\bar{v})$. Again, from the assumptions, it follows that $W(\phi_t(\bar{v})) < W(\phi_0(\bar{v})) = W(\bar{v})$ for all $t \in J(\bar{v})$ with $t > 0$. Fix some $\bar{t} \in J(\bar{v})$ with $\bar{t} > 0$. Then $W(\phi_{\bar{t}}(\bar{v})) < W(\bar{v})$. Since, by Lemma A.1(b), $\phi: Q \to U$, and hence $W \circ \phi: Q \to \mathbf{R}$, is continuous and $(\bar{t}, \bar{v}) \in Q$, it follows that $W(\phi_{\bar{t}}(x)) < W(\bar{v})$ for any x in some neighborhood of $\bar{v}$ contained in Q (note that here we are using the fact that Q is an open set in view of Lemma A.1(b)). In particular, since $\lim_{k \to \infty} \phi_{t^k}(\bar{x}) = \bar{v}$, we have $W(\phi_{\bar{t}}(\phi_{t^k}(\bar{x}))) < W(\bar{v})$ for all k sufficiently large. By Lemma A.1(a) we have $\phi^k_{\bar{t}+t}(\bar{x}) = \phi_{\bar{t}}(\phi_{t^k}(\bar{x}))$ for all k sufficiently large, and hence, it follows that $W(\phi_{\bar{t}+t^k}(\bar{x})) < W(\bar{v})$ for all k sufficiently large. But this contradicts the fact that $W(\bar{v}) \leqslant W(\phi_t(\bar{x}))$ for all $t \in J(\bar{x})$. □

6. Acknowledgements. The author wishes to thank Richard Stone for many useful discussions which have been invaluable toward the development of this work and to express his gratitude to the two anonymous referees for their patient reading of a prior version of this manuscript and for their very useful comments. Thanks are also due to members of AT & T Advanced Decision Support Systems for providing the motivation for the work presented in this paper and, in particular, to Jun-Min Liu, Mauricio G. C. Resende and Aydin Ulkucu.

References

[1] Adler, I. and Monteiro, R. D. C. (1991). Limiting Behavior of the Affine Scaling Continuous Trajectories for Linear Programming Problems. *Math. Programming* **50** 29–51.

[2] Barnes, E. R (1986). A Variation on Karmarkar's Algorithm for Solving Linear Programming Problems. *Math Programming* **36** 174–182.

[3] Bayer, D. A. and Lagarias, J. C. (1989). The Nonlinear Geometry of Linear Programming. Part I: Affine and Projective Scaling Trajectories. Part II: Legendre Transform Coordinates and Central Trajectories. *Trans. Amer Math. Soc.* **314** 499–581.

[4] Dikin. I. I (1967). Iterative Solution of Problems of Linear and Quadratic Programming. *Soviet Math. Dokl* **8** 674–675.

[5] Fiacco, A. and McCormick, G. (1968). *Nonlinear Programming· Sequential Unconstrained Minimization Techniques.* John Wiley and Sons, New York

[6] Gonzaga, C. (1988). Polynomial Affine Algorithms for Linear Programming. Report ES-139/88, COPPE, Universidade Federal do Rio de Janeiro, Rio de Janeiro, Brazil.

[7] Hirsch. M. W. and Smale, S. (1974). *Differential Equations, Dynamical Systems, and Linear Algebra.* Academic Press, New York.

[8] Karmarkar, N. (1984). A New Polynomial Time Algorithm for Linear Programming. *Combinatorica* **4** 373–395.

[9] Megiddo, N. (1989). Pathways to the Optimal Set in Linear Programming, in N. Megiddo (Ed.) *Progress in Mathematical Programming: Interior-Point Algorithms and Related Methods.* Springer-Verlag Berlin and New York, 131–158.

[10] ______ and Shub, M. (1989). Boundary Behavior of Interior Point Algorithms for Linear Programming. *Math Oper. Res.* **14** 97–146.

[11] Monteiro, R. D C. (1988). Convergence and Boundary Behavior of the Projective Scaling Trajectories for Linear Programming, in J. C. Lagarias and M. J Todd (Eds.), *Proc. AMS-IME-SIAM Res. Conf. on Math. Developments Arising from Linear Programming.* Bowdoin College, Brunswick, ME.

[12] ______, Adler, I. and Resende, M. G. C. (1990). A Polynomial-Time Primal-Dual Affine Scaling Algorithm for Linear and Convex Quadratic Programming and its Power Series Extension. *Math. Oper. Res.* **15** 191–214.

[13] Schrijver, A. (1986). *Theory of Linear and Integer Programming* John Wiley and Sons, New York.

[14] Shub, M. (1987). On the Asymptotic Behavior of the Projective Rescaling Algorithm for Linear Programming. *J Complexity* **3** 258–269.

[15] Vanderbei, R J., Meketon, M. S. and Freedman, B. A. (1986). A Modification of Karmarkar's Linear Programming Algorithm *Algorithmica* **1** 395–407.

[16] Witzgall, C., Boggs, P. and Domich, P. (1988). On the Convergence Behavior of Trajectories for Linear Programming, in J. C. Lagarias and M. J. Todd (Eds.), *Proc. AMS-IME-SIAM Res. Conf. on Math. Developments Arising from Linear Programming.* Bowdoin College, Brunswick, ME.

SIE DEPARTMENT, UNIVERSITY OF ARIZONA, TUCSON, ARIZONA 85721

AT & T BELL LABORATORIES, CRAWFORDS CORNER ROAD, HOLMDEL, NEW JERSEY 07733