

THE NONLINEAR GEOMETRY OF LINEAR PROGRAMMING. II LEGENDRE TRANSFORM COORDINATES AND CENTRAL TRAJECTORIES

D. A. BAYER AND J. C. LAGARIAS

ABSTRACT. Karmarkar's projective scaling algorithm for solving linear programming problems associates to each objective function a vector field defined in the interior of the polytope of feasible solutions of the problem. This paper studies the set of trajectories obtained by integrating this vector field, called *P-trajectories*, as well as a related set of trajectories, called *A-trajectories*. The *A-trajectories* arise from another linear programming algorithm, the affine scaling algorithm. The affine and projective scaling vector fields are each defined for linear programs of a special form, called *standard form* and *canonical form*, respectively.

These trajectories are studied using a nonlinear change of variables called *Legendre transform coordinates*, which is a projection of the gradient of a logarithmic barrier function. The Legendre transform coordinate mapping is given by rational functions, and its inverse mapping is algebraic. It depends only on the constraints of the linear program, and is a one-to-one mapping for canonical form linear programs. When the polytope of feasible solutions is bounded, there is a unique point mapping to zero, called the *center*.

The *A-trajectories* of standard form linear programs are linearized by the Legendre transform coordinate mapping. When the polytope of feasible solutions is bounded, they are the complete set of geodesics of a Riemannian geometry isometric to Euclidean geometry. Each *A-trajectory* is part of a real algebraic curve.

Each *P-trajectory* for a canonical form linear program lies in a plane in Legendre transform coordinates. The *P-trajectory* through $\mathbf{0}$ in Legendre transform coordinates, called the *central P-trajectory*, is part of a straight line, and is contained in the *A-trajectory* through $\mathbf{0}$, called the *central A-trajectory*. Each *P-trajectory* is part of a real algebraic curve.

The central *A-trajectory* is the locus of centers of a family of linear programs obtained by adding an extra equality constraint of the form $\langle \mathbf{c}, \mathbf{x} \rangle = \mu$. It is also the set of minima of a parametrized family of logarithmic barrier functions. Power-series expansions are derived for the central *A-trajectory*, which is also the central *P-trajectory*. These power-series have a simple recursive form and are useful in developing "higher-order" analogues of Karmarkar's algorithm.

A-trajectories are defined for a general linear program. Using this definition, it is shown that the limit point \mathbf{x}_∞ of a central *A-trajectory* on the boundary of the feasible solution polytope P is the center of the unique face of P containing \mathbf{x}_∞ in its relative interior.

Received by the editors October 8, 1986 and, in revised form, June 9, 1987 and March 25, 1988.
1980 *Mathematics Subject Classification* (1985 Revision). Primary 90C05; Secondary 52A40, 34A34.

The first author was partially supported by ONR contract N00014-87-K0214.

The central trajectory of a combined primal-dual linear program has a simple set of polynomial relations determining it as an algebraic curve. These relations are a relaxed form of the complementary slackness conditions. This central trajectory algebraically projects onto the central trajectories of both the primal and dual linear programs, and this gives an algebraic correspondence between points on the positive parts of the central trajectories of the primal and dual linear programs.

Two Lagrangian dynamical systems with simple Lagrangians are shown to have A -trajectories as q -trajectories. The Hamiltonian dynamical systems associated to these Lagrangian systems are completely integrable.

1. INTRODUCTION

In 1984 Narendra Karmarkar [K] introduced a new linear programming algorithm which he proved ran in polynomial time in the worst case. This algorithm, which we call the *projective scaling algorithm*, takes a series of piecewise linear steps in the relative interior of the polytope of feasible solutions of the linear programming problem. The step direction is computed using a vector field defined on the relative interior of the polytope of feasible solutions. This vector field, which we call the *projective scaling vector field*, depends on the linear program's constraints and on the objective function. The projective scaling vector field is defined for linear programs of a special form which we call *canonical form*, and Karmarkar uses projective transformations to compute this vector field direction.

Our viewpoint is that a fundamental object underlying the projective scaling algorithm is the set of trajectories obtained by integrating this vector field, which we call *projective scaling trajectories*, or *P-trajectories*, and that the polynomial-time nature of Karmarkar's algorithm arises from special geometric properties of these trajectories.

The projective scaling algorithm provides one method of approximately following these trajectories. Several authors [Re, V, Go, KMY] have recently developed other linear programming algorithms that follow these trajectories in a different manner.

This series of papers studies the *P-trajectories* and a related family of curves which we call *affine scaling trajectories* or *A-trajectories* for short. The affine scaling trajectories are associated to another vector field, the *affine scaling vector field*, which arises in connection with another interior-point linear programming method, the *affine scaling algorithm*, which was originally proposed by I. I. Dikin [D1, D2] in 1967, and rediscovered by many people including [B, VMF].

In part I we defined the affine and projective scaling vector fields and showed that *P-trajectories* of a *canonical form linear program*:

$$\begin{cases} \text{minimize } \langle \mathbf{c}, \mathbf{x} \rangle, \\ A\mathbf{x} = \mathbf{0}, \\ \langle \mathbf{e}, \mathbf{x} \rangle = n, \\ \mathbf{x} \geq \mathbf{0}, \end{cases}$$

having $\mathbf{e} = (1, 1, \dots, 1)^T$ as a feasible solution were the *radial projections* of A -trajectories of the *homogeneous standard form linear program*:

$$\begin{cases} \text{minimize } \langle \mathbf{c}, \mathbf{x} \rangle, \\ A\mathbf{x} = \mathbf{0}, \\ \mathbf{x} \geq \mathbf{0}, \end{cases}$$

whose polytope of feasible solutions is a cone. Consequently P -trajectories can be studied by studying the A -trajectories of a related linear programming problem.

This paper studies these trajectories using a nonlinear change of coordinates which we call *Legendre transform coordinates*. Legendre transform coordinates are given by a projection of the gradient of a logarithmic barrier function determined by the linear program's constraints. We show that Legendre transform coordinates linearize the A -trajectories. As a consequence when the polytope \mathbf{P} of feasible solutions is bounded the set of A -trajectories for all objective functions forms a complete set of geodesics for a Riemannian geometry on the relative interior of \mathbf{P} which is isometric to Euclidean geometry. P -trajectories are partially linearized by Legendre transform coordinates in the sense that each P -trajectory lies in a plane in Legendre transform coordinates. (This is explained further in part III, where it is shown that P -trajectories are linearized by a related mapping, *projective Legendre transform coordinates*.) The Legendre transform coordinate mapping is a rational mapping, and consequently the A -trajectories are parts of real algebraic curves. P -trajectories are also proved to be parts of real algebraic curves. We use Legendre transform coordinates to derive power-series expansions for A -trajectories.

Associated to any set of constraints H having a bounded polytope P_H of feasible solutions is a special point \mathbf{x}_H which we call the *center* of P_H . It is the point mapped to $\mathbf{0}$ in Legendre transform coordinates, and coincides with the "analytical center" of Sonnevend [So1, So2]. For each objective function we call the A -trajectory (resp. P -trajectory) through the center the *central A -trajectory* (resp. *central P -trajectory*). We show the central P -trajectory coincides with the central A -trajectory. We give several other characterizations of the central A -trajectory, one of which is that it is the set of solutions to a set of fixed point problems of a parametrized family of logarithmic barrier functions, the *logarithmic barrier function trajectory*.

We describe interpretations due to the second author of A -trajectories as $\dot{\mathbf{q}}$ -trajectories of Lagrangian dynamical systems whose associated Hamiltonian dynamical systems are completely integrable.

We are indebted to Mike Todd for many comments improving the exposition of this paper. The results of §§1–9 were presented at MSRI in January 1986.

2. SUMMARY OF RESULTS

The affine scaling vector fields and projective scaling vector fields are defined for linear programs of special forms which we call standard form and canonical form, respectively.

A *standard form linear program* is one of the following form:

$$(2.1) \quad \begin{cases} \text{minimize } \langle \mathbf{c}, \mathbf{x} \rangle, \\ A\mathbf{x} = \mathbf{b}, \\ \mathbf{x} \geq \mathbf{0}. \end{cases}$$

We call the constraints of such a linear program *standard form constraints*. We say such a linear program is in *strict standard form* if it has a feasible solution $\mathbf{x} = (x_1, \dots, x_n)$ with all $x_i > 0$, and we call the corresponding constraints *strict standard form constraints*. Strict standard form constraints have the nice property that the relative interior $\text{Rel-Int}(P)$ of the polytope P of feasible solutions has

$$\text{Rel-Int}(P) = P \cap \text{Int}(\mathbf{R}_+^n),$$

so consists exactly of vectors \mathbf{x} in P with $\mathbf{x} > \mathbf{0}$.

A *canonical form linear program* is one of the following form:

$$(2.2) \quad \begin{cases} \text{minimize } \langle \mathbf{c}, \mathbf{x} \rangle, \\ A\mathbf{x} = \mathbf{0}, \\ \langle \mathbf{e}, \mathbf{x} \rangle = n, \\ \mathbf{x} \geq \mathbf{0}, \end{cases}$$

such that $\mathbf{e} = (1, 1, \dots, 1)^T$ is a feasible solution, i.e., $A\mathbf{e} = \mathbf{0}$. Constraints of this form are called *canonical form constraints*. A canonical form linear program is a special kind of strict standard form linear program.

The affine and projective scaling vector fields are defined using affine and projective transformations to rescale the linear programs, in a manner described in part I, §4. The *affine scaling vector field* $\mathbf{v}_A(\mathbf{x}; \mathbf{c})$ for the objective function $\langle \mathbf{c}, \mathbf{x} \rangle$ at an interior feasible point \mathbf{x} of the strict standard form linear program (2.1) is given by

$$\mathbf{v}_A(\mathbf{x}; \mathbf{c}) = -X\pi_{(AX)^\perp}(X\mathbf{c})$$

where

$$X = \text{diag}(\mathbf{x}) = \begin{bmatrix} x_1 & & \\ & \ddots & \\ & & x_n \end{bmatrix}$$

is a diagonal matrix and $\pi_{(AX)^\perp}$ denotes orthogonal projection onto the subspace $(AX)^\perp$. In particular

$$\pi_{(AX)^\perp} = I - XA^T(AX^2A^T)^{-1}AX.$$

The *projective scaling vector field* $\mathbf{v}_P(\mathbf{x}; \mathbf{c})$ for the objective function $\langle \mathbf{c}, \mathbf{x} \rangle$ at an interior feasible point \mathbf{x} of a canonical form linear program (2.2) is given by

$$(2.4) \quad \mathbf{v}_P(\mathbf{x}; \mathbf{c}) = -X\pi_{(AX)^\perp}(X\mathbf{c}) + (1/n)\langle X\mathbf{e}, \pi_{(AX)^\perp}(X\mathbf{c}) \rangle X\mathbf{e}.$$

The affine scaling trajectories and projective scaling trajectories are obtained by integrating these vector fields. The *affine scaling trajectory* or *A-trajectory* $T_A(\mathbf{x}; \mathbf{c}, A, \mathbf{b})$ of the standard form linear program (2.1) is the point-set determined by the solution to the *affine scaling differential equation*:

$$\begin{cases} d\mathbf{x}(t)/dt = \mathbf{v}_A(\mathbf{x}(t); \mathbf{c}), \\ \mathbf{x}(0) = \mathbf{x}, \end{cases}$$

extended to the maximal range of the parameter t for which a solution is defined. The *projective scaling trajectory*, or *P-trajectory* $T_P(\mathbf{x}; \mathbf{c}, A)$ of the canonical form problem (2.2) is the point-set determined by the *projective scaling differential equation*:

$$\begin{cases} d\mathbf{x}(t)/dt = \mathbf{v}_P(\mathbf{x}(t); \mathbf{c}), \\ \mathbf{x}(0) = \mathbf{x}, \end{cases}$$

extended to the maximal range of t for which a solution is defined.

§3 introduces a nonlinear change of variables, *Legendre transform coordinates*, to study *A-trajectories* and *P-trajectories*. The Legendre transform coordinate mapping $\phi_H(\mathbf{x})$ is defined for any set H of inequality constraints

$$\langle \mathbf{a}_j, \mathbf{x} \rangle \geq b_j, \quad 1 \leq j \leq m$$

on \mathbf{R}^n , and is defined on the relative interior of the polytope of feasible solutions P_H . It is a projection of the gradient of a logarithmic barrier function associated to the constraints that are not constant on P_H onto the subspace D_H of feasible directions. The feasible direction subspace D_H is determined by the constraints that are constant on P_H . The mapping $\phi_H(\mathbf{x})$ is given by rational functions, and its range is the relative interior of a certain polyhedral cone C_H in D_H . The range is all of D_H in the case that the polytope P_H is bounded. The Legendre transform mapping is one-to-one if the constraints H are *full rank*, that is, if $\mathbf{R}^n = \mathbf{R}[\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_m]$. In that case the inverse mapping ϕ_H^{-1} is well defined and is an algebraic mapping. Legendre transform coordinates transform contravariantly under invertible affine mappings $\mathbf{y} = \mathbf{J}(\mathbf{x}) = \mathbf{L}\mathbf{x} + \mathbf{m}$, so that

$$\phi_H(\mathbf{x}) = L^T \phi_{L(H)}(\mathbf{J}(\mathbf{x})),$$

where L^T denotes the transpose of L . §3 also gives an explicit formula for the Legendre transform coordinates $\phi_H(\mathbf{x})$ for a set of strict standard form constraints:

$$(2.5) \quad \phi_H(\mathbf{x}) = -\pi_{A^\perp}(X^{-1}\mathbf{e}) = -(I - A(AA^T)^{-1}A^T) \begin{bmatrix} 1/x_1 \\ 1/x_2 \\ \vdots \\ 1/x_n \end{bmatrix},$$

For strict standard form problems the mapping $\phi_H(\mathbf{x})$ is one-to-one, and its range is the relative interior of the cone

$$C_H = -\pi_{A^\perp}(\mathbf{R}_+^n),$$

where \mathbf{R}_+^n is the positive orthant in \mathbf{R}^n .

§4 determines the effect on tangent vectors of the Legendre transform mapping for a strict standard form linear program. The tangent spaces in the domain and range can both be identified with A^\perp and we show for $\mathbf{v} \in A^\perp$ that at the point \mathbf{x} one has

$$d\phi_H(\mathbf{v}) = \pi_{A^\perp}(X^{-2}\mathbf{v}),$$

and

$$d\phi_H^{-1}(\mathbf{v}) = X\pi_{(AX)^\perp}(X\mathbf{v}).$$

§4 shows that Legendre transform coordinates linearize the affine scaling trajectories. Theorem 4.1 asserts that

$$d\phi_H(\mathbf{v}_A(\mathbf{x}; \mathbf{c})) = -\pi_{A^\perp}(\mathbf{c}),$$

so that the affine scaling vector field is constant. Hence the images $\phi_H(T_A(\mathbf{x}; \mathbf{c}))$ of A -trajectories in Legendre transform coordinates are (parts of) a family of parallel straight lines in the direction $-\pi_{A^\perp}(\mathbf{c})$.

An immediate consequence of this linearization is that the set of affine scaling trajectories for all objective functions is naturally viewed as the complete set of geodesics for a Riemannian geometry on the relative interior of the polytope of feasible solutions. This geometry is the pullback of Euclidean geometry on the Legendre transform coordinate space, which is \mathbf{R}^n if the polytope of feasible solutions is bounded.

Another immediate consequence of the linearization is that each A -trajectory is part of a real algebraic curve, because the mapping ϕ_H^{-1} is algebraic. In part I we showed that each P -trajectory is algebraically related to another A -trajectory, so it follows that each P -trajectory is part of a real algebraic curve as well.

§5 computes the projective scaling vector field for a canonical form linear program in Legendre transform coordinates. Theorem 5.1 shows that

$$d\phi_H(\mathbf{v}_P(\mathbf{x}; \mathbf{c})) = -\pi_{[\frac{A}{e}]^\perp}(\mathbf{c}) - \frac{1}{n}\langle X\mathbf{e}, \pi_{(AX)^\perp}(X\mathbf{c}) \rangle \phi_H(\mathbf{x}).$$

This vector field has a constant component in the direction

$$\mathbf{c}^* = \pi_{[\frac{A}{e}]^\perp}(\mathbf{c})$$

and a component pointing radially. As a consequence each P -trajectory $\phi(T_P(\mathbf{x}; \mathbf{c}, A))$ in Legendre transform coordinates lies in a plane in the Legendre transform coordinate space.

A set of linear program constraints H having a bounded polytope P_H of feasible solutions has a unique point \mathbf{x}_H whose Legendre transform coordinates are $\phi_H(\mathbf{x}_H) = \mathbf{0}$. We call this point the *center* of P_H . This notion coincides with the “analytical center” introduced by Sonnevend [So1, So2].

The *central A-trajectory* for a given linear program having a bounded polytope of feasible solutions is that A -trajectory passing through the center of

P_H . More generally, for a standard form problem (2.1) we define the *central A-trajectory* $T_A(\mathbf{c}; A, \mathbf{b})$ for the objective function $\langle \mathbf{c}, \mathbf{x} \rangle$ by

$$T_A(\mathbf{c}; A, \mathbf{b}) = \{\mathbf{x}: \phi_H(\mathbf{x}) = t\pi_{A^\perp}(\mathbf{c}) \text{ where } t \in \mathbf{R}\}.$$

This definition makes sense even for standard form problems with unbounded polyhedra, which do not possess a center.

§6 gives several characterizations of central A -trajectories. First, the central P -trajectory of a canonical form linear program is contained in the central A -trajectory, and in the case of a normalized objective function they coincide. Second, the central A -trajectory $T_A(\mathbf{c}; A, \mathbf{b})$ of a standard form problem is exactly the set of centers \mathbf{x}_{H_μ} of the linear programming constraints

$$\begin{cases} A\mathbf{x} = \mathbf{b}, \\ \mathbf{x} \geq \mathbf{0}, \end{cases}$$

together with the extra equality constraint

$$\langle \mathbf{c}, \mathbf{x} \rangle = \mu,$$

for the values $\mu^- < \mu < \mu^+$, where μ^+ and μ^- are the maximum and minimum of the objective function $\langle \mathbf{c}, \mathbf{x} \rangle$ on the polytope of feasible solutions, respectively. Third, if the feasible solution polytope is bounded then that part of the central A -trajectory of a strict standard form linear program on one side on the center is the *logarithmic barrier function trajectory*, described as the set $\mathbf{x}(\mu)$ of solutions to the following parametrized family of nonlinear minimization problems:

$$\begin{cases} \text{minimize } \langle \mathbf{c}, \mathbf{x} \rangle - \mu \left(\sum_{i=1}^n \log x_i \right), \\ A\mathbf{x} = \mathbf{b}, \\ \mathbf{x} \geq \mathbf{0}, \end{cases}$$

for $0 < \mu < \infty$.

§7 derives two power-series expansions for A -trajectories. The first power-series expansion applies to an arbitrary A -trajectory of a standard form problem and takes as power-series parameter a Legendre transform coordinate parameter t , defined by

$$\phi_H(\mathbf{x}(t)) = (t - t_0)\pi_{A^\perp}(\mathbf{c}).$$

The second power-series expansion applies to the central A -trajectory of a canonical form linear program and takes the value z of the objective function as power-series parameter, so that $z = \langle \mathbf{c}, \mathbf{x}(z) \rangle$, and

$$\phi_H(\mathbf{x}(z)) = f(z)\pi_{[\mathcal{F}]^\perp}(\mathbf{c}),$$

for a certain scalar function $f(z)$. These power-series expansions have a simple form, suitable for computation, which yield “higher-order” analogues of Karmarkar’s algorithm, cf. [AKRV, KLSW].

§8 defines A -trajectories and central A -trajectories for a general linear program on \mathbf{R}^n given in the *inequality form*

$$(2.6) \quad \begin{cases} \text{minimize } \langle \mathbf{c}, \mathbf{x} \rangle, \\ \langle \mathbf{a}_j, \mathbf{x} \rangle \geq b_j; \quad 1 \leq j \leq m, \end{cases}$$

which is of full rank. An A -trajectory consists of those points having

$$\phi_H(\mathbf{x}) = - \sum_{j=1}^m \frac{\mathbf{a}_j}{\langle \mathbf{a}_j, \mathbf{x} \rangle - b_j} = \mathbf{v} + \mu \mathbf{c},$$

where \mathbf{v} is a fixed vector and the parameter μ varies. §8 gives three different definitions of A -trajectories and proves they are all equivalent.

§9 studies central trajectories for linear programs having a full-dimensional polytope of feasible solutions. It shows that if H is a set of inequality constraints and H' is another set of constraints obtained by adding extra constraints of the form $\pm \langle \mathbf{c}, \mathbf{x} \rangle \geq b$, for various different values of b , then the central A -trajectory $T_A(\mathbf{c}, H')$ attached to H' is *contained in* the central A -trajectory $T_A(\mathbf{c}, H)$ attached to H . In the case that P_H is a bounded polytope, the center of H' is generally in a different place on the central trajectory than the center of H . Reneger [Re] bases a linear programming algorithm on the idea of approximately following a series of such centers along the central trajectory to an optimal point; the series of centers is obtained by adding extra constraints of the form $-\langle \mathbf{c}, \mathbf{x} \rangle \geq \mu$ where μ is a parameter whose value is changed at each step, and the centers are followed using Newton's method. §9 also shows that the limit point \mathbf{x}_∞ of a central A -trajectory $T_A(\mathbf{c}, H)$ on the boundary ∂P_H of the polytope P_H of feasible solutions is the center of the unique face of ∂P_H in which \mathbf{x}_∞ is a relative interior point.

§10 relates the central trajectories of a pair of dual linear programming problems. Consider the following pair (P) and (D) of dual linear programs:

$$(P) \quad \begin{cases} \text{minimize } \langle \mathbf{c}, \mathbf{x} \rangle, \\ A^T \mathbf{x} \geq \mathbf{b}, \end{cases}$$

and

$$(D) \quad \begin{cases} \text{minimize } -\langle \mathbf{b}, \mathbf{y} \rangle, \\ A\mathbf{y} = \mathbf{c}, \\ \mathbf{y} \geq \mathbf{0}. \end{cases}$$

The combined primal-dual linear program (PD) is:

$$(PD) \quad \begin{cases} \text{minimize } \langle \mathbf{c}, \mathbf{x} \rangle - \langle \mathbf{b}, \mathbf{y} \rangle, \\ A^T \mathbf{x} \geq \mathbf{b}, \\ A\mathbf{y} = \mathbf{c}, \\ \mathbf{y} \geq \mathbf{0}. \end{cases}$$

The *central trajectory* of (PD) orthogonally projects onto the positive half of the central trajectory of (P) in the \mathbf{x} -variables, and onto the positive half of the central trajectory of (D) in the \mathbf{y} -variables. This gives rise to a one-to-one algebraic correspondence between points on the positive half of the central

trajectory of (P) and points on the positive half of the central trajectory of (D) (see Theorem 10.1). This correspondence was previously observed by Osbourne [Os] in the context of logarithmic barrier function trajectories.

§11 gives polynomial ideals of relations which specify algebraic curves that contain A -trajectories. The central A -trajectory of the combined primal-dual linear program (PD) satisfies a particularly simple set of relations. It is the set of polynomial relations in the polynomial ring $C[x, y, u, \mu]$ generated by the relations

$$\begin{aligned} A^T x - u &= b, \\ Ay &= c, \\ y_j u_j &= \mu, \quad 1 \leq j \leq m. \end{aligned}$$

The last set of equations is a parametrically relaxed form of the complementary slackness conditions for (PD) . Megiddo [m2] studies this curve as a logarithmic barrier function trajectory, and a linear programming algorithm based on following it is given in [KMY].

The Legendre transform mapping plays an important role in classical mechanics, where it is used to convert a dynamical system given in Lagrangian form to an equivalent Hamiltonian dynamical system. This leads to the question whether or not there are such dynamical systems of a simple form giving rise to A -trajectories.

§12 describes results due to the second author describing two simple Lagrangian dynamical systems where \dot{q} -trajectories coincide with A -trajectories. The corresponding Hamiltonian dynamical systems are shown to be completely integrable.

3. LEGENDRE TRANSFORM COORDINATES

Legendre transform coordinates are a nonlinear change of variable associated to a set of linear programming constraints. Let H denote a set of linear inequality constraints on \mathbf{R}^n :

$$(3.1) \quad \langle \mathbf{a}_j, \mathbf{x} \rangle \geq b_j, \quad 1 \leq j \leq m.$$

Let P_H denote the polytope of feasible solutions of H , and let D_H denote the vector space of *feasible directions* of H , i.e.,

$$D_H = \{\lambda(\mathbf{x}_1 - \mathbf{x}_2) : \mathbf{x}_1, \mathbf{x}_2 \in P_H \text{ and } \lambda \in \mathbf{R}\}.$$

Let M_H denote the *affine hull* of P_H , i.e.,

$$M_H = \{\mathbf{x}_0 + \mathbf{x}_1 : \mathbf{x}_0 \in P_H \text{ and } \mathbf{x}_1 \in D_H\}.$$

The *dimension* $d = d(H)$ is defined by

$$d(H) = \dim(P_H) = \dim(D_H) = \dim(M_H)$$

and the *rank* $r = r(H)$ is the dimension of the vector space $[\mathbf{a}_1, \dots, \mathbf{a}_m]$ spanned by the vectors \mathbf{a}_i that are normal to the inequality constraint boundaries. A set of constraints is *full dimensional* if $d(H) = n$ and is of *full rank* if

$r(H) = n$. Any set of constraints H for which P_H is bounded is necessarily of full rank.

We first define the Legendre transform coordinate mapping $\phi_H(\mathbf{x})$ in the full-dimensional case when $\text{Int}(P_H)$ is nonempty. Associate to H the *logarithmic barrier function*

$$(3.2) \quad f_H(\mathbf{x}) = - \sum_{j=1}^m \log(\langle \mathbf{a}_j, \mathbf{x} \rangle - b_j),$$

which is defined for $\mathbf{x} \in \text{Int}(P_H)$. The *Legendre transform mapping* $\phi_H(\mathbf{x})$ is the gradient of $f_H(\mathbf{x})$, which is

$$(3.3) \quad \phi_H(\mathbf{x}) = \nabla f_H(\mathbf{x}) = - \sum_{j=1}^m \frac{\mathbf{a}_j}{\langle \mathbf{a}_j, \mathbf{x} \rangle - b_j}.$$

To define the Legendre transform coordinate mapping in the case that P_H is nonempty and of dimension less than n , we need some more definitions. A constraint $\langle \mathbf{a}_j, \mathbf{x} \rangle \geq b_j$ in H is *nonsingular* if there exists a feasible point \mathbf{x}_0 with $\langle \mathbf{a}, \mathbf{x}_0 \rangle > b_j$; otherwise it is *singular*. (That is, the singular constraints function as equality constraints.) Let H_n denote the set of nonsingular constraints in H ; by renumbering the constraints, we may suppose they are

$$(3.4) \quad \langle \mathbf{a}_j, \mathbf{x} \rangle \geq b_j, \quad 1 \leq j \leq m^*,$$

with $m^* \leq m$. The nonsingular constraint polytope P_{H_n} is always full dimensional. The *Legendre transform coordinate mapping* $\phi_H(\mathbf{x})$ is defined on the relative interior $\text{Rel-Int}(P_H)$ of P_H and is defined as the projection onto the feasible direction space D_H of the Legendre transform mapping $\phi_{H_n}(\mathbf{x})$ of the nonsingular constraint set H_n , which is:

$$(3.5) \quad \begin{aligned} \phi_H(\mathbf{x}) &= \pi_{D_H}(\phi_{H_n}(\mathbf{x})) = \pi_{D_H}(\nabla f_{H_n}(\mathbf{x})) \\ &= \pi_{D_H} \left(- \sum_{j=1}^{m^*} \frac{\mathbf{a}_j}{\langle \mathbf{a}_j, \mathbf{x} \rangle - b_j} \right). \end{aligned}$$

Here π_{D_H} denotes orthogonal projection onto the feasible direction subspace D_H .

The Legendre transform coordinate mapping is most naturally viewed as a mapping onto a suitable quotient space of the dual space $(\mathbf{R}^n)^*$. This “coordinate-free” definition is given in Appendix A. In it the gradient $\nabla f_H(\mathbf{x})$ is replaced by the differential $df_H(\mathbf{x})$, which is an element of $(\mathbf{R}^n)^*$, the set of linear functionals $l: \mathbf{R}^n \rightarrow \mathbf{R}$. The gradient ∇f_H is obtained from the differential df_H using the (not natural) isomorphism $(\mathbf{R}^n)^* \rightarrow \mathbf{R}^n$ given by the transpose map: $l_y \in (\mathbf{R}^n)^*$ corresponds to $\mathbf{y} \in \mathbf{R}^n$ where $l_y(\mathbf{x}) = \langle \mathbf{y}, \mathbf{x} \rangle$. This mapping also identifies quotient spaces of $(\mathbf{R}^n)^*$ with subspaces of \mathbf{R}^n , and quotient mappings on $(\mathbf{R}^n)^*$ with projection operators on \mathbf{R}^n . We use the coordinatized definition for $\phi_H(\mathbf{x})$ because it is convenient for explicit computations.

The basic properties of the Legendre transform mapping in the full-dimensional case are as follows.

Theorem 3.1. *Let H be a set of constraints*

$$\langle \mathbf{a}_j, \mathbf{x} \rangle \geq b_j, \quad 1 \leq j \leq m,$$

in \mathbf{R}^n such that P_H has nonempty interior. The Legendre transform coordinate map

$$(3.6) \quad \phi_H(\mathbf{x}) = - \sum_{j=1}^m \frac{\mathbf{a}_j}{\langle \mathbf{a}_j, \mathbf{x} \rangle - b_j}$$

has domain $\text{Int}(P_H)$ and range $\text{Rel-Int}(C_H)$, where C_H is the cone

$$(3.7) \quad C_H = \mathbf{R}^+[-\mathbf{a}_1, -\mathbf{a}_2, \dots, -\mathbf{a}_m].$$

If H is of full rank then C_H is full dimensional and ϕ_H is a one-to-one mapping.

Proof. One has

$$\text{Int}(P_H) = \{\mathbf{x}: \langle \mathbf{a}_j, \mathbf{x} \rangle > b_j \text{ for } 1 \leq j \leq m\}$$

and

$$(3.8) \quad \text{Rel-Int}(C_H) = \left\{ \mathbf{x}: \mathbf{x} = - \sum_{j=1}^m \mu_j \mathbf{a}_j \text{ with all } \mu_j > 0 \right\}.$$

Formula (3.6) shows that $\mathbf{x} \in \text{Int}(P_H)$ implies that $\phi_H(\mathbf{x}) \in \text{Rel-Int}(C_H)$.

To show that the mapping is onto, let $\mathbf{c} \in \text{Rel-Int}(C_H)$ and write $\mathbf{c} = - \sum_{j=1}^m \mu_j \mathbf{a}_j$, with all $\mu_j > 0$. Consider the function

$$(3.9) \quad g(\mathbf{x}) = -\langle \mathbf{c}, \mathbf{x} \rangle - \sum_{j=1}^m \log(\langle \mathbf{a}_j, \mathbf{x} \rangle - b_j).$$

We claim that $g(\mathbf{x})$ attains a global minimum \mathbf{x}_0 on $\text{Int}(P_H)$. If so, then necessarily $\nabla g(\mathbf{x}_0) = \mathbf{0}$. Since

$$\nabla g(\mathbf{x}) = -\mathbf{c} + \phi_H(\mathbf{x}),$$

it follows that $\phi_H(\mathbf{x}_0) = \mathbf{c}$ as required.

We prove the claim. It is clear that $g(\mathbf{x})$ attains a global minimum if P_H is bounded, for in that case every term on the right side of (3.9) is bounded below on $\text{Int}(P_H)$, and $g(\mathbf{x}) \rightarrow \infty$ as \mathbf{x} approaches the boundary of P_H . The case that P_H is unbounded requires a more involved argument. We rewrite the formula for $g(\mathbf{x})$ as

$$(3.10) \quad g(\mathbf{x}) = \sum_{j=1}^m \left(-\frac{1}{m} \langle \mathbf{c}, \mathbf{x} \rangle - \log(\langle \mathbf{a}_j, \mathbf{x} \rangle - b_j) \right)$$

and prove that each term in the sum is bounded below on $\text{Int}(P_H)$. Since

$$-\langle \mathbf{c}, \mathbf{x} \rangle = \sum_{j=1}^m \mu_j (\langle \mathbf{a}_j, \mathbf{x} \rangle - b_j) + \sum_{j=1}^m \mu_j b_j$$

with all $\mu_j > 0$, on letting $B = \sum \mu_j b_j$ we have

$$-\frac{1}{m}\langle \mathbf{c}, \mathbf{x} \rangle \geq \frac{1}{m} \sum \mu_j b_j = \frac{1}{m} B,$$

so that $-\langle \mathbf{c}, \mathbf{x} \rangle$ is bounded below. Then since (3.10) gives

$$(3.11) \quad -\langle \mathbf{c}, \mathbf{x} \rangle \geq \mu_j (\langle \mathbf{a}_j, \mathbf{x} \rangle - b_j) + B,$$

we have

$$-\frac{1}{m}\langle \mathbf{c}, \mathbf{x} \rangle - \log(\langle \mathbf{a}_j, \mathbf{x} \rangle - b_j) \geq -\frac{1}{m}\langle \mathbf{c}, \mathbf{x} \rangle - \log\left(-\frac{1}{\mu_j}\langle \mathbf{c}, \mathbf{x} \rangle - \frac{1}{\mu_j}B\right)$$

which is bounded below. If

$$W = \{\mathbf{x}: \langle \mathbf{c}, \mathbf{x} \rangle = 0 \text{ and } \langle \mathbf{a}_j, \mathbf{x} \rangle = 0 \text{ for } 1 \leq j \leq m^*\},$$

then $\mathbf{x}_0 \in P_H$ implies that the flat $\mathbf{x}_0 + W \subseteq P_H$, and $g(\mathbf{x})$ is constant on this flat. Let W^\perp denote the orthogonal complement of W , and set $\bar{P}_H = P_H \cap W^\perp$. Then $P_H = \bar{P}_H \oplus W$ and $g(\mathbf{x})$ attains all values in its range on $\text{Rel-Int}(\bar{P}_H)$. Finally one checks that $g(\mathbf{x}) \rightarrow \infty$ in \bar{P}_H as \mathbf{x} approaches $\partial\bar{P}_H$ because some $-\log(\langle \mathbf{a}_j, \mathbf{x} \rangle - b_j) \rightarrow \infty$, and $g(\mathbf{x}) \rightarrow \infty$ as $\|\mathbf{x}\| \rightarrow \infty$ in \bar{P}_H because $-\langle \mathbf{c}, \mathbf{x} \rangle \rightarrow \infty$. (The last assertion holds because in P_H as $\|\mathbf{x}\| \rightarrow \infty$ either $-\langle \mathbf{c}, \mathbf{x} \rangle \rightarrow \infty$ or else some $\langle \mathbf{a}_j, \mathbf{x} \rangle - b_j \rightarrow \infty$, in which case $-\langle \mathbf{c}, \mathbf{x} \rangle \rightarrow \infty$ by (3.11).) This implies that $g(\mathbf{x})$ attains a global minimum in $\text{Rel-Int}(\bar{P}_H)$, which is also its global minimum in $\text{Int}(P_H)$. The claim is proved.

If H is full rank, then C_H is full dimensional. Also the Jacobian $\nabla\phi_H(\mathbf{x})$ is the Hessian $\nabla^2 f_H(\mathbf{x})$, which is

$$(3.12) \quad \nabla\phi_H(\mathbf{x}) = \sum_{j=1}^m \frac{1}{(\langle \mathbf{a}_j, \mathbf{x} \rangle - b_j)^2} \mathbf{a}_j \mathbf{a}_j^T.$$

Each term $\mathbf{a}_j \mathbf{a}_j^T$ is a positive semidefinite symmetric matrix. The full-rank hypothesis guarantees that $\nabla\phi_H(\mathbf{x})$ is positive definite for each $\mathbf{x} \in \text{Int}(P_H)$. Indeed if $\mathbf{y} \in \mathbf{R}^n$ then

$$\mathbf{y}^T \nabla\phi_H(\mathbf{x}) \mathbf{y} = \sum_{j=1}^m \frac{1}{(\langle \mathbf{a}_j, \mathbf{x} \rangle - b_j)^2} \langle \mathbf{a}_j, \mathbf{y} \rangle^2.$$

Since $\mathbf{R}^n = [\mathbf{a}_1, \dots, \mathbf{a}_m]$ there exists some $\langle \mathbf{a}_j, \mathbf{y} \rangle > 0$ so this gives $\mathbf{y}^T \nabla\phi_H(\mathbf{x}) \mathbf{y} > 0$ as required. It is well known [BT] that this positive-definiteness condition implies that the logarithmic barrier function $f_H(\mathbf{x})$ is strictly convex on $\text{Int}(P_H)$ and that $\phi_H(\mathbf{x}) = \nabla f_H(\mathbf{x})$ is one-to-one on $\text{Int}(P_H)$. \square

The Legendre transform mapping $\phi_H = \nabla f_H$ is a mapping given by rational functions, so is an algebraic mapping, and therefore its inverse mapping $\phi_H^{-1}(\mathbf{y})$ is given by algebraic functions. The rational map $\phi_H(\mathbf{x})$ defined by (3.3) makes sense outside the domain $\text{Int}(P_H)$; in fact it is defined on all of \mathbf{C}^n outside of the

hyperplanes $\{\mathbf{x}: \langle \mathbf{a}_j, \mathbf{x} \rangle = b_j\}$. In this paper we are concerned with the mapping ϕ_H restricted to the real domain $\text{Int}(P_H)$. However some of its properties may best be understood in terms of its behavior on \mathbb{C}^n .

To illustrate the Legendre transform coordinate mapping ϕ_H and its inverse mapping ϕ_H^{-1} , consider the following one-dimensional example. The constraint set H is

$$\begin{cases} x \geq 0, \\ -x \geq -1, \end{cases}$$

which has $\text{Int}(P_H) = \{x: 0 < x < 1\}$. The logarithmic barrier function $f_H(\mathbf{x}) = -\log x - \log(1-x)$ yields

$$\phi_H(x) = \nabla f_H(\mathbf{x}) = -\frac{1}{x} + \frac{1}{1-x}.$$

Theorem 3.1 asserts that $\phi_H: (0, 1) \rightarrow (-\infty, \infty)$ is one-to-one and onto. A simple computation shows that if $y = \phi_H(x)$ then

$$x = \phi_H^{-1}(y) = \frac{y + 2 - \sqrt{y^2 + 4}}{2y},$$

where the minus sign is taken in the square root to give the branch of ϕ_H^{-1} that lies in $\text{Int}(P_H)$. This example shows that ϕ_H^{-1} is not a rational function in general.

Legendre transform coordinates transform contravariantly under invertible affine transformations.

Theorem 3.2. *Let H be a set of constraints*

$$\langle \mathbf{a}_j, \mathbf{x} \rangle \geq b_j, \quad 1 \leq j \leq m,$$

in \mathbb{R}^n such that P_H has nonempty interior. Let $\mathbf{y} = \mathbf{J}(\mathbf{x}) = L\mathbf{x} + \mathbf{m}$ be an invertible affine transformation with inverse $\mathbf{x} = \mathbf{J}^{-1}(\mathbf{y}) = L^{-1}\mathbf{y} - L^{-1}\mathbf{m}$, and let $\mathbf{J}(H)$ denote the transformed set of constraints

$$(3.13) \quad \langle \mathbf{a}_j, \mathbf{J}^{-1}(\mathbf{y}) \rangle \geq b_j, \quad 1 \leq j \leq m.$$

Then $P_{\mathbf{J}(H)} = \mathbf{J}(P_H)$ and the following diagram commutes:

$$(3.14) \quad \begin{array}{ccc} \text{Int}(P_H) & \xrightarrow{\mathbf{J}} & \text{Int}(P_{\mathbf{J}(H)}) \\ \phi_H \downarrow & & \downarrow \phi_{\mathbf{J}(H)} \\ \text{Rel-Int}(C_H) & \xleftarrow{L^T} & \text{Rel-Int}(C_{\mathbf{J}(H)}) \end{array}$$

Proof. The relation $P_{\mathbf{J}(H)} = \mathbf{J}(P_H)$ is immediate. Next we claim that

$$f_{\mathbf{J}(H)}(\mathbf{J}(\mathbf{x})) = f_H(\mathbf{x}),$$

for all $x \in \text{Int}(P_H)$. Indeed if $\mathbf{y} = \mathbf{J}(\mathbf{x})$ then

$$f_{\mathbf{J}(H)}(\mathbf{J}(\mathbf{x})) = -\sum_{j=1}^m \log(\langle \mathbf{a}_j, \mathbf{J}^{-1}(\mathbf{J}(\mathbf{x})) \rangle - b_j) = f_H(\mathbf{x}).$$

Now differentials transform contravariantly under invertible affine transformations, i.e., for any function $f(\mathbf{x})$, if $\bar{f}(\mathbf{y}) = f(\mathbf{J}^{-1}(\mathbf{y}))$ then

$$(3.15) \quad (df)_{\mathbf{x}}(\mathbf{v}) = \mathbf{J}^*((d\bar{f})_{\mathbf{J}(\mathbf{x})})(\mathbf{J}(\mathbf{v})),$$

for \mathbf{v} a tangent vector at \mathbf{x} , where $\mathbf{J}^*: (\mathbf{R}^n)^* \rightarrow (\mathbf{R}^n)^*$ is the linear transformation *adjoint* to \mathbf{J} , defined by

$$\mathbf{J}^*(l)(\mathbf{x}) = l(\mathbf{J}(\mathbf{x}) - \mathbf{J}(\mathbf{0})) = l(L\mathbf{x}),$$

where $l \in (\mathbf{R}^n)^*$ is a linear functional $l: \mathbf{R}^n \rightarrow \mathbf{R}$. Under the identification of $(\mathbf{R}^n)^*$ with \mathbf{R}^n by $l_{\mathbf{y}}(\mathbf{x}) = \langle \mathbf{y}, \mathbf{x} \rangle$ one has $\mathbf{J}^*(\mathbf{y}) = L^T \mathbf{y}$ and $(df)_{\mathbf{x}} = (\nabla f)_{\mathbf{x}}$. Thus (3.14) follows as a special case of (3.15). \square

We call the mapping $\phi_{\mathbf{H}}(\mathbf{x})$ the Legendre transform coordinate mapping because it is related to the Legendre transformation used in the theory of convexity (see [F, R1, R2]), which is an extension of the Legendre transformation used in differential equations (see [Ar, CH, Ln]). To explain this relation, let the pair (f, D) consist of a real-valued function f defined on a domain D in \mathbf{R}^n such that:

- (i) D is open and convex and f is strictly convex and continuously differentiable on D .
- (ii) For any sequence $\{\mathbf{x}_i\}$ in D with $\lim_{i \rightarrow \infty} \mathbf{x}_i = \bar{\mathbf{x}}$ a point on ∂D one has $\|df(\mathbf{x}_i)\| \rightarrow \infty$.

To any pair (f, D) with these properties one associates a pair (f^c, D^c) in which f^c is the (Fenchel) conjugate function $f^c: (\mathbf{R}^n)^* \rightarrow \mathbf{R} \cup \{+\infty\}$ defined by

$$f^c(l) = \sup\{l(\mathbf{x}) - f(\mathbf{x}) : \mathbf{x} \in D\},$$

and $D^c = \{l: f^c(l) < \infty\}$. Rockafellar calls (f^c, D^c) the Legendre transform of (f, D) and proves [R2, Theorem 26.5] that it has the following properties:

- (1) The map $\mathbf{x} \rightarrow (df)_{\mathbf{x}}$ is a one-to-one map of D onto D^c .
- (2) The pair (f^c, D^c) has properties (i), (ii), so that (f^{cc}, D^{cc}) is defined, and $f^{cc} = f$, $D^{cc} = D$.
- (3) $(df^c)_{\mathbf{y}}$ is the inverse map to $(df)_{\mathbf{x}}$.

The proof of Theorem 3.1 shows that if $P_{\mathbf{H}}$ is full dimensional and \mathbf{H} is of full rank then $(f_{\mathbf{H}}, \text{Int}(P_{\mathbf{H}}))$ satisfies (i), (ii), the Legendre transform coordinate map $\phi_{\mathbf{H}}$ is the differential $(df_{\mathbf{H}})_{\mathbf{x}}$, and $D^c = \text{Int}(C_{\mathbf{H}})$. By Rockafellar's theorem the pair $(f_{\mathbf{H}}^c, \text{Int}(C_{\mathbf{H}}))$ exists and $(df_{\mathbf{H}}^c)_{\mathbf{y}}$ inverts the Legendre transform coordinate mapping. It seems a difficult problem to find $f_{\mathbf{H}}^c$ explicitly, however.

Now we treat the Legendre transform coordinate mapping $\phi_{\mathbf{H}}(\mathbf{x})$ in the case that $P_{\mathbf{H}}$ is a lower-dimensional polytope.

Theorem 3.3. *Let \mathbf{H} be a set of constraints in \mathbf{R}^n whose nonsingular constraints H_n are*

$$(3.16) \quad \langle \mathbf{a}_j, \mathbf{x} \rangle \geq b_j, \quad 1 \leq j \leq m^*.$$

The Legendre transform coordinate mapping $\phi_H(\mathbf{x})$ has domain $\text{Rel-Int}(P_H)$ and range $\text{Rel-Int}(C_H)$ where

$$C_H = \pi_{D_H}(C_{H_n}) \quad \text{and} \quad C_{H_n} = \mathbf{R}^+[-\mathbf{a}_1, -\mathbf{a}_2, \dots, -\mathbf{a}_{m^*}].$$

If H is of full rank then $\dim(C_H) = \dim(P_H)$ and ϕ_H is a one-to-one mapping.

Proof. This is proved in the full-dimensional case in Theorem 3.1, so we may suppose that $d = \dim(P_H) < n$. We reduce to the full-dimensional case by an affine change of variables. Take a one-to-one affine transformation $\mathbf{J}: \mathbf{R}^d \rightarrow \mathbf{R}^n$ such that $\mathbf{J}(\mathbf{R}^d) = M_H$ the affine hull of P_H . Let $\mathbf{y} = \mathbf{J}(\mathbf{x}) = \mathbf{L}\mathbf{x} + \mathbf{m}$ and note that the associated linear mapping $\mathbf{y} = \mathbf{L}\mathbf{x}$ maps \mathbf{R}^d to the feasible direction subspace D_H . Now define a set of constraints $\bar{H}_n = \mathbf{J}^{-1}(H_n)$ in \mathbf{R}^d by

$$\langle \mathbf{a}_j, \mathbf{J}(\mathbf{x}) \rangle \geq b_j, \quad 1 \leq j \leq m^*.$$

The polytope $P_{\bar{H}_n}$ is easily checked to be full-dimensional in \mathbf{R}^d , and

$$(3.17) \quad \mathbf{J}(P_{\bar{H}_n}) = P_{H_n} \cap M_H = P_H,$$

and one also has

$$(3.18) \quad f_{H_n}(\mathbf{J}(\mathbf{x})) = f_{\bar{H}_n}(\mathbf{x}), \quad \mathbf{x} \in \mathbf{R}^d.$$

We claim that the following diagram commutes:

$$(3.19) \quad \begin{array}{ccc} \text{Int}(P_{\bar{H}_n}) & \xrightarrow{\mathbf{J}} & \text{Rel-Int}(P_H) \\ \phi_{\bar{H}_n} \downarrow & & \downarrow \phi_{H_n} \\ \text{Rel-Int}(C_{\bar{H}_n}) & \xleftarrow{\mathbf{L}^T} & \text{Rel-Int}(C_{H_n}) \\ & \nwarrow \mathbf{L}^T & \downarrow \pi_{D_H} \\ & & \text{Rel-Int}(C_H) \end{array}$$

The commutativity of the top square in (3.19) follows from the general transformation property of gradients under a one-to-one affine change of variable (which generalizes (3.15)). The bottom triangle in (3.19) commutes because

$$(3.20) \quad \mathbf{L}^T \mathbf{y} = \mathbf{0} \quad \text{for all } \mathbf{y} \in D_H^\perp.$$

This holds since for all $\mathbf{x} \in \mathbf{R}^d$ one has $\mathbf{L}\mathbf{x} \in D_H$ so that $\langle \mathbf{L}^T \mathbf{y}, \mathbf{x} \rangle = \langle \mathbf{y}, \mathbf{L}\mathbf{x} \rangle = 0$.

Now by Theorem 3.1 the map $\phi_{\bar{H}_n}$ is one-to-one and onto its range $\text{Rel-Int}(C_{\bar{H}_n})$, hence the commutativity of (3.19) guarantees that the mapping $\phi_H = \pi_{D_H} \circ \phi_{H_n}$ maps onto $\text{Rel-Int}(C_H)$.

Finally observe that if H is full rank in \mathbf{R}^n then the constraints \bar{H}_n are full rank in \mathbf{R}^d . By Theorem 3.1 the mapping $\phi_{\bar{H}_n}$ is one-to-one, and then (3.19) implies that ϕ_H is also one-to-one, since \mathbf{J} is onto. \square

Legendre transform coordinates transform in a simple way under affine transformations in the general case.

Theorem 3.4. *Let H be a set of constraints in \mathbf{R}^n*

$$\langle \mathbf{a}_j, \mathbf{x} \rangle \geq b_j, \quad 1 \leq j \leq m,$$

having D_H as its subspace of feasible directions, and let M_H be the affine hull of P_H . Let $\mathbf{y} = \mathbf{J}(\mathbf{x}) = \mathbf{L}\mathbf{x} + \mathbf{m}$ be an affine mapping from \mathbf{R}^n to \mathbf{R}^k which is one-to-one on the domain M_H . Let \mathbf{J}^{-1} denote any affine mapping from \mathbf{R}^k to \mathbf{R}^n such that $\mathbf{J}^{-1}(\mathbf{J}(\mathbf{x})) = \mathbf{x}$ for all $\mathbf{x} \in M_H$ and let $\mathbf{J}(H)$ denote the set of constraints

$$\langle \mathbf{a}_j, \mathbf{J}^{-1}(\mathbf{y}) \rangle \geq b_j, \quad 1 \leq j \leq m,$$

on \mathbf{R}^k . Then $D_{\mathbf{J}(H)} = \mathbf{J}(D_H)$, $P_{\mathbf{J}(H)} = \mathbf{J}(P_H)$, and the following diagram commutes:

$$(3.21) \quad \begin{array}{ccc} \text{Rel-Int}(P_H) & \xrightarrow{\mathbf{J}} & \text{Rel-Int}(P_{\mathbf{J}(H)}) \\ \downarrow \phi_H & & \downarrow \phi_{\mathbf{J}(H)} \\ \text{Rel-Int}(C_H) & \xleftarrow{L^T} & \text{Rel-Int}(C_{\mathbf{J}(H)}) \end{array}$$

We remark that although the choice of \mathbf{J}^{-1} is not necessarily unique, the Legendre transform mapping $\phi_{\mathbf{J}(H)}$ is well defined independent of this choice.

Proof. We can find a full-dimensional linear program \bar{H} on \mathbf{R}^d , where $d = d(H)$, and two injective affine mappings $\mathbf{J}_1: \mathbf{R}^d \rightarrow \mathbf{R}^n$ and $\mathbf{J}_2: \mathbf{R}^d \rightarrow \mathbf{R}^k$ such that $\mathbf{J}_1(\bar{H}) = H$, $\mathbf{J}_2(\bar{H}) = \mathbf{J}(\bar{H})$, and such that the top triangle in the following diagram commutes:

$$\begin{array}{ccccc} & & \text{Int}(P_{\bar{H}}) & & \\ & \swarrow \mathbf{J}_1 & & \searrow \mathbf{J}_2 & \\ \text{Rel-Int}(P_{\bar{H}}) & \xrightarrow{\quad \mathbf{J} \quad} & \text{Rel-Int}(P_{\mathbf{J}(\bar{H})}) & & \\ & \searrow \phi_{\bar{H}} & & \swarrow \phi_{\mathbf{J}(\bar{H})} & \\ & & \text{Rel-Int}(C_{\bar{H}}) & & \\ & \swarrow L_1^T & & \searrow L_2^T & \\ \text{Rel-Int}(C_H) & \xleftarrow{L^T} & \text{Rel-Int}(C_{\mathbf{J}(H)}) & & \end{array}$$

Here $\mathbf{J}_i(\mathbf{x}) = L_i \mathbf{x} + \mathbf{m}_i$. Then the whole diagram commutes, for the two sides commute by (3.19), and the relation $L^T L_2^T = L_1^T$ follows from $\mathbf{J} \circ \mathbf{J}_1 = \mathbf{J}_2$. This proves (3.21). Finally since \mathbf{J} and ϕ_H are onto maps, (3.21) determines $\phi_{\mathbf{J}(H)}$ uniquely, so it must be well defined independent of the choice of \mathbf{J}^{-1} . \square

Given a set of inequality constraints H of full rank the map ϕ_H is one-to-one so there is at most one point \mathbf{x} such that $\phi_H(\mathbf{x}) = \mathbf{0}$. If it exists we call it the *center of H* and denote it by \mathbf{x}_H . Theorem 3.4 implies that the center is an affine invariant of H in the sense that

$$(3.22) \quad \mathbf{J}(\mathbf{x}_H) = \mathbf{x}_{\mathbf{J}(H)}$$

holds for any affine transformation that is one-to-one on the domain P_H . In the full-dimensional case the center is defined by $\nabla f_H(\mathbf{x}_H) = \mathbf{0}$, which is equivalent to the condition that \mathbf{x}_H maximizes the function

$$\prod_{j=1}^m (\langle \mathbf{a}_j, \mathbf{x} \rangle - b_j)$$

on $\text{Int}(P_H)$, so \mathbf{x}_H coincides with the “analytical center” of Sonnevend [So1, So2].

Now we find Legendre transform coordinates for the constraints of a (strict) standard form linear program. Recall that *standard form constraints* are given by

$$(3.23) \quad \begin{cases} A\mathbf{x} = \mathbf{b}, \\ \mathbf{x} \geq \mathbf{0}, \end{cases}$$

and that such a set of constraints is in *strict standard form* if it has a feasible solution $\mathbf{x} = (x_1, \dots, x_n)$ with all $x_i > 0$. For a linear program in strict standard form the relative interior $\text{Rel-Int}(P_H)$ of the polytope P_H of feasible solutions is nonempty and is $P_H \cap \text{Int}(\mathbf{R}_+^n)$, which is

$$\begin{cases} A\mathbf{x} = \mathbf{b}, \\ \mathbf{x} > \mathbf{0}. \end{cases}$$

A set of standard form constraints H is always of full rank. For a set of strict standard form constraints H the nonsingular constraints H_n are

$$x_i \geq 0, \quad 1 \leq i \leq n,$$

and the associated logarithmic barrier function is

$$f_{H_n}(\mathbf{x}) = - \sum_{i=1}^n \log x_i.$$

Hence

$$\phi_{H_n}(\mathbf{x}) = -X^{-1}\mathbf{e} = - \begin{bmatrix} 1/x_1 \\ \vdots \\ 1/x_n \end{bmatrix}$$

where $X = \text{diag}(x_1, \dots, x_n)$ is a diagonal matrix.

Theorem 3.3 yields the following result, stated as a theorem for ease of reference.

Theorem 3.5. *Let H denote a set of strict standard form linear programming constraints,*

$$\begin{cases} Ax = \mathbf{b}, \\ \mathbf{x} \geq \mathbf{0}. \end{cases}$$

The Legendre transform mapping ϕ_H having domain $\text{Rel-Int}(P_H)$ is defined by

$$(3.24) \quad \phi_H(\mathbf{x}) = -\pi_{A^\perp}(X^{-1}\mathbf{e}).$$

It is a rational mapping which is one-to-one and onto its range $\text{Rel-Int}(C_H)$. If the feasible solution set P_H is bounded then its range is all of A^\perp . If \mathbf{e} is in P_H , and \mathbf{e}^T is in the row space of A then $\phi_H(\mathbf{e}) = \mathbf{0}$, so that \mathbf{e} is the center of P_H . \square

The mapping $\phi_H(\mathbf{x})$ is given explicitly in matrix form as

$$(3.25) \quad \phi_H(\mathbf{x}) = -(I - A^T(AA^T)^{-1}A)X^{-1}\mathbf{e}$$

provided that the matrix A has full row rank so that $(AA^T)^{-1}$ exists. This formula shows that the entries of $\phi_H(\mathbf{x})$ are homogeneous rational functions of \mathbf{x} of degree -1 . Theorem 3.5 implies that for a strict standard form problem the mapping ϕ_H is real-analytic, and that the inverse mapping

$$\phi_H^{-1}: \text{Rel-Int}(C_H) \rightarrow \text{Rel-Int}(P_H)$$

is algebraic, and that ϕ_H^{-1} is a real-analytic diffeomorphism of $\text{Rel-Int}(C_H)$ onto $\text{Rel-Int}(P_H)$.

We compute the effect of the Legendre transform mapping ϕ_H on tangent vectors. The tangent space $T(\text{Rel-Int}(P_H))_{\mathbf{x}}$ at a point \mathbf{x} of $\text{Rel-Int}(P_H)$ is identified with A^\perp by viewing P_H as embedded in \mathbf{R}^n , and the tangent space $T(A^\perp)_{\mathbf{x}}$ to A^\perp has a similar identification with A^\perp . Let

$$(d\phi_H)_{\mathbf{x}}: T(\text{Rel-Int}(P_H))_{\mathbf{x}} \rightarrow T(A^\perp)_{\mathbf{x}}$$

denote the differential of ϕ_H .

Theorem 3.6. *For any set H of strict standard form linear programming constraints and any $\mathbf{x} \in \text{Rel-Int}(P_H)$ the mapping*

$$(3.26) \quad (d\phi_H)_{\mathbf{x}}: T(\text{Rel-Int}(P_H))_{\mathbf{x}} \rightarrow T(A^\perp)_{\phi(\mathbf{x})}$$

is a linear isomorphism. In particular, on identifying both of these tangent spaces with A^\perp one has for $\mathbf{v} \in A^\perp$ that

$$(3.27) \quad (d\phi_H)_{\mathbf{x}}(\mathbf{v}) = \pi_{A^\perp}(X^{-2}\mathbf{v}),$$

and if $\mathbf{y} = \phi_H(\mathbf{x})$ then

$$(3.28) \quad (d\phi_H^{-1})_{\mathbf{y}}(\mathbf{v}) = X\pi_{(AX)^\perp}(X\mathbf{v}).$$

Proof. For \mathbf{v} in A^\perp and \mathbf{x} in $\text{Rel-Int}(P_H)$ the curve $\mathbf{c}(t) = \mathbf{x} + t\mathbf{v}$ lies in $\text{Rel-Int}(P_H)$ for $-\varepsilon < t < \varepsilon$ with ε small enough, and has \mathbf{v} as tangent vector at \mathbf{x} . Then if $X = \text{diag}(\mathbf{x})$ and $V = \text{diag}(\mathbf{v})$ then for $|t|$ small enough we have

$$\begin{aligned}\phi_H(\mathbf{x} + t\mathbf{v}) &= -\pi_{A^\perp}((X + tV)^{-1}\mathbf{e}) \\ &= -\pi_{A^\perp}((I + tX^{-1}V)^{-1}X^{-1}\mathbf{e}) \\ &= -\pi_{A^\perp}(X^{-1}\mathbf{e} - tX^{-2}V\mathbf{e} + O(t^2)) \\ &= \phi_H(\mathbf{x}) + t\pi_{A^\perp}(X^{-2}\mathbf{v}) + O(t^2),\end{aligned}$$

from which (3.27) follows. Here we use the identity

$$(X + tV)^{-1} = (I + tX^{-1}V^{-1})^{-1}X^{-1}$$

which is valid for small enough $|t|$ since the diagonal matrices X and V commute and X is invertible.

We define the linear operators $L_i: \mathbf{R}^n \rightarrow \mathbf{R}^n$ by $L_1(\mathbf{w}) = \pi_{A^\perp}(X^{-2}\mathbf{w})$ and $L_2(\mathbf{w}) = X\pi_{AX^\perp}(X\mathbf{w})$. It is easy to check that both these operators have range spaces contained in A^\perp . The formula (3.28) for $(d\phi_H^{-1})_y$ is equivalent to the algebraic identity

$$(3.29) \quad L_2(L_1(\mathbf{v})) = \mathbf{v}, \quad \text{all } \mathbf{v} \in A^\perp.$$

That is, the (generally noninvertible) linear operators L_1 and L_2 when restricted to the domain A^\perp are invertible and are mutual inverses of each other. To verify (3.29) suppose that $\mathbf{v} \in A^\perp$ and set

$$\tilde{\mathbf{v}} = L_1(\mathbf{v}) = \pi_{A^\perp}(X^{-2}\mathbf{v}).$$

Using $\pi_{A^\perp} = I - A(AA^T)^{-1}A$ we have

$$\begin{aligned}(3.30) \quad L_2 \circ L_1(\mathbf{v}) &= L_2(\tilde{\mathbf{v}}) = X\pi_{(AX)^\perp}(X\pi_{A^\perp}(X^{-2}\mathbf{v})) \\ &= X\pi_{(AX)^\perp}(X^{-1}\mathbf{v}) - X\pi_{(AX)^\perp}(XA^T(AA^T)^{-1}AX^{-2}\mathbf{v}).\end{aligned}$$

For the first term on the right of (3.30) we find on expanding $\pi_{(AX)^\perp}$ that

$$(3.31) \quad X\pi_{(AX)^\perp}(X^{-1}\mathbf{v}) = XX^{-1}\mathbf{v} - X^2A(AX^2A^T)^{-1}A\mathbf{v} = \mathbf{v},$$

using the fact that $A\mathbf{v} = \mathbf{0}$ since \mathbf{v} is in A^\perp . For the second term on the right in (3.30), on expanding $\pi_{(AX)^\perp}$ we obtain

$$\begin{aligned}X(XA^T(AA^T)^{-1}AX^{-2}\mathbf{v}) \\ - X(XA^T[(AX^2A^T)^{-1}AX(XA^T)](AA^T)^{-1}AX^{-2}\mathbf{v}) = 0,\end{aligned}$$

using the fact that the product in brackets is the identity. Combining the last three equations proves (3.29) holds. \square

4. AFFINE SCALING VECTOR FIELD IN LEGENDRE TRANSFORM COORDINATES

We take as given a linear program in *standard form*:

$$(4.1) \quad \begin{cases} \text{minimize } \langle \mathbf{c}, \mathbf{x} \rangle, \\ A\mathbf{x} = \mathbf{b}, \\ \mathbf{x} \geq \mathbf{0} \end{cases}$$

such that there exists a feasible solution $\mathbf{x} = (x_1, \dots, x_n)$ with all $x_i > 0$. In part I, Lemma 4.1, it is shown that the *affine scaling vector field* $\mathbf{v}_A(\mathbf{x}; \mathbf{c})$ for a strict standard form linear program is

$$(4.2) \quad \mathbf{v}_A(\mathbf{x}; \mathbf{c}) = -X\pi_{(AX)^\perp}(X\mathbf{c}),$$

where $X = \text{diag}(\mathbf{x})$. The *affine scaling vector* $\mathbf{v}_A(\mathbf{x}; \mathbf{c})$ lies in A^\perp , as may be verified using (4.2). In part I, Lemma 4.1 it is also shown that the vector field $\mathbf{v}_A(\mathbf{x}; \mathbf{c})$ depends only on the component $\pi_{A^\perp}(\mathbf{c})$ of \mathbf{c} in the A^\perp -direction, i.e.,

$$(4.3) \quad \mathbf{v}_A(\mathbf{x}; \mathbf{c}) = \mathbf{v}_A(\mathbf{x}; \pi_{A^\perp}(\mathbf{c})).$$

We have the following result.

Theorem 4.1. *Given a set H of standard form constraints*

$$\begin{cases} A\mathbf{x} = \mathbf{b}, \\ \mathbf{x} \geq \mathbf{0}, \end{cases}$$

having a feasible point \mathbf{x} with all $x_i > 0$, so that $\text{Rel-Int}(P_H) = P_H \cap \text{Int}(\mathbf{R}_+^n)$. Let $\mathbf{v}_A(\mathbf{x}; \mathbf{c}) = -X\pi_{(AX)^\perp}(X\mathbf{c})$ denote the affine scaling vector field associated to the objective function $\langle \mathbf{c}, \mathbf{x} \rangle$. If $\phi_H: \text{Rel-Int}(P_H) \rightarrow A^\perp$ is the Legendre transform map $\phi_H(\mathbf{x}) = -\pi_{A^\perp}(X^{-1}\mathbf{e})$ then

$$(4.4) \quad (d\phi_H)_x(\mathbf{v}_A(\mathbf{x}; \mathbf{c})) = -\pi_{A^\perp}(\mathbf{c}).$$

Proof. The formula for $\mathbf{v}_A(\mathbf{x}; \mathbf{c})$ and Theorem 3.6 yield

$$\begin{aligned} \mathbf{v}_A(\mathbf{x}; \pi_{A^\perp}(\mathbf{c})) &= -X\pi_{(AX)^\perp}(X\pi_{A^\perp}(\mathbf{c})) \\ &= (d\phi_H^{-1})_y(-\pi_{A^\perp}(\mathbf{c})), \end{aligned}$$

where $\mathbf{y} = \phi_H(\mathbf{x})$. Then using (4.3) gives

$$\begin{aligned} (d\phi_H)_x(\mathbf{v}_A(\mathbf{x}; \mathbf{c})) &= (d\phi_H)_x(\mathbf{v}_A(\mathbf{x}; \pi_{A^\perp}(\mathbf{c}))) \\ &= (d\phi_H)_x((d\phi_H^{-1})_y(-\pi_{A^\perp}(\mathbf{c}))) \\ &= -\pi_{A^\perp}(\mathbf{c}), \end{aligned}$$

proving the theorem. \square

This result immediately shows that A -trajectories are parts of straight lines in Legendre transform coordinates.

Corollary 4.1a. *Let a set H of strict standard form constraints be given, and let $T_A(\mathbf{x}; \mathbf{c})$ be the affine scaling trajectory associated to the objective function $\langle \mathbf{c}, \mathbf{x} \rangle$ that goes through $\mathbf{x} \in \text{Rel-Int}(P_H)$. Let $\phi_H: \text{Rel-Int}(P_H) \rightarrow A^\perp$ be the Legendre transform coordinate mapping. If the polytope P_H of feasible solutions is bounded then the image of $T_A(\mathbf{x}; \mathbf{c})$ in Legendre transform coordinates is a straight line in A^\perp in the direction*

$$(4.5) \quad \mathbf{c}^* = -\pi_{A^\perp}(\mathbf{c}).$$

If P_H is unbounded, then the range R_H of ϕ_H is the relative interior of a cone in A^\perp and the image of $T_A(\mathbf{x}; \mathbf{c})$ in Legendre transform coordinates is the intersection of a straight line with direction $\pi_{A^\perp}(\mathbf{c})$ with R_H , and is either a half-line or a line segment.

Proof. Formula (4.4) shows that the transformed vector field is constant in Legendre transform coordinates, so that its integral curves are parts of straight lines. \square

Corollary 4.1a has the immediate consequence that we can define a global metric geometry on $\text{Rel-Int}(P_H)$ of a strict standard form problem having a bounded polytope of feasible solutions such that the A -trajectories are geodesics.

Corollary 4.1b. *Given a set H of strict standard form constraints*

$$\begin{cases} A\mathbf{x} = \mathbf{b}, \\ \mathbf{x} \geq \mathbf{0}, \end{cases}$$

whose polytope P_H of feasible solutions is bounded, there is a global metric $d_H: \text{Rel-Int}(P_H) \times \text{Rel-Int}(P_H) \rightarrow \mathbf{R}$ such that every geodesic with respect to $d_H(\cdot, \cdot)$ is an A -trajectory associated to a suitable objective function. Conversely every A -trajectory for any objective function not constant on P_H is a geodesic of $d(\cdot, \cdot)$. This metric geometry is isometric to Euclidean geometry on \mathbf{R}^d with $d = \dim(A^\perp)$, hence is geodesically complete.

Proof. The Legendre transform coordinate space is A^\perp , using Theorem 3.5. Now we can take any complete metric geometry $\tilde{d}(\cdot, \cdot)$ on A^\perp having straight lines as unique geodesics and pull it back to $\text{Rel-Int}(P_H)$ using the inverse Legendre transform map ϕ_H^{-1} , i.e., we define

$$d_H(\mathbf{x}_1, \mathbf{x}_2) \equiv \tilde{d}(\phi_H(\mathbf{x}_1), \phi_H(\mathbf{x}_2)).$$

We may use Euclidean geometry on A^\perp , or more generally any (global) affine rescaling of Euclidean geometry. \square

Note that with this choice the global metric geometry on the Legendre transform coordinate space arises from a Riemannian geometry. The pullback global metric geometry on $\text{Rel-Int}(P_H)$ then also arises from a Riemannian geometry.

A third consequence of Theorem 4.1 is that each A -trajectory is part of a real algebraic curve.

Corollary 4.1c. *Every A -trajectory of a standard form linear programming problem is part of a real algebraic curve in the linear program's coordinate system.*

Proof. Any standard form linear programming problem H may be brought to strict standard form by dropping all variables x_i which are identically zero on P_H . Consequently it suffices to prove the corollary for strict standard form linear programs only. For strict standard form problems the result follows directly from Corollary 4.1a, since each A -trajectory is the inverse image under a rational mapping of a line segment, a half-line, or a straight line. \square

All P -trajectories are parts of a real algebraic curve, because each P -trajectory is algebraically related to an A -trajectory of a standard form problem.

Corollary 4.1d. *Every P -trajectory of a canonical form linear programming problem is part of a real algebraic curve in the linear program's coordinate system.*

Proof. Part I, Theorem 7.1 showed that any P -trajectory $T_P(\mathbf{x}; \mathbf{c}, A)$ of a canonical form problem is the radial projection of the A -trajectory $T_A(\mathbf{x}; \mathbf{c}, A, \mathbf{0})$ of the associated strict standard form linear program obtained by dropping the constraint $\langle \mathbf{e}, \mathbf{x} \rangle = n$, i.e.,

$$T_P(\mathbf{x}; \mathbf{c}, A) = \{n\mathbf{y}/\langle \mathbf{e}, \mathbf{y} \rangle : \mathbf{y} \in T_A(\mathbf{x}; \mathbf{c}, A, \mathbf{0})\}.$$

Since this is an algebraic relation and this A -trajectory is part of a real algebraic curve by Corollary 4.1c, so is $T_P(\mathbf{x}; \mathbf{c}, A)$. \square

5. PROJECTIVE SCALING VECTOR FIELD IN LEGENDRE TRANSFORM COORDINATES

Consider a linear program in *canonical form*:

$$(5.1) \quad \begin{cases} \text{minimize } \langle \mathbf{c}, \mathbf{x} \rangle, \\ A\mathbf{x} = \mathbf{0}, \\ \langle \mathbf{e}, \mathbf{x} \rangle = n, \\ \mathbf{x} \geq \mathbf{0}, \end{cases}$$

where \mathbf{e} is a feasible solution. Part I, Lemma 4.3 showed that the *projective scaling vector field* $\mathbf{v}_P(\mathbf{x}; \mathbf{c}, A)$ for a canonical form linear program with objective function $\langle \mathbf{c}, \mathbf{x} \rangle$ is

$$(5.2) \quad \mathbf{v}_P(\mathbf{x}; \mathbf{c}, A) = -X\pi_{(AX)^\perp}(X\mathbf{c}) + (1/n)\langle X\mathbf{e}, \pi_{(AX)^\perp}(X\mathbf{c}) \rangle X\mathbf{e},$$

where $X = \text{diag}(\mathbf{x})$. We normally abbreviate $\mathbf{v}_P(\mathbf{x}; \mathbf{c}, A)$ to $\mathbf{v}_P(\mathbf{x}; \mathbf{c})$, omitting the explicit dependence on the constraints. The projective scaling vector field $\mathbf{v}_P(\mathbf{x}; \mathbf{c})$ for \mathbf{x} in $\text{Rel-Int}(P_H)$ lies in $[\frac{A}{\mathbf{e}^T}]^\perp$, as may be verified from (5.2). Part I, Lemma 4.3 showed $\mathbf{v}_P(\mathbf{x}; \mathbf{c})$ depends only on the component $\pi_{A^\perp}(\mathbf{c})$ of \mathbf{c} in the A^\perp -direction, i.e.,

$$(5.3) \quad \mathbf{v}_P(\mathbf{x}; \mathbf{c}) = \mathbf{v}_P(\mathbf{x}; \pi_{A^\perp}(\mathbf{c})).$$

This vector field *does* depend on the component of \mathbf{c} in the \mathbf{e} -direction, even though this component of the objective function remains constant on the polytope P_H due to the constraint $\langle \mathbf{c}, \mathbf{x} \rangle = n$.

Now we compute the projective scaling vector field in Legendre transform coordinates.

Theorem 5.1. *Let H be a set of canonical form linear programming constraints, and let $\mathbf{v}_P(\mathbf{x}; \mathbf{c})$ be the projective scaling vector field associated to the objective function $\langle \mathbf{c}, \mathbf{x} \rangle$ at $\mathbf{x} \in \text{Rel-Int}(P_H)$, given by*

$$(5.4) \quad \mathbf{v}_P(\mathbf{x}; \mathbf{c}) = -X\pi_{(AX)^\perp}(X\mathbf{c}) + (1/n)\langle X\mathbf{e}, \pi_{(AX)^\perp}(X\mathbf{c}) \rangle X\mathbf{e}.$$

If $\phi_H: \text{Rel-Int}(P_H) \rightarrow A^\perp$ is the Legendre transform mapping given by

$$(5.5) \quad \phi_H(\mathbf{x}) = -\pi_{[\frac{A}{\mathbf{e}}]^\perp}(X^{-1}\mathbf{e}),$$

then

$$(5.6) \quad (d\phi_H)_x(\mathbf{v}_P(\mathbf{x}; \mathbf{c})) = -\pi_{[\frac{A}{\mathbf{e}}]^\perp}(\mathbf{c}) - (1/n)\langle X\mathbf{e}, \pi_{(AX)^\perp}(X\mathbf{c}) \rangle \phi(\mathbf{x}).$$

Proof. Apply Theorem 3.2 to the formula for $\mathbf{v}_P(\mathbf{x}; \mathbf{c})$ above to obtain

$$(5.7) \quad (d\phi_H)_x(\mathbf{v}_P(\mathbf{x}; \mathbf{c})) = -\pi_{[\frac{A}{\mathbf{e}}]^\perp}(X^{-1}\pi_{(AX)^\perp}(X\mathbf{c})) \\ + (1/n)\langle X\mathbf{e}, \pi_{(AX)^\perp}(X\mathbf{c}) \rangle \pi_{[\frac{A}{\mathbf{e}}]^\perp}(X^{-1}\mathbf{e}).$$

To simplify the first term on the right of (5.9), we use

$$(5.8) \quad X^{-1}\pi_{(AX)^\perp}(X\mathbf{c}) = X^{-1}(I - XA^T(AX^2A^T)^{-1}AX)X\mathbf{c} \\ = \mathbf{c} - A^T\mathbf{w},$$

where

$$(5.9) \quad \mathbf{w} = (AX^2A^T)^{-1}AX^2\mathbf{c}.$$

Now $\pi_{A^\perp}(A^T\mathbf{w}) = \mathbf{0}$, either by noting that $A^T\mathbf{w}$ lies in the row space of A , or by directly calculating

$$\pi_{A^\perp}(A^T\mathbf{w}) = (I - A^T(AA^T)^{-1}A)A^T(AX^2A^T)^{-1}AX^2\mathbf{c} = \mathbf{0}.$$

Thus (5.8) yields

$$-\pi_{[\frac{A}{\mathbf{e}}]^\perp}(X^{-1}\pi_{(AX)^\perp}(X\mathbf{c})) = -\pi_{[\frac{A}{\mathbf{e}}]^\perp}(\mathbf{c}) + \pi_{(\mathbf{e}^T)^\perp}(\pi_{A^\perp}(A^T\mathbf{w})) \\ = -\pi_{[\frac{A}{\mathbf{e}}]^\perp}(\mathbf{c}),$$

using the fact that $\pi_{[\frac{A}{\mathbf{e}}]^\perp} = \pi_{(\mathbf{e}^T)^\perp}\pi_{A^\perp}$ since $A\mathbf{e} = \mathbf{0}$. Substituting this equality and the identity (5.5) into (5.7) completes the proof. \square

We remark that the vector $\mathbf{w} = (AX^2A^T)^{-1}AX^2\mathbf{c}$ has a natural interpretation as a set of dual variables (see Todd-Burrell [TB]).

Theorem 5.1 shows that the projective scaling vector field in Legendre transform coordinates is a vector sum of the constant vector field $-\mathbf{c}^*$ where $\mathbf{c}^* = \pi_{[\frac{A}{\mathbf{e}}]^\perp}(\mathbf{c})$, together with a vector field that points radially toward or away from the origin $\mathbf{0}$ at each point in the Legendre transform coordinate space. This has the following consequence.

Corollary 5.1 a. *For a canonical form linear program every P -trajectory with objective function $\langle \mathbf{c}, \mathbf{x} \rangle$ viewed in the Legendre transform coordinate space lies in a plane containing the line $L_{\mathbf{c}^*} = \{\mathbf{c}^* t: -\infty < t < \infty\}$, where $\mathbf{c}^* = \pi_{[\frac{\mathbf{A}}{\mathbf{e}}]^\perp}(\mathbf{c})$.*

Proof. Consider the P -trajectory passing through a point $\mathbf{y} = \phi_{\mathbf{H}}(\mathbf{x})$ in Legendre transform coordinates. The vector field $(d\phi_{\mathbf{H}})_{\mathbf{x}}(\mathbf{v}_P(\mathbf{x}; \mathbf{c}))$ has $L_{\mathbf{c}^*}$ as an invariant set, so that if \mathbf{y} is on $L_{\mathbf{c}^*}$ then the P -trajectory stays on $L_{\mathbf{c}^*}$ (see also Theorem 6.1). If \mathbf{y} is not on $L_{\mathbf{c}^*}$ then the plane determined by \mathbf{y} and $L_{\mathbf{c}^*}$ in Legendre transform coordinates is an invariant set of the vector field $(d\phi_{\mathbf{H}})_{\mathbf{x}}(\mathbf{v}_P(\mathbf{x}; \mathbf{c}))$ and the P -trajectory in Legendre transform coordinates $\phi_{\mathbf{H}}(T_P(\mathbf{x}; \mathbf{c}))$ must lie in it. \square

The projective scaling vector field $\mathbf{v}_P(\mathbf{x}; \mathbf{c})$ depends on the component of \mathbf{c} in the \mathbf{e} -direction. An objective function $\langle \mathbf{c}, \mathbf{x} \rangle$ is *normalized* if $\langle \mathbf{c}, \mathbf{x} \rangle \geq 0$ for all feasible solutions and $\langle \mathbf{c}, \mathbf{x} \rangle = 0$ for some feasible solution. Normalized objective functions play a special role in Karmarkar's projective scaling algorithm, and Karmarkar's convergence proof only applies to normalized objective functions. The criterion that an objective function $\langle \mathbf{c}, \mathbf{x} \rangle$ be normalized uniquely determines its component in the \mathbf{e} -direction, as follows (proved in part I, Lemma 4.4). Given any objective function \mathbf{c} of a canonical form linear program there is a unique normalized objective function \mathbf{c}_N such that \mathbf{c}_N lies in \mathcal{A}^\perp , and $\pi_{[\frac{\mathbf{A}}{\mathbf{e}}]^\perp}(\mathbf{c}) = \pi_{[\frac{\mathbf{A}}{\mathbf{e}}]^\perp}(\mathbf{c}_N)$. In fact if $\mathbf{c}^* = \pi_{[\frac{\mathbf{A}}{\mathbf{e}}]^\perp}(\mathbf{c})$ and \mathbf{x}_{opt} is an optimal solution then

$$\mathbf{c}_N = \mathbf{c}^* - (1/n)\langle \mathbf{c}^*, \mathbf{x}_{\text{opt}} \rangle \mathbf{e}.$$

It can be proved that an objective function is normalized if and only if in Legendre transform coordinates all P -trajectories are asymptotically parallel to the central P -trajectory as they approach an optimal solution, see part III.

6. CENTRAL TRAJECTORIES

The Legendre transform coordinate map $\phi_{\mathbf{H}}$ associated to a set of linear programming constraints \mathbf{H} whose polytope $P_{\mathbf{H}}$ of feasible solutions is bounded determines a unique point $\mathbf{x}_{\mathbf{H}}$ such that

$$(6.1) \quad \phi_{\mathbf{H}}(\mathbf{x}_{\mathbf{H}}) = \mathbf{0},$$

which we call the *center* of \mathbf{H} . By Theorem 3.5 a canonical form linear program always has a bounded polytope $P_{\mathbf{H}}$ and the point \mathbf{e} is its center. We call the A -trajectory and P -trajectory through the center \mathbf{e} of a canonical form problem with objective function $\langle \mathbf{c}, \mathbf{x} \rangle$ the *central A -trajectory with objective function $\langle \mathbf{c}, \mathbf{x} \rangle$* , denoted $T_A(\mathbf{c}, A)$, and the *central P -trajectory with objective function $\langle \mathbf{c}, \mathbf{x} \rangle$* , denoted by $T_P(\mathbf{c}, A)$. The trajectory $T_A(\mathbf{c}, A)$ is $T_A(\mathbf{e}; \mathbf{c}, [\frac{\mathbf{A}}{\mathbf{e}}], [\frac{\mathbf{0}}{n}])$ in the notation of part I.

We first show that central A -trajectories contain (and usually coincide with) central P -trajectories.

Theorem 6.1. *The central P -trajectory $T_P(\mathbf{c}, A)$ of a canonical form linear program is contained in the central A -trajectory $T_A(\mathbf{c}, A)$ for the same linear program. They coincide if and only if the projective scaling vector field $\mathbf{v}_P(\mathbf{x}; \mathbf{c})$ does not vanish for all \mathbf{x} on the central A -trajectory $T_A(\mathbf{c}, A)$.*

Proof. Let $\mathbf{c}^* = \pi_{[\frac{\mathbf{c}}{\|\mathbf{c}\|}]^\perp}(\mathbf{c})$. By Corollary 4.1a the central A -trajectory is simply the line $-\mathbf{c}^* t$ in Legendre transform coordinates, i.e.,

$$(6.2) \quad T_A(\mathbf{e}; \mathbf{c}, A) = \{\mathbf{x}: \pi_{[\frac{\mathbf{c}}{\|\mathbf{c}\|}]^\perp}(X^{-1}\mathbf{e}) = -t\mathbf{c}^*, \quad -\infty < t < \infty\}.$$

Let $\mathbf{x}_t = X_t \mathbf{e}$ denote the unique point on $T_A(\mathbf{e}; \mathbf{c}, A)$ with

$$\phi_H(\mathbf{x}_t) = -\pi_{[\frac{\mathbf{c}}{\|\mathbf{c}\|}]^\perp}(X_t^{-1}\mathbf{e}) = -t\mathbf{c}^*.$$

By Theorem 5.1 the projective scaling vector field at \mathbf{x}_t in Legendre transform coordinates is

$$(d\phi_H)_{\mathbf{x}_t}(\mathbf{v}_P(\mathbf{x}; \mathbf{c}, A)) = (-1 + (t/n)\langle X_t \mathbf{e}, \pi_{(AX_t)^\perp} X_t \mathbf{c} \rangle) \mathbf{c}^*.$$

Hence in Legendre transform coordinates the projective scaling differential equation trajectories starting at a point on the line $\mathbf{c}^* t$ remain on the line $\mathbf{c}^* t$. Furthermore the usual existence theorems for ordinary differential equations tell us that the trajectory starting at $\mathbf{0}$ in Legendre transform coordinates extends to the nearest zeros of the projective scaling vector field on $\mathbf{c}^* t$ on both sides of $\mathbf{0}$, so that if $t^+ = \min\{t: t > 0 \text{ and } \mathbf{v}_P(\mathbf{x}_t; \mathbf{c}, A) = \mathbf{0}\}$ and $t^- = \max\{t: t < 0 \text{ and } \mathbf{v}_P(\mathbf{x}_t; \mathbf{c}, A) = \mathbf{0}\}$ then

$$(6.3) \quad T_P(\mathbf{c}, A) = T_P(\mathbf{e}; \mathbf{c}, A) = \{\mathbf{x}_t: -\pi_{[\frac{\mathbf{c}}{\|\mathbf{c}\|}]^\perp}(\mathbf{x}_t) = t\mathbf{c}^* \text{ and } t^+ < t < t^-\}.$$

In particular $T_P(\mathbf{c}, A) = T_A(\mathbf{c}, A)$ if and only if the projective scaling vector field $\mathbf{v}_P(\mathbf{x}; \mathbf{c}, A)$ does not vanish on the central A -trajectory $T_A(\mathbf{c}, A)$. \square

If $\langle \mathbf{c}, \mathbf{x} \rangle$ is a normalized objective function, then we showed (part I, Lemma 4.5) that the projective scaling vector field never vanishes. By Theorem 6.1 the central P -trajectory and the central A -trajectory coincide in this case. If $\langle \mathbf{c}, \mathbf{x} \rangle$ is *not* a normalized objective function then the projective scaling vector field can have at most one critical point, which is always on the central trajectory, and in this case the central P -trajectory is strictly contained in the central A -trajectory; we do not consider this case further.

Now we define central A -trajectories for strict standard form linear programs and give two characterizations of them. Consider a linear program H in strict standard form:

$$(6.4) \quad \begin{cases} \text{minimize } \langle \mathbf{c}, \mathbf{x} \rangle, \\ A\mathbf{x} = \mathbf{b}, \\ \mathbf{x} \geq \mathbf{0} \end{cases}$$

having the feasible solution $\mathbf{x} = (x_1, \dots, x_n)$ with all $x_i > 0$. The *central A -trajectory* $T_A(\mathbf{c}, A, \mathbf{b})$ is defined by

$$(6.5) \quad T_A(\mathbf{c}, A, \mathbf{b}) = \{\mathbf{x}: \pi_{A^\perp}(X^{-1}\mathbf{e}) = -\pi_{A^\perp}(\mathbf{c})t, \\ \text{for all } t \text{ with } \pi_{A^\perp}(\mathbf{c})t \in \text{Rel-Int}(C_H)\}.$$

This definition reduces to (6.2) for a canonical form problem, and also makes sense for standard form problems having an unbounded domain P_H . The central A -trajectory $T_A(\mathbf{c}, A, \mathbf{b})$ so defined is a *curve* if $\mathbf{c}^* = \pi_{A^\perp}(\mathbf{c}) \in \text{Rel-Int}(C_H)$ with $\mathbf{c}^* \neq \mathbf{0}$; it is the *center* $\{\mathbf{x}_H\}$ if $\mathbf{c}^* = \mathbf{0} \in \text{Rel-Int}(C_H)$; it is the *empty set* if $\mathbf{c}^* \notin \text{Rel-Int}(C_H)$.

The central trajectory $T_A(\mathbf{c}, A, \mathbf{b})$ is an affine invariant in the sense that if $\Psi_{D^{-1}}(\mathbf{x}) = D^{-1}\mathbf{x}$ is an affine transformation with all $d_i > 0$ then

$$(6.6) \quad \Psi_{D^{-1}}(T_A(\mathbf{c}, A, \mathbf{b})) = T_A(D\mathbf{c}, AD, \mathbf{b}).$$

This follows from Theorem 3.2.

One characterization of the central A -trajectory $T_A(\mathbf{c}, A, \mathbf{b})$ for a strict standard form linear program is that it is the locus of centers of a parameterized family of standard form linear programs having an extra sliding equality constraint of the form $\langle \mathbf{c}, \mathbf{x} \rangle = \mu$ added.

Theorem 6.2. *Let H denote the set of strict standard form linear program constraints:*

$$\begin{cases} A\mathbf{x} = \mathbf{b} \\ \mathbf{x} \geq \mathbf{0} \end{cases}$$

and let H_μ denote these constraints together with the extra constraint $\langle \mathbf{c}, \mathbf{x} \rangle = \mu$. Set $\mu_+ = \max(\langle \mathbf{c}, \mathbf{x} \rangle : \mathbf{x} \in P_H)$ and $\mu_- = \min(\langle \mathbf{c}, \mathbf{x} \rangle : \mathbf{x} \in P_H)$, and suppose $\mu_- < \mu_+$. Let $\mathbf{x}(\mu)$ denote the center of P_{H_μ} if one exists. Then the central trajectory $T_A(\mathbf{c}, A, \mathbf{b})$ is the set of centers $\mathbf{x}(\mu)$ of P_{H_μ} for $\mu_- < \mu < \mu_+$, i.e.,

$$(6.7) \quad T(\mathbf{c}, A, \mathbf{b}) = \{\mathbf{x}(\mu) : \mu_- < \mu < \mu_+\}.$$

Before beginning the proof we remark that if P_{H_μ} is unbounded for some value of μ , then it is unbounded for all values of μ . In this case there are no centers \mathbf{x}_μ and Theorem 6.2 asserts that $T(\mathbf{c}, A, \mathbf{b})$ is the empty set.

Proof. By definition if \mathbf{x} is in $T_A(\mathbf{c}, A, \mathbf{b})$ then \mathbf{x} is in $\text{Rel-Int}(P_H)$ and

$$(6.8) \quad \phi_H(\mathbf{x}) = -\pi_{A^\perp}(X^{-1}\mathbf{e}) = -\pi_{A^\perp}(\mathbf{c})t$$

for some real t . The condition $\mu^- < \mu^+$ shows that $\langle \mathbf{c}, \mathbf{x} \rangle$ is not constant on P_H , so that for \mathbf{x} in $\text{Rel-Int}(P_H)$ one has $\langle \mathbf{c}, \mathbf{x} \rangle = \mu$ with $\mu_- < \mu < \mu_+$. Now H_μ is a set of strict standard form constraints and $\mathbf{x} \in \text{Rel-Int}(H_\mu)$, and

$$\phi_{H_\mu}(\mathbf{x}) = -\pi_{[A]^\perp}(X^{-1}\mathbf{e}).$$

Set $\tilde{\mathbf{c}} = \pi_{A^\perp}(\mathbf{c})$ and observe that

$$\pi_{[A]^\perp} \equiv \pi_{[\tilde{\mathbf{c}}]^\perp} \equiv \pi_{\tilde{\mathbf{c}}^T} \circ \pi_{A^\perp},$$

where $\pi_{\tilde{\mathbf{c}}^T}$ and π_{A^\perp} commute since $A\tilde{\mathbf{c}} = \mathbf{0}$. Hence one has

$$(6.9) \quad \begin{aligned} \phi_{H_\mu}(\mathbf{x}) &= -\pi_{(\tilde{\mathbf{c}}^T)^\perp}(\pi_{A^\perp}(X^{-1}\mathbf{e})) \\ &= \pi_{(\tilde{\mathbf{c}}^T)^\perp}(\phi_H(\mathbf{x})). \end{aligned}$$

Using (6.8) one obtains

$$\phi_{H_\mu}(\mathbf{x}) = \pi_{(\tilde{\mathbf{c}}^T)^\perp}(-\tilde{\mathbf{c}}t) = \mathbf{0}.$$

Hence H_μ has a center $\mathbf{x}(\mu)$ and $\mathbf{x} = \mathbf{x}(\mu)$. This proves that $T(\mathbf{c}, A, \mathbf{b}) \subseteq \{\mathbf{x}_\mu: \mu_- < \mu < \mu_+\}$. For the reverse inclusion if $\phi_{H_\mu}(\mathbf{x}) = \mathbf{0}$ then (6.9) yields

$$\pi_{(\tilde{\mathbf{c}}^T)^\perp}(\phi_H(\mathbf{x}(\mu))) = \mathbf{0},$$

so $\phi_H(\mathbf{x}(\mu)) = \tilde{\mathbf{c}}t$ for some t . \square

Theorem 6.2 has the important consequence that the objective function value $\mu = \langle \mathbf{c}, \mathbf{x}(\mu) \rangle$ always provides a natural parameterization of the central A -trajectory, i.e., there is at most one point \mathbf{x} on the central A -trajectory with a given value of $\langle \mathbf{c}, \mathbf{x} \rangle$.

Second, we show that all points on the central A -trajectory on one side of the center are the solutions of a parameterized family of nonlinear fixed point problems described by a logarithmic barrier function.

Theorem 6.3. *Let H denote a set of strict standard form linear program constraints*

$$\begin{cases} A\mathbf{x} = \mathbf{b}, \\ \mathbf{x} \geq \mathbf{0}. \end{cases}$$

Suppose that P_H is bounded and that $\langle \mathbf{c}, \mathbf{x} \rangle$ is not constant on P_H . Then the nonlinear minimization problem

$$(6.10) \quad \text{minimize } \langle \mathbf{c}, \mathbf{x} \rangle - \mu \left(\sum_{i=1}^n \log x_i \right)$$

over $\mathbf{x} \in \text{Rel-Int}(P_H)$ attains its minimum at a unique point $\mathbf{x}(\mu)$ for $0 < \mu < \infty$.

The Legendre transform coordinates of $\mathbf{x}(\mu)$ are

$$(6.11) \quad \phi_H(\mathbf{x}(\mu)) = -(1/\mu)\pi_{A^\perp}(\mathbf{c}),$$

so that $\mathbf{x}(\mu)$ lies on the central A -trajectory $T_A(\mathbf{c}, A, \mathbf{b})$.

Proof. Since $\nabla^2 g_\mu(\mathbf{x}) = \mu X^{-2}$ where $X = \text{diag}(x_1, \dots, x_n)$ is positive definite on $\text{Rel-Int}(P_H)$, it follows that $g_\mu(\mathbf{x})$ is strictly convex and hence has at most one critical point in $\text{Rel-Int}(P_H)$, which is a global minimum if it exists. We find Lagrange multiplier conditions for a critical point $\mathbf{x}(\mu)$ of $g_\mu(\mathbf{x})$, for $\mu > 0$. Set

$$L_\mu(\mathbf{x}, \boldsymbol{\lambda}) = g_\mu(\mathbf{x}) - \boldsymbol{\lambda}^T (A\mathbf{x} - \mathbf{b}).$$

The Lagrange multiplier conditions are

$$(6.12a) \quad dL_\mu/d\mathbf{x} = \mathbf{c} - \mu X^{-1}\mathbf{e} - A^T\boldsymbol{\lambda} = \mathbf{0},$$

$$(6.12b) \quad dL_\mu/d\boldsymbol{\lambda} = A\mathbf{x} - \mathbf{b} = \mathbf{0}.$$

Consequently any critical point $\mathbf{x}(\mu)$ must satisfy

$$(6.13) \quad X^{-1}\mathbf{e} = (1/\mu)(\mathbf{c} - A^T\boldsymbol{\lambda}),$$

so that

$$(6.14) \quad \phi_H(\mathbf{x}(\mu)) = -\pi_{A^\perp}(X^{-1}\mathbf{e}) = -(1/\mu)\pi_{A^\perp}(\mathbf{c}).$$

By hypothesis $\pi_{A^\perp}(\mathbf{c}) \neq \mathbf{0}$, and there is a unique point $\mathbf{x}(\mu)$ in $\text{Rel-Int}(P_H)$ satisfying (6.14), which is on the positive part of the central trajectory; furthermore there is then a unique λ such that (6.12) holds, so that $\mathbf{x}(\mu)$ is a critical point. It is the unique global minimum of $g_\mu(\mathbf{x})$, and taking $0 < \mu < \infty$ one obtains all points on the positive side of the center. \square

Theorem 6.3 holds also for standard form linear programs having an unbounded polytope P_H of feasible solutions, provided that the objective function $\langle \mathbf{c}, \mathbf{x} \rangle$ is not constant on P_H and that $-\pi_{A^\perp}(\mathbf{c})$ is in the range $\text{Rel-Int}(C_H)$ of the Legendre transform coordinate mapping, so that (6.14) is solvable and the minimization problem (6.10) then has a unique solution.

7. POWER-SERIES EXPANSIONS FOR A -TRAJECTORIES

Since A -trajectories are parts of real algebraic curves, they have locally convergent power-series expansions about any (nonsingular) point of the curve. We derive two power-series expansions: the first is for an arbitrary A -trajectory of a standard form problem in terms of a parameter measuring Euclidean distance in Legendre transform coordinates. The second power-series expansion is for the central A -trajectory of a canonical form problem at the center, and uses the objective function value as the power-series parameter.

Theorem 7.1. *Given a strict standard form linear program*

$$\begin{cases} \text{minimize } \langle \mathbf{c}, \mathbf{x} \rangle, \\ A\mathbf{x} = \mathbf{b}, \\ \mathbf{x} \geq \mathbf{0}, \end{cases}$$

and feasible point $\mathbf{x}_0 = X_0\mathbf{e}$ with $\mathbf{x}_0 > \mathbf{0}$. Set $\mathbf{c}^ = \pi_{A^\perp}(\mathbf{c})$, and parametrize the A -trajectory $T_A(\mathbf{x}_0; \mathbf{c})$ using Legendre transform coordinates by*

$$(7.1) \quad \phi_H(\mathbf{x}_t) = \phi_H(\mathbf{x}_0) - t\mathbf{c}^*.$$

Then the power-series expansion

$$(7.2) \quad \mathbf{x}_t = \mathbf{x}_0 + \sum_{k=1}^{\infty} \mathbf{v}_k t^k$$

has coefficient vectors $\mathbf{v} = \mathbf{v}_k(\mathbf{c}^)$ computed recursively as follows. Let $V_k = \text{diag}(\mathbf{v}_k)$, and $\Phi_0 = I$, and initialize with*

$$(7.3a) \quad \mathbf{v}_1 = X_0\pi_{(AX_0)^\perp}(X_0\mathbf{c}^*),$$

$$(7.3b) \quad \Phi_1 = -X_0^{-1}V_1.$$

The recursion step determining \mathbf{v}_k and Φ_k is:

$$(7.4a) \quad \mathbf{v}_k = -X_0 \pi_{(AX_0)^\perp} \left(X_0 \sum_{j=1}^{k-1} \pi_{A^\perp} (X_0^{-2} \Phi_j V_{k-j} \mathbf{e}) \right),$$

$$(7.4b) \quad \Phi_k = -X_0^{-1} \sum_{j=0}^{k-1} \Phi_j V_{k-j}.$$

The matrices Φ_k are diagonal matrices.

Proof. Define $X_t = \text{diag}(\mathbf{x}_t)$, and the power-series expansion (7.2) to be determined is equivalent to

$$(7.5) \quad X_t = X_0 + \sum_{k=1}^{\infty} V_k t^k$$

where the V_k are diagonal matrices. The vector $\mathbf{x}_t - \mathbf{x}_0$ lies in the subspace $A^\perp = \{\mathbf{x}: A\mathbf{x} = \mathbf{0}\}$; hence (7.2) yields

$$(7.6) \quad \pi_{A^\perp} (V_k \mathbf{e}) = \mathbf{0} \quad \text{for } k \geq 1.$$

Define the diagonal matrices Φ_k by

$$(7.7) \quad X_t^{-1} = X_0^{-1} \sum_{k=0}^{\infty} \Phi_k t^k.$$

Now (7.5) gives

$$X_t^{-1} = X_0^{-1} \left(I + X_0^{-1} \left(\sum_{k=1}^{\infty} V_k t^k \right) \right)^{-1},$$

since diagonal matrices commute. Comparing the last two equations shows that

$$\left(\sum_{k=0}^{\infty} \Phi_k t^k \right) \left(I + X_0^{-1} \left(\sum_{k=1}^{\infty} V_k t^k \right) \right) = I.$$

Evaluating coefficients of powers of t in this formula yields $\Phi_0 = I$, and

$$\Phi_k + \sum_{j=0}^{k-1} \Phi_j X_0^{-1} V_{k-j} = \mathbf{0}, \quad k \geq 1.$$

This yields the recursion

$$(7.8) \quad \Phi_k = -X_0^{-1} \left(\sum_{j=0}^{k-1} \Phi_j V_{k-j} \right),$$

since diagonal matrices commute, and proves (7.3b) and (7.4b).

Using the definition of Legendre transform coordinates and the definition (7.7) one obtains

$$\begin{aligned}\phi_H(\mathbf{x}_t) &= -\pi_{A^\perp}(X_t^{-1}\mathbf{e}) \\ &= -\pi_{A^\perp}\left(X_0^{-1}\left(I + \sum_{k=1}^{\infty}\Phi_k t^k\right)\mathbf{e}\right) \\ &= \phi_H(\mathbf{x}_0) - \pi_{A^\perp}\left(\sum_{k=1}^{\infty}(X_0^{-1}\Phi_k\mathbf{e})t^k\right).\end{aligned}$$

Comparing this with (7.1) and equating powers of t gives

$$(7.9a) \quad \pi_{A^\perp}(X_0^{-1}\Phi_1\mathbf{e}) = \mathbf{c}^*,$$

$$(7.9b) \quad \pi_{A^\perp}(X_0^{-1}\Phi_k\mathbf{e}) = \mathbf{0}, \quad \text{all } k \geq 2.$$

We have $\Phi_1 = -X_0^{-1}V_1$ by (7.8) so that

$$(7.10) \quad \pi_{A^\perp}(X_0^{-2}V_1\mathbf{e}) = \pi_{A^\perp}(X_0^{-1}\Phi_1\mathbf{e}) = \mathbf{c}^*.$$

The condition (7.6) asserts that $\mathbf{v}_1 = V_1\mathbf{e}$ is in A^\perp , and Theorem 3.6 now shows that (7.10) has a unique solution which is given by

$$\mathbf{v}_1 = X_0\pi_{(AX_0)^\perp}(X_0\mathbf{c}^*),$$

and this proves (7.3a). Next, for $k \geq 2$ substitute (7.8) into (7.9) to obtain

$$(7.11) \quad \pi_{A^\perp}(X_0^{-2}V_k\mathbf{e}) = -\sum_{j=1}^{k-1}\pi_{A^\perp}(X_0^{-2}\Phi_j V_{k-j}\mathbf{e}).$$

Now $\mathbf{v}_k = V_k\mathbf{e}$ is in A^\perp , so applying Theorem 3.6 to (7.11) yields

$$\mathbf{v}_k = -X_0\pi_{(AX_0)^\perp}\left(X_0\sum_{j=1}^{k-1}\pi_{A^\perp}(X_0^{-2}\Phi_j V_{k-j}\mathbf{e})\right),$$

which is (7.9a) and completes the proof. (Note here that $\Phi_j V_{k-j}\mathbf{e}$ is usually not in A^\perp , so one cannot simplify the right-hand side of this equation further using (3.31).) \square

The formulae of Theorem 7.1 simplify considerably when the initial point \mathbf{x}_0 is \mathbf{e} , as happens in a canonical form linear program. The recursive formulae for $k = 1$ are then

$$(7.12a) \quad \mathbf{v}_1 = -\mathbf{c}^*,$$

$$(7.12b) \quad \Phi_1 = -V_1,$$

and for $k \geq 2$ are

$$(7.12c) \quad \mathbf{v}_k = -\pi_{A^\perp} \left(\sum_{j=1}^{k-1} \Phi_j \mathbf{v}_{k-j} \right),$$

$$(7.12d) \quad \Phi_k = - \sum_{j=0}^{k-1} \Phi_j V_{k-j}.$$

In particular only the single projection π_{A^\perp} is involved in computing each \mathbf{v}_k .

Now we derive a power-series formula for the central A -trajectory of a canonical form linear program about the point \mathbf{e} , using the value $z = \langle \mathbf{x}, \mathbf{e} \rangle$ of the objective function as the power-series parameter. The remarks after Theorem 6.2 imply that the objective function value parameter z is a monotone decreasing function of the Legendre transform coordinate parameter t .

Theorem 7.2. *Given a canonical form linear program in \mathbf{R}^n*

$$\begin{cases} \text{minimize } \langle \mathbf{c}^*, \mathbf{x} \rangle, \\ A\mathbf{x} = \mathbf{0}, \\ \langle \mathbf{e}, \mathbf{x} \rangle = n, \\ \mathbf{x} \geq \mathbf{0}, \end{cases}$$

such that $\mathbf{e} = (1, 1, \dots, 1)$ is feasible and \mathbf{c}^ is in $[\frac{A}{\mathbf{e}^T}]^\perp$. Define $z = \langle \mathbf{c}^*, \mathbf{x} \rangle$, and parametrize the central A -trajectory $T_A(\mathbf{e}; \mathbf{c}^*)$ as $\{\mathbf{x}(z)\}$, with $\mathbf{x}(0) = \mathbf{e}$. Then the power-series expansion around $z = 0$ is given by*

$$(7.13) \quad \mathbf{x}(z) = \mathbf{e} + \sum_{k=1}^{\infty} \mathbf{v}_k^* z^k$$

where the coefficient vectors $\mathbf{v}_k^ = \mathbf{v}_k^*(\mathbf{c}^*)$ are computed recursively as follows. Let $V_k^* = \text{diag}(\mathbf{v}_k^*)$ and $\Phi_0^* = I$, initialized with*

$$(7.14a) \quad \alpha_1 = -\langle \mathbf{c}^*, \mathbf{c}^* \rangle^{-1},$$

$$(7.14b) \quad \mathbf{v}_1^* = -\alpha_1 \mathbf{c}^*,$$

$$(7.14c) \quad \Phi_1^* = -V_1^*.$$

The recursion step determines α_k , \mathbf{v}_k^ , and the diagonal matrices Φ_k^* by*

$$(7.15a) \quad \alpha_k = \alpha_1 \left\langle \mathbf{c}^*, \sum_{j=1}^{k-1} \Phi_j^* V_{k-j}^* \mathbf{e} \right\rangle,$$

$$(7.15b) \quad \mathbf{v}_k^* = -\pi_{[\frac{A}{\mathbf{e}^T}]^\perp} \left(\sum_{j=1}^{k-1} \Phi_j^* V_{k-j}^* \mathbf{e} \right) - \alpha_k \mathbf{c}^*,$$

$$(7.15c) \quad \Phi_k^* = - \sum_{j=0}^{k-1} \Phi_j^* V_{k-j}^*.$$

Proof. By definition $\mathbf{x}(z)$ is that point on the central A -trajectory for which

$$(7.16) \quad z = \langle \mathbf{c}^*, \mathbf{x}(z) \rangle,$$

and $\mathbf{x}(0) = \mathbf{e}$. Since $\mathbf{x}(z)$ is on the central A -trajectory one has

$$(7.17a) \quad \phi_H(\mathbf{x}(z)) = -f(z)\mathbf{c}^*,$$

where

$$(7.17b) \quad f(z) = \sum_{k=0}^{\infty} \alpha_k z^k$$

converges in a neighborhood of $z = 0$ since ϕ_H and $\mathbf{x}(z)$ are both analytic mappings. Note that if \mathbf{x}_t is the Legendre transform parameterization of the central A -trajectory, with $\phi_H(\mathbf{x}_t) = -t\mathbf{c}^*$, then (7.17) yields

$$t = t(z) = \sum_{k=0}^{\infty} \alpha_k z^k.$$

This gives a conversion formula from a power-series in t to one in z .

Before beginning the calculations observe first that since $\mathbf{x}(z)$ is in $[\frac{4}{e^7}]^\perp$, (7.13) implies that

$$(7.18) \quad \pi_{[\frac{4}{e^7}]^\perp}(\mathbf{v}_k^*) = \mathbf{v}_k^*.$$

Substitute the power-series expansion (7.13) into (7.16) and equate powers of z to obtain

$$(7.19a) \quad \langle \mathbf{c}^*, \mathbf{v}_1^* \rangle = 1,$$

$$(7.19b) \quad \langle \mathbf{c}^*, \mathbf{v}_k^* \rangle = 0, \quad \text{all } k \geq 2.$$

To evaluate $\phi_H(\mathbf{x}(z))$, let $X_z^* = \text{diag}(\mathbf{x}(z))$ and $V_k^* = \text{diag}(\mathbf{v}_k^*)$. The power-series expansion (7.13) is equivalent to

$$X_z^* = I + \sum_{k=1}^{\infty} V_k^* z^k.$$

Set

$$(7.20) \quad (X_z^*)^{-1} = \sum_{k=0}^{\infty} \Phi_k^* z^k,$$

and one easily finds as in Theorem 7.1 that $\Phi_0^* = I$ and

$$(7.21) \quad \Phi_k^* = - \sum_{j=0}^{k-1} \Phi_j^* V_{k-j}^*, \quad \text{all } k \geq 1,$$

which proves (7.14c) and (7.15c). Now compute

$$\begin{aligned}\phi_H(\mathbf{x}(z)) &= \pi_{[\frac{A}{\epsilon^*}]^\perp}((X_z^*)^{-1}\mathbf{e}) \\ &= -\pi_{[\frac{A}{\epsilon^*}]^\perp}\left(\mathbf{e} + \sum_{k=1}^{\infty}(\Phi_k^*\mathbf{e})z^k\right) \\ &= -\sum_{k=1}^{\infty}\pi_{[\frac{A}{\epsilon^*}]^\perp}(\Phi_k^*\mathbf{e})z^k.\end{aligned}$$

Equating this with (7.17) term by term yields $\alpha_0 = 0$ and

$$(7.22) \quad \alpha_k \mathbf{c}^* = \pi_{[\frac{A}{\epsilon^*}]^\perp}(\Phi_k^*\mathbf{e}), \quad \text{all } k \geq 1.$$

Now (7.21) gives $\Phi_1^* = -V_1^*$ hence $\Phi^*\mathbf{e} = -\mathbf{v}_1^*$ so that the last equation for $k = 1$ yields

$$(7.23) \quad \alpha_1 \mathbf{c}^* = -\pi_{[\frac{A}{\epsilon^*}]^\perp}(\mathbf{v}_1^*).$$

Multiplying this equation by $(\mathbf{c}^*)^T$ and using (7.19a) yields

$$\alpha_1 \langle \mathbf{c}^*, \mathbf{c}^* \rangle = -\langle \mathbf{c}^*, \pi_{[\frac{A}{\epsilon^*}]^\perp}(\mathbf{v}_1^*) \rangle = -\langle \mathbf{c}^*, \mathbf{v}_1^* \rangle = -1.$$

This proves (7.14a). Also $\pi_{[\frac{A}{\epsilon^*}]^\perp}(\mathbf{v}_1^*) = \mathbf{v}_1^*$, so that (7.23) gives $\mathbf{v}_1^* = -\alpha_1 \mathbf{c}^*$, which proves (7.14b). To derive the recursion step, multiply (7.22) on the left by $(\mathbf{c}^*)^T$ to get

$$(7.24) \quad \begin{aligned}\alpha_k \langle \mathbf{c}^*, \mathbf{c}^* \rangle &= \langle \mathbf{c}^*, \pi_{[\frac{A}{\epsilon^*}]^\perp}(\Phi_k^*\mathbf{e}) \rangle \\ &= -\left\langle \mathbf{c}^*, \sum_{j=0}^{k-1} \Phi_j^* V_{k-j}^* \mathbf{e} \right\rangle,\end{aligned}$$

using (7.21) and the fact that \mathbf{c}^* is in $[\frac{A}{\epsilon^*}]^\perp$. Now we note that

$$\langle \mathbf{c}^*, \pi_{[\frac{A}{\epsilon^*}]^\perp}(\Phi_0^* V_k^* \mathbf{e}) \rangle = \langle \mathbf{c}^*, \Phi_0^* V_k^* \mathbf{e} \rangle = \langle \mathbf{c}^*, \mathbf{v}_k^* \rangle = 0,$$

using (7.18) and (7.19b). Substituting this in (7.24) yields

$$\alpha_k \langle \mathbf{c}^*, \mathbf{c}^* \rangle = -\left\langle \mathbf{c}^*, \sum_{j=1}^{k-1} \Phi_{k-j}^* V_j^* \mathbf{e} \right\rangle,$$

which proves (7.15a). Finally (7.21) and (7.22b) give

$$-\alpha_k \mathbf{c}^* = \pi_{[\frac{A}{\epsilon^*}]^\perp}(V_k^* \mathbf{e}) + \sum_{j=1}^{k-1} \pi_{[\frac{A}{\epsilon^*}]^\perp}(\Phi_{k-1}^* V_j^* \mathbf{e}).$$

Using (7.18) this yields

$$\mathbf{v}_k^* = -\pi_{[\frac{A}{\epsilon^*}]^\perp}\left(\sum_{j=1}^{k-1} \Phi_{k-j}^* V_j^* \mathbf{e}\right) - \alpha_k \mathbf{c}^*,$$

which proves (7.15b) and completes the proof. \square

The formula $t = \sum_{k=0}^{\infty} \alpha_k z^k$ expresses t in terms of z . By reversion of power-series there is an inverse formula

$$z = z(t) = \sum_{k=0}^{\infty} \beta_k t^k.$$

In fact one has

$$\begin{aligned} z = \langle \mathbf{c}^*, \mathbf{x}_t \rangle &= \left\langle \mathbf{c}^*, \mathbf{e} + \sum_{k=1}^{\infty} \mathbf{v}_k t^k \right\rangle \\ &= \sum_{k=1}^{\infty} \langle \mathbf{c}^*, \mathbf{v}_k \rangle t^k, \end{aligned}$$

so that $\beta_0 = 0$ and

$$\beta_k = \langle \mathbf{c}^*, \mathbf{v}_k \rangle, \quad \text{all } k \geq 1,$$

where the \mathbf{v}_k may be computed using the formulae (7.12) with A replaced by $[\frac{A}{\epsilon^T}]$.

One can consider “higher-order” analogues of Karmarkar’s projective scaling algorithm, based on taking steps following truncated power-series expansions to the central A -trajectory. The choice of the power-series parameter plays a critical role in the performance of such algorithms. A second-order power-series implementation of these ideas, using the coordinate for which the objective function decreases (locally) most rapidly as the local power-series parameter, is described in [AKRV]. An average 25% decrease in the number of iterations of this second-order method compared to the projective scaling method was empirically observed on a standard set of 30 test problems. Other power-series implementations are described in [KLSW].

8. A -TRAJECTORIES FOR FULL RANK LINEAR PROGRAMS

It is possible to define A -trajectories for an arbitrary linear program. We consider in detail the special case of a linear program in \mathbf{R}^n in *inequality form*:

$$(8.1) \quad \begin{cases} \text{minimize } \langle \mathbf{c}, \mathbf{x} \rangle - c_0 \\ \langle \mathbf{a}_j, \mathbf{x} \rangle \geq b_j, & 1 \leq j \leq m, \end{cases}$$

which also satisfy the conditions:

(H1) *The polytope P_H is full dimensional.*

(H2) *The vectors $[\mathbf{a}_1, \dots, \mathbf{a}_m]$ span \mathbf{R}^n .*

A linear program satisfying (H2) is of full rank; the condition (H2) automatically holds if P_H is bounded. We define A -trajectories for this linear program by finding a one-to-one affine mapping $\mathbf{J}: \mathbf{R}^n \rightarrow \mathbf{R}^m$ which maps the linear program into a standard form linear program in \mathbf{R}^m , whose equality constraints

$$(8.2) \quad A\mathbf{y} = \mathbf{b}$$

describe the image $\{\mathbf{J}(\mathbf{x}) : \mathbf{x} \in \mathbf{R}^n\}$, and which has an objective function $\langle \tilde{\mathbf{c}}, \mathbf{y} \rangle - \tilde{c}_0$ for which

$$(8.3) \quad \langle \tilde{\mathbf{c}}, \mathbf{J}(\mathbf{x}) \rangle - \tilde{c}_0 = \langle \mathbf{c}, \mathbf{x} \rangle - c_0$$

for all $\mathbf{x} \in \mathbf{R}^n$. We define the A -trajectory $T_A(\mathbf{x}_0; \mathbf{c}, H)$ to be the pullback under \mathbf{J}^{-1} of the A -trajectory of the standard form problem through $\mathbf{J}(\mathbf{x}_0)$, which is $\mathbf{J}^{-1}(T_A(\mathbf{J}(\mathbf{x}_0); \tilde{\mathbf{c}}, A, \mathbf{b}))$.

To verify that this definition makes sense we must show that we can find a suitable \mathbf{J} and $\langle \tilde{\mathbf{c}}, \mathbf{y} \rangle - \tilde{c}_0$, and that $T_A(\mathbf{x}_0; \mathbf{c}, H)$ is well defined independent of the choice of \mathbf{J} and $\langle \tilde{\mathbf{c}}, \mathbf{y} \rangle - \tilde{c}_0$. A suitable mapping $\mathbf{y} = \mathbf{J}(\mathbf{x})$ is given using the *slack variables*:

$$(8.4) \quad y_j = \langle \mathbf{a}_j, \mathbf{x} \rangle - b_j,$$

since the transformed inequality constraints are now

$$(8.5) \quad y_j \geq 0, \quad 1 \leq j \leq m,$$

so the problem is in strict standard form. Since \mathbf{J} is one-to-one we can find a projection $\boldsymbol{\pi} \circ \mathbf{J} = \mathbf{I}$ and define $\langle \tilde{\mathbf{c}}, \mathbf{y} \rangle - \tilde{c}_0 = \langle \mathbf{c}, \boldsymbol{\pi}(\mathbf{y}) \rangle - c_0$. We show that the trajectories are well defined independent of the choice of \mathbf{J} and $\tilde{\mathbf{c}}$ in the proof of Theorem 8.1 below.

A second definition of A -trajectories for linear programs (8.1) which satisfy conditions (H1), (H2), uses rescaling ideas. Associate to this linear program's constraints H the *logarithmic barrier function*

$$(8.6) \quad f_H(\mathbf{x}) = - \sum_{j=1}^m \log(\langle \mathbf{a}_j, \mathbf{x} \rangle - b_j).$$

In this case the Legendre transform coordinates $\phi_H(\mathbf{x})$ are given by

$$(8.7) \quad \phi_H(\mathbf{x}) = \nabla f_H(\mathbf{x}) = - \sum_{j=1}^m \frac{\mathbf{a}_j}{\langle \mathbf{a}_j, \mathbf{x} \rangle - b_j}$$

and its Hessian is

$$(8.8) \quad \nabla^2 f_H(\mathbf{x}) = \nabla \phi_H(\mathbf{x}) = \sum_{j=1}^m \frac{1}{(\langle \mathbf{a}_j, \mathbf{x} \rangle - b_j)^2} \mathbf{a}_j \mathbf{a}_j^T$$

which is positive definite for all $\mathbf{x} \in \text{Int}(P_H)$. At a given point \mathbf{x}_0 the Taylor-series approximation to $f_H(\mathbf{x})$ is

$$(8.9) \quad f_H(\mathbf{x}) = f_H(\mathbf{x}_0) + \langle \nabla f_H(\mathbf{x}_0), \mathbf{x} - \mathbf{x}_0 \rangle + \frac{1}{2} (\mathbf{x} - \mathbf{x}_0)^T \nabla^2 f_H(\mathbf{x}_0) (\mathbf{x} - \mathbf{x}_0) + O(\|\mathbf{x} - \mathbf{x}_0\|^3).$$

Now choose an affine rescaling of variables to make the quadratic term in the Taylor-series spherical. Set $\tilde{\mathbf{x}} = \mathbf{J}(\mathbf{x})$ with $\mathbf{J}(\mathbf{x}) = K(\mathbf{x} - \mathbf{x}_0)$, where K is an invertible matrix chosen so that

$$H = \nabla^2 f_H(\mathbf{x}_0) = K^T K,$$

which is always possible since H is a positive-definite symmetric matrix [HK, p. 247]. The transformed potential function has the Taylor-series expansion:

$$f_H(\tilde{\mathbf{x}}) = f_H(\mathbf{x}_0) + \langle (K^{-1}) \nabla f_H(\mathbf{x}_0), \tilde{\mathbf{x}} \rangle + \langle \tilde{\mathbf{x}}, \tilde{\mathbf{x}} \rangle + O(\|\tilde{\mathbf{x}}\|^3).$$

Then

$$(8.10) \quad \langle \mathbf{c}, \mathbf{x} \rangle - c_0 = \langle (K^{-1})^T \mathbf{c}, \tilde{\mathbf{x}} \rangle + (\langle \mathbf{c}, \mathbf{x}_0 \rangle - c_0).$$

The *affine scaling direction field* is the pullback under \mathbf{J}^{-1} of the negative of the transformed objective function direction $-(K^{-1})^T \mathbf{c}$, which is

$$(8.11) \quad \begin{aligned} \mathbf{v}_A(\mathbf{x}_0; \mathbf{c}, H) &\equiv (\mathbf{J}^{-1})_* (-(K^{-1})^T \mathbf{c}) \\ &= -K^{-1} (K^{-1})^T \mathbf{c} \\ &= -(\nabla^2 f_H(\mathbf{x}_0))^{-1} \mathbf{c}. \end{aligned}$$

The A -trajectory $T_A(\mathbf{x}_0; \mathbf{c}, H)$ is obtained by integrating this vector field starting at \mathbf{x}_0 .

A third definition of A -trajectories uses Legendre transform coordinates:

$$(8.12) \quad T_A(\mathbf{x}_0; \mathbf{c}, H) = \{\mathbf{x} \in \text{Int}(P_H) : \phi_H(\mathbf{x}) = \phi_H(\mathbf{x}_0) + t\mathbf{c} \text{ for some } t \in \mathbf{R}\}.$$

These three definitions are equivalent.

Theorem 8.1. *These three definitions define the same set of A -trajectories.*

Proof. (i) \Leftrightarrow (iii). Start with the first definition, and choose any injective affine mapping $\mathbf{J}: \mathbf{R}^n \rightarrow \mathbf{R}^m$ that maps the linear program to the standard form problem with constraints $\tilde{H} = \mathbf{J}(H)$ given by

$$\begin{cases} \mathbf{A}\mathbf{y} = \mathbf{b}, \\ y_j \geq 0, \quad 1 \leq j \leq m, \end{cases}$$

and objective function $\langle \tilde{\mathbf{c}}, \mathbf{y} \rangle - \tilde{c}_0$ satisfying

$$\langle \tilde{\mathbf{c}}, \mathbf{J}(\mathbf{x}) \rangle - \tilde{c}_0 = \langle \mathbf{c}, \mathbf{x} \rangle - c_0$$

for all $\mathbf{x} \in \mathbf{R}^n$. The associated Legendre transform mapping is

$$\phi_{\tilde{H}}(\mathbf{y}) = -\pi_{A^\perp}(Y^{-1}\mathbf{e})$$

where $Y = \text{diag}(y_1, \dots, y_m)$. By Theorem 3.4 the following diagram commutes:

$$(8.13) \quad \begin{array}{ccc} \text{Rel-Int}(P_H) & \xrightarrow{\mathbf{J}} & \text{Rel-Int}(P_{\tilde{H}}) \\ \downarrow \phi_H & & \downarrow \phi_{\tilde{H}} \\ \text{Rel-Int}(C_H) & \xleftarrow{\mathbf{J}^*} & \text{Int}(C_{\tilde{H}}) \end{array}$$

Here \mathbf{J}^* is the *adjoint mapping* to \mathbf{J} , defined by

$$\langle \mathbf{J}^*(\mathbf{y}), \mathbf{x} \rangle = \langle \mathbf{y}, \mathbf{J}(\mathbf{x}) - \mathbf{J}(\mathbf{0}) \rangle$$

for all $\mathbf{x} \in \mathbf{R}^n$, $\mathbf{y} \in \mathbf{R}^m$. Theorem 3.4 also asserts that the mappings are one-to-one and onto in (8.13).

By Corollary 4.1a the image of the A -trajectories $T_A(\mathbf{y}; \tilde{\mathbf{c}}, A, \mathbf{b})$ under ϕ_H is a family of parallel lines with slope $\pi_{A^\perp}(\tilde{\mathbf{c}})$. Since \mathbf{J}^* is affine and one-to-one its image in \mathbf{R}^n is also a family of parallel lines. To show these lines coincide with those given by definition (iii) it suffices to prove that the slope of these lines is \mathbf{c} . By definition of the adjoint we have

$$\langle \mathbf{J}^*(\pi_{A^\perp}(\tilde{\mathbf{c}})), \mathbf{x} \rangle = \langle \pi_{A^\perp}(\tilde{\mathbf{c}}), \mathbf{J}(\mathbf{x}) - \mathbf{J}(\mathbf{0}) \rangle,$$

for all $\mathbf{x} \in \mathbf{R}^n$. Since $\mathbf{J}(\mathbf{x}) - \mathbf{J}(\mathbf{0})$ lies in A^\perp ,

$$\langle \pi_{A^\perp}(\tilde{\mathbf{c}}), \mathbf{J}(\mathbf{x}) - \mathbf{J}(\mathbf{0}) \rangle = \langle \tilde{\mathbf{c}}, \mathbf{J}(\mathbf{x}) - \mathbf{J}(\mathbf{0}) \rangle.$$

Comparing this with (8.3) yields

$$\langle \mathbf{J}^*(\pi_{A^\perp}(\tilde{\mathbf{c}})), \mathbf{x} \rangle = \langle \mathbf{c}, \mathbf{x} \rangle + c_1,$$

for all $\mathbf{x} \in \mathbf{R}^n$, which implies that $c_1 = 0$ and $\mathbf{J}^*(\pi_{A^\perp}(\tilde{\mathbf{c}})) = \mathbf{c}$, completing the argument.

(ii) \Leftrightarrow (iii). We compute the affine scaling vector field $\mathbf{v}_A = \mathbf{v}_A^{(3)}(\mathbf{x}; \mathbf{c}, H)$ associated to the third definition, which is defined by

$$\nabla f_H(\mathbf{x} + \varepsilon \mathbf{v}_A) = \phi_H(\mathbf{x}) - \varepsilon \mathbf{c} + O(\varepsilon^2).$$

This just asserts that $\nabla \phi_H(\mathbf{x}) \mathbf{v}_A = -\mathbf{c}$, which by (8.8) yields

$$\mathbf{v}_A = -(\nabla^2 f_H(\mathbf{x}))^{-1} \mathbf{c}.$$

This coincides with (8.11), so that the second and third definitions are equivalent. \square

One can similarly define A -trajectories for a general linear program of the form (8.1), which may have a lower-dimensional polytope of feasible solutions but is assumed to be of full rank. The analogue of the first definition above is as follows. The full rank condition guarantees that the slack variable transformation $\mathbf{J}: \mathbf{R}^n \rightarrow \mathbf{R}^{m^*}$ given by

$$y_j = \langle \mathbf{a}_j, \mathbf{x} \rangle - b_j, \quad 1 \leq j \leq m^*,$$

is one-to-one. It transforms the linear program (8.1) to a standard form linear program in \mathbf{R}^{m^*} , and one defines A -trajectories to be the pullback by \mathbf{J}^{-1} of the A -trajectories of this standard form problem. There are also extensions of the second and third definitions of A -trajectories given earlier to this case. The extension of the third definition is

$$(8.14) \quad T_A(\mathbf{x}_0; \mathbf{c}, H) = \{\mathbf{x} \in \text{Rel-Int}(P_H): \phi_H(\mathbf{x}) = \phi_H(\mathbf{x}_0) + t\mathbf{c}^* \text{ for } t \in \mathbf{R}\}$$

where $\mathbf{c}^* = \pi_{D_H}(\mathbf{c})$. The obvious extension of Theorem 8.1 holds for all full rank linear programs.

One can define A -trajectories for nonfull rank linear programs, using (8.14). In this case the resulting A -trajectories are not curves but are higher-dimensional objects; we omit the details.

9. CENTRAL A -TRAJECTORIES FOR FULL RANK LINEAR PROGRAMS

We define and study central A -trajectories for an inequality form linear program:

$$(9.1) \quad \begin{cases} \text{minimize } \langle \mathbf{c}, \mathbf{x} \rangle - c_0, \\ \langle \mathbf{a}_j, \mathbf{x} \rangle \geq b_j, & 1 \leq j \leq m, \end{cases}$$

which is of full rank. We define the *central A -trajectory* $T_A(\mathbf{c}, H)$ by

$$(9.2) \quad T_A(\mathbf{c}, H) = \{\mathbf{x} \in \text{Int}(P_H) : \phi_H(\mathbf{x}) = t\mathbf{c} \text{ for some } t \in \mathbf{R}\}.$$

An alternate definition is that if $\mathbf{J} : \mathbf{R}^n \rightarrow \mathbf{R}^m$ transforms the linear program (9.1) to the standard form linear program

$$\begin{cases} \text{minimize } \langle \tilde{\mathbf{c}}, \mathbf{y} \rangle - \tilde{c}_0, \\ A\mathbf{y} = \mathbf{b}, \\ y_j \geq 0, & 1 \leq j \leq m, \end{cases}$$

then the central A -trajectory $T_A(\mathbf{c}, H)$ is the pullback by \mathbf{J}^{-1} of the central A -trajectory of this transformed problem, i.e.,

$$T_A(\mathbf{c}, H) = \mathbf{J}^{-1}(T_A(\tilde{\mathbf{c}}, A, \mathbf{b})).$$

The equivalence of these two definitions easily follows from (8.13).

Adding (or removing) constraints of the form $\pm \langle \mathbf{c}, \mathbf{x} \rangle \geq b$ does not move the central A -trajectory with objective function $\langle \mathbf{c}, \mathbf{x} \rangle$.

Theorem 9.1. *Let H be a full rank set of constraints*

$$\langle \mathbf{a}_i, \mathbf{x} \rangle \geq b_i, \quad 1 \leq i \leq m,$$

in \mathbf{R}^n . Let H' denote H together with an additional set of constraints of the form

$$\begin{aligned} \langle \mathbf{c}, \mathbf{x} \rangle &\geq b_j^+, & 1 \leq j \leq m^+, \\ \langle -\mathbf{c}, \mathbf{x} \rangle &\geq b_j^-, & 1 \leq j \leq m^-, \end{aligned}$$

which are nonsingular constraints in H' . Then the central A -trajectory $T_A(\mathbf{c}; H')$ is contained in the central A -trajectory $T_A(\mathbf{c}, H)$.

Proof. Since $P_{H'} \subseteq P_H$ we have $D_{H'} \subseteq D_H$ and the nonsingularity hypothesis guarantees that $D_{H'} \equiv D_H$, so that

$$\mathbf{c}_{H'}^* \equiv \pi_{D_{H'}}(\mathbf{c}) = \pi_{D_H}(\mathbf{c}) \equiv \mathbf{c}_H^*.$$

For any \mathbf{x} in $\text{Rel-Int}(P_H)$ one has

$$\phi_{H'}(\mathbf{x}) = \pi_{D_H} \left(- \sum_{i=1}^m \frac{\mathbf{a}_i}{\langle \mathbf{a}_i, \mathbf{x} \rangle - b_i} - \sum_{j=1}^{m^+} \frac{\mathbf{c}}{\langle \mathbf{c}, \mathbf{x} \rangle - b_j^+} + \sum_{j=1}^{m^-} \frac{\mathbf{c}}{\langle \mathbf{c}, \mathbf{x} \rangle + b_j^-} \right).$$

This gives

$$(9.3) \quad \phi_{H'}(\mathbf{x}) = \phi_H(\mathbf{x}) + \mu(\mathbf{x})\mathbf{c}_{H'}^*,$$

where

$$\mu(\mathbf{x}) = - \sum_{j=1}^{m^+} \frac{1}{\langle \mathbf{c}, \mathbf{x} \rangle - b_j^+} + \sum_{j=1}^{m^-} \frac{1}{\langle \mathbf{c}, \mathbf{x} \rangle + b_j^-}.$$

Consequently if $\mathbf{x} \in T_A(\mathbf{c}, H')$ then $\phi_{H'}(\mathbf{x}) = t\mathbf{c}_{H'}^* = t\mathbf{c}_H^*$ which with (9.3) shows that $\phi_H(\mathbf{x}) = t'\mathbf{c}_H^*$ for some t' , so that $\mathbf{x} \in T_A(\mathbf{c}, H)$. \square

Although the central trajectory itself is not moved by adding extra constraints of the form $\pm \langle \mathbf{c}, \mathbf{x} \rangle \geq b_i$, added constraints generally move the center of H' to a different place along the central trajectory from that of H , assuming that H and H' have centers. As mentioned earlier, Renegar [Re] develops a linear programming algorithm that approximately follows a sequence of centers produced in this manner along the central trajectory to an optimal point.

Now we show that the limit point \mathbf{x}_∞ of the central A -trajectory $T_A(\mathbf{c}, H)$ on the boundary $\partial P_H = P_H - \text{Rel-Int}(P_H)$ is the center of the largest dimensional face of ∂P_H containing \mathbf{x}_∞ .

Theorem 9.2. *Let H be a full rank set of constraints*

$$(9.4) \quad \langle \mathbf{a}_i, \mathbf{x} \rangle \geq b_i, \quad 1 \leq i \leq m,$$

in \mathbf{R}^n . Suppose that the central A -trajectory $T_A(\mathbf{c}; H)$ is nonempty with

$$(9.5) \quad T_A(\mathbf{c}, H) = \{\mathbf{x}_t; \phi_H(\mathbf{x}_t) = t\mathbf{c}^* \text{ with } t^- < t < \infty\}.$$

Then $\mathbf{x}_\infty = \lim_{t \rightarrow \infty} \mathbf{x}_t$ exists. Set $\mu = \langle \mathbf{c}, \mathbf{x}_\infty \rangle$, and let H' be H together with the constraints

$$(9.6a) \quad \langle \mathbf{c}, \mathbf{x} \rangle \geq \mu,$$

$$(9.6b) \quad -\langle \mathbf{c}, \mathbf{x} \rangle \geq -\mu.$$

Then \mathbf{x}_∞ is the center of $P_{H'}$. Furthermore $P_{H'}$ is the unique face of P_H containing \mathbf{x}_∞ in its relative interior.

Proof. If $\mathbf{c}^* = \pi_{D_H}(\mathbf{c}) = \mathbf{0}$ then $T_A(\mathbf{c}, H) = \mathbf{x}_H$ is the center of H , and $\mathbf{x}_t = \mathbf{x}_H$ so $\mathbf{x}_\infty = \mathbf{x}_H$. Also $\langle \mathbf{c}, \mathbf{x} \rangle$ is constant on H so the constraints (9.6) are singular constraints, hence $P_{H'} = P_H$ and the theorem holds trivially in this case.

Next suppose that $\mathbf{c}_H^* \neq \mathbf{0}$, so that $T_A(\mathbf{c}; H)$ is a curve by the full rank hypothesis. Since Legendre transform coordinates only become unbounded approaching $\partial P_H = P_H - \text{Rel-Int}(P_H)$, and since \mathbf{x}_t parametrizes an algebraic curve, the limit \mathbf{x}_∞ exists.

Let H_n denote the set of nonsingular constraints of H , which we may suppose without loss of generality are

$$\langle \mathbf{a}_i, \mathbf{x} \rangle \geq b_i, \quad 1 \leq i \leq m^*.$$

Then the feasible direction subspace $D_H = [\mathbf{a}_{m^*+1}, \dots, \mathbf{a}_m]^\perp$. The added constraints (9.7) are singular constraints of H' and its feasible direction subspace is

$$D_{H'} = [\mathbf{a}_{m^*+1}, \dots, \mathbf{a}_m, \mathbf{c}]^\perp = [\mathbf{a}_{m^*+1}, \dots, \mathbf{a}_m, \mathbf{c}_H^*]^\perp.$$

Hence $\pi_{D_{H'}} = \pi_{(\mathbf{c}^*)^\perp} \pi_{D_H}$, since $\mathbf{c}^* \in D_H$ implies that $\pi_{(\mathbf{c}^*)^\perp}$ commutes with π_{D_H} . Consequently by definition of \mathbf{x}_t we have

$$\pi_{D_{H'}} \phi_H(\mathbf{x}_t) = \pi_{(\mathbf{c}^*)^\perp} (\pi_{D_H} (\phi_H(\mathbf{x}_t))) = \pi_{(\mathbf{c}^*)^\perp} (t\mathbf{c}^*) = \mathbf{0}.$$

Then

$$(9.7) \quad \pi_{H'}(\mathbf{x}_\infty) = \lim_{t \rightarrow \infty} \pi_{D_{H'}}(\phi_H(\mathbf{x}_t)) = \mathbf{0}.$$

The definition of μ guarantees that \mathbf{x}_∞ is in $P_{H'}$, and (9.7) implies that \mathbf{x}_∞ must be in $\text{Rel-Int}(P_{H'})$ and is the center $\mathbf{x}_{H'}$ of H' .

Finally the constraints (9.6) are just the equality constraint $\langle \mathbf{c}, \mathbf{x} \rangle = \mu$. This constraint does not contain any relative interior point of P_H ; hence the constraints H' must cut out a face of P_H , so $P_{H'}$ is a face of P_H . Each point \mathbf{x} of P_H lies in the relative interior of a unique face of P_H , so the theorem follows. \square

Since the range space of R_H of the Legendre transform coordinate mapping is the relative interior of a cone, it follows that in (9.5) the limiting value t^- must be either 0 or $-\infty$. If $t^- = 0$ then $\lim_{t \rightarrow 0^+} \mathbf{x}_t$ does not exist, and the corresponding central trajectory is unbounded in the linear program's coordinates.

10. CENTRAL TRAJECTORIES AND LINEAR PROGRAMMING DUALITY

There is a simple rational mapping from the central trajectory of a linear program to the central trajectory of its dual linear program. This mapping appears in Osborne [Os, Theorem 2.2], where it is stated in terms of logarithmic barrier functions.

Consider a linear program in \mathbf{R}^n in inequality form:

$$(P) \quad \begin{cases} \text{minimize } \langle \mathbf{c}, \mathbf{x} \rangle, \\ A^T \mathbf{x} \geq \mathbf{b}, \end{cases}$$

where $A = [\mathbf{a}_1, \dots, \mathbf{a}_m]$. The corresponding dual linear program in \mathbf{R}^m is

$$(D) \quad \begin{cases} \text{minimize } -\langle \mathbf{b}, \mathbf{y} \rangle, \\ A\mathbf{y} = \mathbf{c}, \\ \mathbf{y} \geq \mathbf{0}. \end{cases}$$

We say that the dual linear programs (P) and (D) are *transverse* if both objective functions are nonconstant on their polytopes of feasible solutions. In the case that (P) has an interior feasible point the transversality condition is that $\mathbf{c} \neq \mathbf{0}$ and $\pi_{(A^T)^\perp}(\mathbf{b}) \neq \mathbf{0}$.

Let $T_+^{(P)}$ denote the positive half of the central trajectory of (P) , defined by

$$T_+^{(P)} = \{\mathbf{x}(t): \phi_{H_P}(\mathbf{x}(t)) = -t\mathbf{c}, 0 < t < \infty\},$$

where H_P denotes the constraints of (P) , and $T_+^{(D)}$ denotes the positive half of the central trajectory of (D) , defined by

$$T_+^{(D)} = \{\mathbf{y}(t): \phi_{H_D}(\mathbf{y}(t)) = t\pi_{A^\perp}(\mathbf{b}), 0 < t < \infty\},$$

where H_D denotes the constraints of (D) .

Theorem 10.1. *Let (P) and (D) be a pair of dual linear programs such that (P) has an interior feasible point and the pair (P) and (D) are transverse. If $\mathbf{x}(t) \in T_+^{(P)}$ and $\mathbf{u}(t) = A^T \mathbf{x}(t) - \mathbf{b}$ are its corresponding slack variables then the vector $\mathbf{y}(t) \in T_+^{(D)}$ satisfies*

$$(10.1) \quad u_j(t)y_j(t) = 1/t, \quad 1 \leq i \leq m.$$

Proof. Since (P) has an interior feasible point, one has

$$(10.2) \quad \phi_{H_P}(\mathbf{x}(t)) = -\sum_{j=1}^m \frac{\mathbf{a}_j}{\langle \mathbf{a}_j, \mathbf{x}(t) \rangle - b_j} = -A\mathbf{u}(t)^{-1} = -t\mathbf{c},$$

where by definition

$$\mathbf{u}(t)^{-1} = \left(\frac{1}{u_1(t)}, \frac{1}{u_2(t)}, \dots, \frac{1}{u_m(t)} \right).$$

Now define $\mathbf{y} = (1/t)\mathbf{u}(t)^{-1}$, and observe that $\mathbf{y} > \mathbf{0}$ and

$$A\mathbf{y} = -(1/t)\phi_{H_P}(\mathbf{x}(t)) = \mathbf{c},$$

so \mathbf{y} is dual feasible and is in $\text{Rel-Int}(P_{H_D})$. Let $Y = \text{diag}(y_1, \dots, y_m)$. Then one has

$$\begin{aligned} \phi_{H_D}(\mathbf{y}) &= -\pi_{A^\perp}(Y^{-1}\mathbf{e}) = -t\pi_{A^\perp}(\mathbf{u}(t)) \\ &= -t\pi_{A^\perp}(A\mathbf{x}(t) - \mathbf{b}) = t\pi_{A^\perp}(\mathbf{b}). \end{aligned}$$

Thus $\mathbf{y} = \mathbf{y}(t) \in T_+^{(D)}$. \square

As $t \rightarrow \infty$ the point $\mathbf{x}(t) \in T_+^{(P)}$ approaches an optimal solution \mathbf{x}_∞ of (P) , and $\mathbf{y}(t) \in T_+^{(D)}$ approaches an optimal solution \mathbf{y}_∞ of (D) . To see this one need only observe that (10.2) implies that $(\mathbf{y}_\infty, \mathbf{x}_\infty)$ are both feasible and satisfy the complementary slackness conditions:

$$(10.3) \quad y_j u_j = y_j (\langle \mathbf{a}_j, \mathbf{x} \rangle - b_j) = 0, \quad 1 \leq j \leq m.$$

Theorem 10.1 actually gives a rational mapping which when restricted to the domain $T_+^{(P)}$ maps $T_+^{(P)}$ to $T_+^{(D)}$, since for $\mathbf{x} = \mathbf{x}(t)$, one has

$$t = (1/c_i)\phi_{H_P}(\mathbf{x})_i,$$

where $\mathbf{c} = (c_1, \dots, c_n)$ and $c_i \neq 0$. Then using (10.1) one obtains such a rational mapping $\delta^{(P)}$ given by

$$\delta^{(P)}(\mathbf{x}) = \frac{c_i}{\phi_H(\mathbf{x})_i} \left[\frac{1}{\frac{\langle \mathbf{a}_1, \mathbf{x} \rangle - b_1}{\ddots} \frac{1}{\langle \mathbf{a}_m, \mathbf{x} \rangle - b_m}} \right].$$

One can also find a rational mapping $\delta^{(D)}(\mathbf{y})$ which when restricted to $T_+^{(D)}$ maps it to $T_+^{(P)}$. One defines t by $1/t = \phi_{H_D}(\mathbf{y})_j / \pi_{(A^T)^\perp}(\mathbf{b})_j$ where $\pi_{(A^T)^\perp}(\mathbf{b})_j \neq 0$; next one defines $\mathbf{u}(t)$ using (10.1), and then \mathbf{x} is recovered from \mathbf{u} by solving a suitable subset of the linear equations $A^T \mathbf{x} - \mathbf{b} = \mathbf{u}$. One has

$$\delta^{(D)} \circ \delta^{(P)}(\mathbf{x}) = \mathbf{x}, \quad \mathbf{x} \in T_+^{(P)}.$$

The combined primal-dual linear program (PD) on \mathbf{R}^{m+n} is

$$(PD) \quad \begin{cases} \text{minimize } \langle \mathbf{c}, \mathbf{x} \rangle - \langle \mathbf{b}, \mathbf{y} \rangle, \\ A^T \mathbf{x} \geq \mathbf{b}, \\ A\mathbf{y} = \mathbf{c}, \\ \mathbf{y} \geq \mathbf{0}. \end{cases}$$

Let H_{PD} denote the set of constraints of (PD). It is easy to check that the Legendre transform coordinate mapping for (PD) is

$$\phi_{H_{PD}} \left(\begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} \right) = \begin{bmatrix} \phi_{H_P}(\mathbf{x}) \\ \phi_{H_D}(\mathbf{y}) \end{bmatrix},$$

and that the central A -trajectory T_{PD} for (PD) is

$$T_{PD} = \left\{ \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} : \phi_{H_{PD}} \left(\begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix} \right) = t \begin{bmatrix} -\mathbf{c} \\ \pi_{A^\perp}(\mathbf{b}) \end{bmatrix} \right\}.$$

Hence the central A -trajectory of (PD) projects onto the central A -trajectories of (P) and (D), via orthogonal projections onto the \mathbf{x} -variables and \mathbf{y} -variables, respectively. The algebraic correspondence given by $\delta^{(P)}(\mathbf{x})$ associates the \mathbf{x} coordinate of a point $(\mathbf{x}, \mathbf{y})^T$ on the central A -trajectory of (PD) to the \mathbf{y} -coordinate.

By Theorem 10.1 points on the central trajectory $(\mathbf{x}(\mu), \mathbf{y}(\mu))$ of the combined primal-dual linear program (PD) satisfy the system of equations

$$\begin{aligned} A^T \mathbf{x} - \mathbf{u} &= \mathbf{b}, \\ A\mathbf{y} &= \mathbf{c}, \\ y_j u_j &= \mu, \quad 1 \leq j \leq m, \end{aligned}$$

with $0 < \mu < \infty$, where all $u_j > 0$. The path $(\mathbf{x}(\mu), \mathbf{y}(\mu))$ is a special case of the parametric logarithmic barrier function trajectories studied by Megiddo [M]. It can be viewed as a homotopy path in the parameter μ , and Kojima,

Mizumo and Yoshise [KMY] develop a polynomial-time linear programming algorithm that follows this path, starting with $\mu = 1$ and making $\mu \rightarrow 0$.

11. A -TRAJECTORIES AS ALGEBRAIC CURVES

A -trajectories are pairs of real algebraic curves. In this section we give polynomial ideals of relations that cut out these curves.

For any ideal I in the polynomial ring $\mathbb{C}[x_1, \dots, x_n]$ let

$$V(I) = \{(x_1, \dots, x_n) \in \mathbb{C}^n : f(x_1, \dots, x_n) = 0 \text{ for all } f \in I\}$$

denote the algebraic variety cut out by the ideal I .

We can explicitly write down an ideal of relations satisfied by points on an A -trajectory. Consider a linear program in \mathbb{R}^n in inequality form:

$$(P) \quad \begin{cases} \text{minimize } \langle \mathbf{c}, \mathbf{x} \rangle, \\ \langle \mathbf{a}_j, \mathbf{x} \rangle \geq b_j, \quad 1 \leq j \leq m, \end{cases}$$

having a full-dimensional polytope of feasible solutions. By Theorem 8.1 an arbitrary A -trajectory T_A is defined by

$$T_A = \{\mathbf{x} : \phi_H(\mathbf{x}) = t\mathbf{c} + \mathbf{c}_0, t \in \mathbb{R}\}.$$

The complex line $L(\mathbf{c}, \mathbf{c}_0) = \{t\mathbf{c} + \mathbf{c}_0 : t \in \mathbb{C}\}$ in \mathbb{C}^n is the variety associated to the polynomial ideal $I(\mathbf{c}, \mathbf{c}_0)$ generated by $\{\langle \mathbf{c}_i, \mathbf{y} \rangle - c_{0i} : 1 \leq i \leq n-1\}$ in $\mathbb{C}[\mathbf{y}_1, \dots, \mathbf{y}_n]$, where $[\mathbf{c}_1, \dots, \mathbf{c}_{n-1}]$ is any basis of \mathbf{c}^\perp and $c_{0i} = \langle \mathbf{c}_i, \mathbf{c}_0 \rangle$ for $1 \leq i \leq n-1$. Consider the ideal

$$(11.1) \quad I(T_A) = [p_1(\mathbf{x}), \dots, p_{n-1}(\mathbf{x})]$$

in $\mathbb{C}[x_1, \dots, x_n]$ generated by the polynomials

$$p_i(\mathbf{x}) = (\langle \mathbf{c}_i, \phi_H(\mathbf{x}) \rangle - c_{0i}) \prod' (\langle \mathbf{a}_j, \mathbf{x} \rangle - b_j),$$

for $1 \leq i \leq n$, where the prime indicates that the product is over one copy of each of the distinct linear factors among the $\langle \mathbf{a}_j, \mathbf{x} \rangle - b_j$; this product serves to clear denominators in $\phi_H(\mathbf{x})$. The ideal $I(T_A)$ cuts out an algebraic variety $V(I(T_A))$ which contains the A -trajectory T_A . In fact it seems likely that if H is full rank and P_H is full dimensional, and the linear program (P) is nondegenerate, then $V(I(T_A))$ is an irreducible curve in affine space \mathbb{C}^n .

There is a particularly simple ideal of relations satisfied by the central A -trajectory of the linear program (P) , which is obtained from the combined primal-dual linear program (PD) .

Theorem 11.1. *Given a full rank linear program (P) in \mathbb{R}^n having a full-dimensional polytope $P_{(P)}$ of feasible solutions and having $\mathbf{c} \neq \mathbf{0}$. For any point \mathbf{x} on the central A -trajectory of (P) except the center there is a unique point $(\mathbf{x}, \mathbf{y}, \mathbf{u}, \mu)$ satisfying the relations:*

$$(11.2a) \quad A^T \mathbf{x} - \mathbf{u} = \mathbf{b},$$

$$(11.2b) \quad A\mathbf{y} = \mathbf{c},$$

$$(11.2c) \quad y_j u_j = \mu, \quad 1 \leq j \leq m.$$

Conversely if \mathbf{x} is an interior point of $P_{(P)}$ for which there is a (complex-valued) solution $(\mathbf{x}, \mathbf{y}, \mathbf{u}, \mu)$ of these equations, then \mathbf{x} is on the central A -trajectory.

Proof. Let $V(A, \mathbf{b}, \mathbf{c})$ denote the algebraic variety determined by the relations (11.2). Whenever \mathbf{x} is a point on the positive part of the central trajectory of (P) , then there is a real point $(\mathbf{x}, \mathbf{y}, \mathbf{u}, \mu)$ on $V(A, \mathbf{b}, \mathbf{c})$ by Theorem 10.1, having $\mathbf{y} > 0$ and $\mu > 0$. The proof of Theorem 10.1 also shows that if \mathbf{x} is a point on the negative part of the central trajectory of (P) , then there is a real point $(\mathbf{x}, \mathbf{y}, \mathbf{u}, \mu)$ on $V(A, \mathbf{b}, \mathbf{c})$ having $\mathbf{y} < 0$ and $\mu < 0$.

Conversely if $(\mathbf{x}, \mathbf{y}, \mathbf{u}, \mu)$ is a point on $V(A, \mathbf{b}, \mathbf{c})$ and $\mathbf{x} \in \text{Int}(P_{(P)})$ then \mathbf{x} must be on the central trajectory. To see this, one calculates that the equations (11.2) imply that

$$\phi_H(\mathbf{x}) = (1/\mu)\mathbf{c},$$

if $\mu \neq 0$, which shows that μ is real and \mathbf{x} is on the central trajectory. In the remaining case $\mu = 0$ the equations (11.2) are just the complementary slackness conditions for (P) , and can have no solution with $\mathbf{x} \in \text{Int}(P_{(P)})$ if the objective function $\langle \mathbf{c}, \mathbf{x} \rangle$ is nonconstant.

The point $(\mathbf{x}, \mathbf{y}, \mathbf{u}, \mu)$ on $V(A, \mathbf{b}, \mathbf{c})$ with \mathbf{x} lying on the central trajectory is unique, because by (11.2a) the value \mathbf{x} determines \mathbf{u} , and the system (11.2b), (11.2c) is then a full rank system of linear equations in unknowns (\mathbf{y}, μ) , thus determining them uniquely. \square

Theorem 11.1 implies that there is an irreducible algebraic curve $V(T_A)$ in $\mathbf{P}^{2n+m+1}(\mathbf{C})$ whose projection onto \mathbf{x} -space intersected with $\{\mathbf{x}: \mathbf{x} \in \mathbf{R}^n\}$ is a finite set of real curves that includes the central A -trajectory, and $V(T_A)$ is a component in the variety $V(A, \mathbf{b}, \mathbf{c})$. The variety $V(A, \mathbf{b}, \mathbf{c})$ is defined by $n + 2m$ equations in $n + 2m + 1$ variables, so that all components of $V(A, \mathbf{b}, \mathbf{c})$ have dimension at least one, and $V(T_A)$ has dimension exactly one. Under the hypotheses of Theorem 11.1 all components of the variety $V(A, \mathbf{b}, \mathbf{c})$ that intersect the affine space \mathbf{C}^{n+2m+1} have dimension one, but $V(A, \mathbf{b}, \mathbf{c})$ contains higher-dimensional components in the hyperplane at infinity in projective space $\mathbf{P}^{n+2m+1}(\mathbf{C})$. To see this, consider the intersection of $V(A, \mathbf{b}, \mathbf{c})$ with the hyperplane $\mu = 0$. Generically in the affine space \mathbf{C}^{n+2m+1} this intersection consists of exactly $\binom{n}{m}$ points, all real. This holds because the conditions

$$y_j u_j = 0, \quad 1 \leq j \leq m,$$

force at least one of each pair (y_j, u_j) to be zero. Generically no more than n of the u_j can be zero, because if $(n + 1)$ u -variables are zero then the system (11.2a) is generically inconsistent, since it then consists of m linear equations in $m - 1$ unknowns. By similar reasoning no more than $m - n$ of the y -variables must be zero. Hence exactly m of the u -variables are zero and $n - m$ of the y -variables are zero, for which there are $\binom{m}{n}$ choices. Once these are selected, in the generic case all the other variables are uniquely determined by the linear equations (11.2a), (11.2b); thus there are $\binom{m}{n}$ points in the intersection. Now

consider the hyperplane at infinity $z = 0$ in $\mathbf{P}^{n+2m+1}(\mathbf{C})$ intersected with the homogeneous ideal

$$\begin{aligned} A^T \mathbf{x} - \mathbf{u} &= bz, \\ A\mathbf{y} &= cz, \\ y_j u_j &= \mu z, \quad 1 \leq j \leq m. \end{aligned}$$

It is easy to see that for full rank A that $V(A, \mathbf{b}, \mathbf{c})$ contains an $(m - n)$ -dimensional projective space determined by $z = 0$, $\mathbf{y} = \mathbf{0}$, $A^T \mathbf{x} - \mathbf{u} = \mathbf{0}$; it contains many other components of dimension ≥ 2 as well.

It remains an open problem to give generators for an ideal cutting out exactly the irreducible curve $V(T_A)$.

12. LAGRANGIAN DYNAMICAL SYSTEMS PRODUCING A -TRAJECTORIES

The Legendre transform plays an important role in classical mechanics; it converts a dynamical system given in Lagrangian form to one in Hamiltonian form, see [Ar, Ln]. This conversion replaces a system of n second-order differential equations in n variables (Lagrange's equations) with a system of $2n$ first-order differential equations in $2n$ variables (Hamilton's equations). The simple form of A -trajectories and P -trajectories in Legendre transform coordinate space suggests that they might be described by a simple dynamical system in Lagrangian or Hamiltonian form.

This is the case for A -trajectories. Consider linear programs in \mathbf{R}^n in the inequality form:

$$(12.1) \quad \begin{cases} \text{minimize } \langle \mathbf{c}, \mathbf{x} \rangle, \\ \langle \mathbf{a}_j, \mathbf{x} \rangle \geq b_j, \quad 1 \leq j \leq m. \end{cases}$$

Let H denote this set of constraints and let P_H denote its polytope of feasible solutions. We present two Lagrangian dynamical systems in which the evolution of the $\dot{\mathbf{q}}$ -variables follows A -trajectories. The first system has \mathbf{q} -variables in \mathbf{R}^n and yields all A -trajectories for a *fixed* objective function. The second has position variables $(\mathbf{q}, \dot{\mathbf{q}})$ in \mathbf{R}^{2n} and the evolution of the $\dot{\mathbf{q}}$ -variables yields A -trajectories for *all* objective functions. The associated Hamiltonian dynamical systems are globally completely integrable. For simplicity we treat in detail the case that P_H is bounded and has a nonempty interior. At the end of the section we describe what happens in the cases that P_H is unbounded with nonempty interior and when P_H is lower dimensional.

A *Lagrangian dynamical system* in a space \mathbf{R}^n has position variables $\mathbf{q} = (q_1, \dots, q_n)$ and velocity variables $\dot{\mathbf{q}} = (\dot{q}_1, \dots, \dot{q}_n)$. The motion of the system is specified by a Lagrangian function $L(\mathbf{q}, \dot{\mathbf{q}})$ mapping \mathbf{R}^{2n} to \mathbf{R} . *Lagrange's equations of motion* are

$$(12.2a) \quad \frac{d}{dt} \left(\frac{\partial L}{\partial \dot{q}_i} \right) = \frac{\partial L}{\partial q_i}, \quad 1 \leq i \leq n,$$

together with

$$(12.2b) \quad \dot{q}_i = \frac{d}{dt}(q_i), \quad 1 \leq i \leq n.$$

The first Lagrangian dynamical system we consider has position variables \mathbf{q} in \mathbf{R}^n and the Lagrangian

$$(12.3) \quad L_1(\mathbf{q}, \dot{\mathbf{q}}) = \langle \mathbf{c}, \mathbf{q} \rangle - \sum_{j=1}^m \log(\langle \mathbf{a}_j, \dot{\mathbf{q}} \rangle - b_j).$$

We show that the $\dot{\mathbf{q}}$ -trajectories of this Lagrangian system are the complete set of A -trajectories for the linear program (12.1), having a fixed objective function $\langle \mathbf{c}, \mathbf{x} \rangle$.

Theorem 12.1. *Suppose that the set H of linear program constraints*

$$(12.4) \quad \langle \mathbf{a}_j, \mathbf{x} \rangle \geq b_j, \quad 1 \leq j \leq m,$$

has a bounded polytope of feasible solutions in \mathbf{R}^n with nonempty interior. Then for $\mathbf{c} \neq \mathbf{0}$ in \mathbf{R}^n each trajectory $\{\dot{\mathbf{q}}(t): -\infty < t < \infty\}$ of the Lagrangian dynamical system with the Lagrangian

$$L_1(\mathbf{q}, \dot{\mathbf{q}}) = \langle \mathbf{c}, \mathbf{q} \rangle - \sum_{j=1}^m \log(\langle \mathbf{a}_j, \dot{\mathbf{q}} \rangle - b_j)$$

is an A -trajectory of the linear program to minimize the objective function $\langle \mathbf{c}, \mathbf{x} \rangle$ with these constraints. Conversely, every A -trajectory of this linear program is a $\dot{\mathbf{q}}$ -trajectory of this Lagrangian dynamical system.

Proof. Define $p_i = \partial L_1 / \partial \dot{q}_i$ so that the vector $\mathbf{p} = (p_1, \dots, p_n)$ satisfies

$$(12.5) \quad \mathbf{p} \equiv \partial L_1 / \partial \dot{\mathbf{q}}.$$

A direct calculation from (12.2) yields

$$(12.6) \quad \mathbf{p} = - \sum_{j=1}^m \frac{\mathbf{a}_j}{\langle \mathbf{a}_j, \dot{\mathbf{q}} \rangle - b_j} \equiv \phi_H(\dot{\mathbf{q}}),$$

where $\phi_H(\cdot)$ is the Legendre transform coordinate mapping associated to the constraints H of the linear program. Now Lagrange's equations for this Lagrangian are

$$(12.7) \quad \frac{d}{dt}(\phi_H(\dot{\mathbf{q}})) = \mathbf{c}.$$

Hence

$$(12.8) \quad \phi_H(\dot{\mathbf{q}}) = \mathbf{c}t + \mathbf{c}_0,$$

where \mathbf{c}_0 is a vector of initial conditions. Since the polytope P_H of feasible solutions is bounded with nonempty interior, by Theorem 3.5 the mapping $\phi_H: \text{Int}(P_H) \rightarrow \mathbf{R}^n$ is one-to-one and onto and has an analytic global inverse function $\phi_H^{-1}: \mathbf{R}^n \rightarrow \text{Int}(P_H)$. Hence we may invert (12.6) everywhere to get

$$\dot{\mathbf{q}} = \phi_H^{-1}(\mathbf{c}t + \mathbf{c}_0).$$

But this is exactly the definition of an A -trajectory for the objective function $\langle \mathbf{c}, \mathbf{x} \rangle$ for this linear program, by (8.12). Since \mathbf{c}_0 is arbitrary we get all A -trajectories for the fixed objective function $\langle \mathbf{c}, \mathbf{x} \rangle$ this way. \square

We convert this Lagrangian dynamical system to an equivalent Hamiltonian dynamical system. A *Hamiltonian dynamical system* is specified by a Hamiltonian function $H(\mathbf{p}, \mathbf{q})$ defined on (a subset of) \mathbf{R}^{2n} and has variables (\mathbf{p}, \mathbf{q}) with $\mathbf{p} = (p_1, \dots, p_n)^T$. \mathbf{R}^{2n} is usually called *phase space*. This system evolves according to *Hamilton's equations*, which are

$$(12.9) \quad \begin{cases} \dot{p}_i = -\partial H / \partial q_i, & 1 \leq i \leq n, \\ \dot{q}_i = \partial H / \partial p_i, & 1 \leq i \leq n. \end{cases}$$

We define the *Poisson bracket* $\{\psi_1, \psi_2\}$ of two continuously differentiable functions on a subset of \mathbf{R}^{2n} by

$$\{\psi_1, \psi_2\} = \sum_{i=1}^n \left(\frac{\partial \psi_1}{\partial p_i} \frac{\partial \psi_2}{\partial q_i} - \frac{\partial \psi_1}{\partial q_i} \frac{\partial \psi_2}{\partial p_i} \right).$$

A Hamiltonian dynamical system on \mathbf{R}^{2n} is said to be (globally) *completely integrable* if there is an open subset S of phase space \mathbf{R}^{2n} which is an invariant subset of H (i.e., it is a union of trajectories satisfying (12.9)) and there exist n functions F_1, F_2, \dots, F_n defined on S such that

- (a) $\{H, F_i\} = 0$ for $1 \leq i \leq n$,
- (b) $\{F_i, F_j\} = 0$ for $1 \leq i \leq n$,
- (c) the gradients $\{(dF_i)_x : 1 \leq i \leq n\}$ are linearly independent at all points of S .

(See [Mo] for examples.) Any function F satisfying $\{H, F\} = 0$ is called an *integral* of the motion, and F is a *conserved quantity* of the motion in the sense that

$$F_i(\mathbf{p}(t), \mathbf{q}(t)) = \text{constant}$$

holds on any solution $(\mathbf{p}(t), \mathbf{q}(t))$ of Hamilton's equations. Most Hamiltonian systems are *not* completely integrable when $n \geq 2$.

The Legendre transformation converts a Lagrangian dynamical system $L(\mathbf{q}, \dot{\mathbf{q}})$ to a Hamiltonian system $H(\mathbf{p}, \mathbf{q})$ which is equivalent in the sense that the functions $\mathbf{q}(t)$ evolve identically in the two systems. The Hamiltonian associated to a Lagrangian L is

$$(12.10) \quad H(\mathbf{p}, \mathbf{q}) = \sum_{i=1}^n p_i \dot{q}_i - L(\mathbf{q}, \dot{\mathbf{q}}).$$

In this formula one needs to express the \dot{q}_i in terms of the (\mathbf{p}, \mathbf{q}) variables. To do this one takes the defining relations

$$(12.11) \quad p_i \equiv \partial L / \partial \dot{q}_i, \quad 1 \leq i \leq n,$$

and solves this system for the \dot{q}_i to obtain

$$(12.12) \quad \dot{q}_i = f_i(\mathbf{p}, \mathbf{q}), \quad 1 \leq i \leq n.$$

The Hamiltonian is then expressed explicitly as

$$(12.13) \quad H(\mathbf{p}, \mathbf{q}) = \sum_{i=1}^n p_i f_i(\mathbf{p}, \mathbf{q}) - L(\mathbf{q}, \mathbf{f}(\mathbf{p}, \mathbf{q})).$$

This conversion method works in principle but usually not in practice since the inversion (12.11) cannot usually be specified in closed form. (In addition there may be singular points where the inversion cannot be done.)

In the case of the Lagrangian $L_1(\mathbf{q}, \dot{\mathbf{q}})$ in Theorem 12.1 we can carry out this inversion and prove the following result.

Theorem 12.2. *Suppose that the constraints H given by*

$$\langle \mathbf{a}_j, \mathbf{x} \rangle \geq b_j, \quad 1 \leq j \leq m,$$

have a bounded polytope of feasible solutions with nonempty interior. Then the Hamiltonian $H_1(\mathbf{p}, \mathbf{q})$ corresponding to the Lagrangian

$$L_1(\mathbf{q}, \dot{\mathbf{q}}) = \langle \mathbf{c}, \mathbf{q} \rangle - \sum_{j=1}^m \log(\langle \mathbf{a}_j, \dot{\mathbf{q}} \rangle - b_j)$$

is

$$(12.14) \quad H_1(\mathbf{p}, \mathbf{q}) = \langle \mathbf{p}, \phi_H^{-1}(\mathbf{p}) \rangle - \langle \mathbf{c}, \mathbf{q} \rangle - \sum_{j=1}^m \log(\langle \mathbf{a}_j, \phi_H^{-1}(\mathbf{p}) \rangle - b_j)$$

where $\phi_H^{-1}(\cdot)$ is the inverse Legendre transform coordinate map for the constraints H . This Hamiltonian system is completely integrable on the whole phase space \mathbf{R}^{2n} .

Proof. The hypotheses of Theorem 12.1 apply. There we observed that $\mathbf{p} = \partial L / \partial \dot{\mathbf{q}} = \phi_H(\dot{\mathbf{q}})$, so that $\dot{\mathbf{q}} = \phi_H^{-1}(\mathbf{p})$. Substituting the formulae $\dot{q}_i = \phi_H^{-1}(\mathbf{p})_i$ into (12.13) yields formula (12.14).

It remains to prove complete integrability. Take $\{\mathbf{c}^{(1)}, \dots, \mathbf{c}^{(n-1)}\}$ to be a basis of \mathbf{c}^\perp , and set

$$\begin{aligned} F_k(\mathbf{p}, \mathbf{q}) &= \langle \mathbf{c}^{(k)}, \mathbf{p} \rangle, \quad 1 \leq k \leq n-1, \\ F_n(\mathbf{p}, \mathbf{q}) &= H_1(\mathbf{p}, \mathbf{q}). \end{aligned}$$

We claim these functions satisfy conditions (a)–(c) for complete integrability. We need only verify (b) and (c) since $\{H_1, F_i\} = \{F_n, F_i\}$. It is clear that

$$\{F_{k_1}, F_{k_2}\} = 0, \quad 1 \leq k_1, k_2 \leq n-1$$

since $\partial F_k / \partial q_i = 0$ for all i in this case. Now

$$\begin{aligned} \{H_1, F_k\} &= \sum_{i=1}^n \left(\frac{\partial H_1}{\partial p_i} \frac{\partial F_k}{\partial q_i} - \frac{\partial H_1}{\partial q_i} \frac{\partial F_k}{\partial p_i} \right) \\ &= \sum_{i=1}^n c_i c_i^{(k)} = \langle \mathbf{c}^{(k)}, \mathbf{c} \rangle = 0 \end{aligned}$$

for $1 \leq k \leq n-1$, which proves (b). To verify (c) we show that $\{dF_k : 1 \leq k \leq n\}$ are linearly independent at all points of \mathbf{R}^{2n} . Now

$$dF_k = \left[\frac{\partial F_k}{\partial p_1}, \dots, \frac{\partial F_k}{\partial p_n}, \frac{\partial F_k}{\partial q_1}, \dots, \frac{\partial F_k}{\partial q_n} \right]$$

and we obtain

$$(12.15) \quad \begin{bmatrix} dF_1 \\ \vdots \\ dF_{n-1} \\ dF_n \end{bmatrix} = \begin{bmatrix} \mathbf{c}_1^T & \mathbf{0} \\ \vdots & \\ \mathbf{c}_{n-1}^T & \mathbf{0} \\ * & \mathbf{c}^T \end{bmatrix}$$

where the entries in $*$ are functions of (\mathbf{p}, \mathbf{q}) . The linear independence of the rows of (12.15) at all points of \mathbf{R}^{2n} is clear. This proves (c). \square

The last two theorems imply that the $\dot{\mathbf{q}}$ -trajectories of the Hamiltonian $H_1(\mathbf{p}, \mathbf{q})$ are exactly the A -trajectories of the linear program (12.5). This can also be verified directly from Hamilton's equations for $H_1(\mathbf{p}, \mathbf{q})$. Hamilton's equations are

$$(12.16a) \quad \dot{\mathbf{p}} = \mathbf{c},$$

$$\begin{aligned} \dot{\mathbf{q}} &= \frac{\partial H_1}{\partial \mathbf{p}} \\ &= \phi_H^{-1}(\mathbf{p}) + \frac{\partial}{\partial \mathbf{p}}[\phi_H^{-1}(\mathbf{p})] \cdot \mathbf{p} - \frac{\partial}{\partial \mathbf{p}}[\phi_H^{-1}(\mathbf{p})](\phi_H(\phi_H^{-1}(\mathbf{p}))) \end{aligned}$$

$$(12.16b) \quad = \phi_H^{-1}(\mathbf{p}).$$

These yield

$$\mathbf{p}(t) = \mathbf{c}t + \mathbf{c}_0, \quad \dot{\mathbf{q}}(t) = \phi_H^{-1}(\mathbf{c}t + \mathbf{c}_0),$$

so that $\dot{\mathbf{q}}$ -trajectories are A -trajectories for the objective function $\langle \mathbf{c}, \mathbf{x} \rangle$.

The second Lagrangian dynamical system has position variables $(\mathbf{q}, \tilde{\mathbf{q}})$ in \mathbf{R}^{2n} and the Lagrangian

$$(12.17) \quad L_2(\mathbf{q}, \tilde{\mathbf{q}}, \dot{\mathbf{q}}, \dot{\tilde{\mathbf{q}}}) = \frac{1}{2} \langle \mathbf{q} + \dot{\tilde{\mathbf{q}}}, \mathbf{q} + \dot{\tilde{\mathbf{q}}} \rangle - \sum_{j=1}^n \log(\langle \mathbf{a}_j, \dot{\tilde{\mathbf{q}}} \rangle - b_j).$$

The $\dot{\mathbf{q}}$ -trajectories of this dynamical system give all of the A -trajectories.

Theorem 12.3. Suppose that the set H of linear program constraints

$$\langle \mathbf{a}_j, \mathbf{x} \rangle \geq b_j, \quad 1 \leq j \leq m,$$

has a bounded polytope of feasible solutions in \mathbf{R}^n with nonempty interior. Then each trajectory $\{\dot{\mathbf{q}}(t): -\infty < t < \infty\}$ of the Lagrangian dynamical system with Lagrangian

$$L_2 = \frac{1}{2} \langle \mathbf{q} + \dot{\mathbf{q}}, \mathbf{q} + \dot{\mathbf{q}} \rangle - \sum_{j=1}^n \log(\langle \mathbf{a}_j, \dot{\mathbf{q}} \rangle - b_j)$$

is an A -trajectory of a linear program with constraints H and some objective function $\langle \mathbf{c}, \mathbf{x} \rangle$, and conversely all A -trajectories for such linear programs with all possible objective functions $\langle \mathbf{c}, \mathbf{x} \rangle$ arise as $\dot{\mathbf{q}}$ -trajectories of this Lagrangian dynamical system.

Remark. In the case $\mathbf{c} = \mathbf{0}$ these trajectories are constants. We make the theorem hold in this case by defining A -trajectories to be points when $\mathbf{c} = \mathbf{0}$.

Proof. This is an easy computation. We see that

$$\begin{aligned} \frac{\partial L_2}{\partial \mathbf{q}} &= \mathbf{q} + \dot{\mathbf{q}}, & \frac{\partial L_2}{\partial \dot{\mathbf{q}}} &= \mathbf{0}, \\ \frac{\partial L_2}{\partial \dot{\mathbf{q}}} &= \phi_H(\dot{\mathbf{q}}), & \frac{\partial L_2}{\partial \ddot{\mathbf{q}}} &= \mathbf{q} + \dot{\mathbf{q}}. \end{aligned}$$

Hence Lagrange's equations of motion are

$$\frac{d}{dt}(\phi_H(\dot{\mathbf{q}})) = \mathbf{q} + \dot{\mathbf{q}}, \quad \frac{d}{dt}(\mathbf{q} + \dot{\mathbf{q}}) = \mathbf{0}.$$

Consequently one has $\mathbf{q} + \dot{\mathbf{q}} = \mathbf{c}$ where the \mathbf{c} are arbitrary constants of integration, and $\phi_H(\dot{\mathbf{q}}) = \mathbf{c}t + \mathbf{c}_0$. Thus $\dot{\mathbf{q}}(t) = \phi_H^{-1}(\mathbf{c}t + \mathbf{c}_0)$ runs over the full set of A -trajectories for all linear programs with all possible objective functions $\langle \mathbf{c}, \mathbf{x} \rangle$ and fixed constraints H , using (8.12). \square

We convert this Lagrangian system to its equivalent Hamiltonian dynamical system and show that it is completely integrable.

Theorem 12.4. Suppose that the constraints H given by

$$\langle \mathbf{a}_j, \mathbf{x} \rangle \geq b_j, \quad 1 \leq j \leq m,$$

have a bounded polytope of feasible solutions with nonempty interior. Then the Hamiltonian $H_2(\mathbf{p}, \tilde{\mathbf{p}}, \mathbf{q}, \tilde{\mathbf{q}})$ corresponding to the Lagrangian

$$L_1 = \frac{1}{2} \langle \mathbf{q} + \dot{\mathbf{q}}, \mathbf{q} + \dot{\mathbf{q}} \rangle - \sum_{j=1}^m \log(\langle \mathbf{a}_j, \dot{\mathbf{q}} \rangle - b_j)$$

is

$$(12.18) \quad H_2 = \frac{1}{2} \langle \tilde{\mathbf{p}}, \tilde{\mathbf{p}} - 2\mathbf{q} \rangle + \langle \mathbf{p}, \phi_H^{-1}(\mathbf{p}) \rangle + \sum_{j=1}^m \log(\langle \mathbf{a}_j, \phi_H^{-1}(\mathbf{p}) \rangle - b_j)$$

where $\phi_H^{-1}(\cdot)$ is the inverse Legendre transform coordinate mapping for H . This Hamiltonian dynamical system is completely integrable on the invariant open set $\mathbf{R}^{4n} - \{(\mathbf{p}, \tilde{\mathbf{p}}, \mathbf{q}, \tilde{\mathbf{q}}): \tilde{p}_1 = 0\}$.

Proof. One has

$$\mathbf{p} = \partial L / \partial \dot{\mathbf{q}} = \phi_{\mathbf{H}}(\dot{\mathbf{q}}),$$

$$\tilde{\mathbf{p}} = \partial L / \partial \dot{\tilde{\mathbf{q}}} = \mathbf{q} + \dot{\tilde{\mathbf{q}}},$$

since the polytope $P_{\mathbf{H}}$ is bounded and nonempty. Theorem 3.1 says that

$$(12.19a) \quad \dot{\mathbf{q}} = \phi_{\mathbf{H}}^{-1}(\mathbf{p})$$

is valid for all \mathbf{p} and the equation for $\tilde{\mathbf{p}}$ yields

$$(12.19b) \quad \dot{\tilde{\mathbf{q}}} = \tilde{\mathbf{p}} - \mathbf{q}.$$

By definition

$$H_2 = \langle \mathbf{p}, \dot{\mathbf{q}} \rangle + \langle \tilde{\mathbf{p}}, \dot{\tilde{\mathbf{q}}} \rangle - \frac{1}{2} \langle \mathbf{q} + \dot{\tilde{\mathbf{q}}}, \mathbf{q} + \dot{\tilde{\mathbf{q}}} \rangle + \sum_{j=1}^m \log(\langle \mathbf{a}_j, \dot{\mathbf{q}} \rangle - b_j).$$

Substituting (12.19) into this equation and simplifying yields (12.18).

To demonstrate complete integrability, one takes the set of functions $F_k = F_k(\mathbf{p}, \tilde{\mathbf{p}}, \mathbf{q}, \tilde{\mathbf{q}})$ for $1 \leq k \leq 2n$ defined by

$$\begin{aligned} F_k &= \tilde{p}_k, & 1 \leq k \leq n, \\ F_{n+k} &= p_1 \tilde{p}_{k+1} - p_{k+1} \tilde{p}_1, & 1 \leq k \leq n-1, \\ F_{2n} &= H_2. \end{aligned}$$

It is easy to see that $\{F_{k_1}, F_{k_2}\} = 0$ for $1 \leq k_1, k_2 \leq 2n-1$ since these integrals involve no $(\mathbf{q}, \tilde{\mathbf{q}})$ variables. Also $\{F_k, H_2\} = 0$ for $1 \leq k \leq n$ because H_2 contains no $\tilde{\mathbf{q}}_k$ variables. Finally a computation gives

$$\{F_{n+k}, H_2\} = \frac{\partial F_{n+k}}{\partial p_1} \frac{\partial H_2}{\partial q_1} + \frac{\partial F_{n+k}}{\partial p_{k+1}} \frac{\partial H_2}{\partial q_{k+1}} = 0.$$

Since $\{H_2, H_2\} = 0$ this verifies (a) and (b). To prove (c) we evaluate the matrix of differentials

$$dF = \left(\frac{\partial F}{\partial \tilde{p}_1}, \dots, \frac{\partial F}{\partial \tilde{p}_n}, \frac{\partial F}{\partial p_1}, \dots, \frac{\partial F}{\partial p_n}, \frac{\partial F}{\partial \tilde{q}_1}, \dots, \frac{\partial F}{\partial \tilde{q}_n}, \frac{\partial F}{\partial q_1}, \dots, \frac{\partial F}{\partial q_n} \right)$$

to be

$$\begin{bmatrix} dF_1 \\ \vdots \\ \frac{dF_{2n-1}}{dF_{2n}} \\ \vdots \\ \frac{dF_n}{dF_{n+1}} \end{bmatrix} = \begin{bmatrix} \mathbf{I} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \hline -p_2 & p_1 \cdots 0 & \tilde{p}_2 - \tilde{p}_1 \cdots 0 & \\ -p_n & 0 \cdots p_1 & \tilde{p}_n & 0 \cdots -\tilde{p}_1 & \mathbf{0} & \mathbf{0} \\ \hline \tilde{p}_1 - q_1 \cdots \tilde{p}_n - q_n & \phi_{\mathbf{H}}^{-1}(\mathbf{p}) & \mathbf{0} & -\tilde{p}_1 \cdots -\tilde{p}_n \end{bmatrix}$$

where \mathbf{I} is an $n \times n$ identity matrix and $\mathbf{0}$ denotes a zero matrix (of the appropriate size). It is clear that this matrix has full row rank whenever $\tilde{p}_1 \neq 0$.

Finally we note that the open set $\mathbf{R}^{4n} - \{(\mathbf{p}, \tilde{\mathbf{p}}, \mathbf{q}, \tilde{\mathbf{q}}) : \tilde{p}_1 = 0\}$ is an invariant set for H_2 . Indeed, since $F_1 = \tilde{p}_1$ is an integral of the motion, each level set $\tilde{p}_1 = \text{constant}$ is separately an invariant set. \square

What happens when the linear program does not have a bounded, full-dimensional polytope of feasible solutions? In the case that the feasible solution polytope P_H is an unbounded full-dimensional polytope, the Lagrangians L_1 and L_2 are defined identically as in the bounded case, and the trajectories corresponding to A -trajectories occupy only part of the phase space. The image of P_H under the Legendre transform coordinates $\phi_H(P_H)$ occupies an open full-dimensional pointed cone in the \mathbf{p} -coordinates (by Theorem 3.5) and trajectories having \mathbf{p} -coordinate inside this cone correspond to A -trajectories. These trajectories are not defined for all time but only for a time interval (t_0, t_1) , depending on the objective function, with some \mathbf{p} -variable diverging at finite time. These Hamiltonian systems are still completely integrable in an appropriate open invariant subset of the phase space. The remaining case where P_H is a lower-dimensional polytope can in principle be reduced to one of the full-dimensional polytope cases by eliminating variables, i.e., restricting the problem to the lower-dimensional flat in \mathbf{R}^n spanned by P_H .

There are several open questions suggested by these dynamical systems. First, are there analogous completely integrable Hamiltonian dynamical systems describing P -trajectories? Second, these systems identify A -trajectories with \mathbf{q} -trajectories, which correspond to *velocities* in the physical interpretation. What connection (if any) to the linear programming problem do the *position* trajectories (\mathbf{q} -trajectories) have? Third, for a linear program with a bounded polytope P_H in which both $\langle \mathbf{c}, \mathbf{x} \rangle$ and $-\langle \mathbf{c}, \mathbf{x} \rangle$ have unique optimal solutions, all the \mathbf{q} -trajectories of the Lagrangian $L_1(\mathbf{q}, \dot{\mathbf{q}})$ have unique limiting velocities $\dot{\mathbf{q}}(\infty)$ as $t \rightarrow +\infty$ and $\dot{\mathbf{q}}(-\infty)$ as $t \rightarrow -\infty$ (which correspond to optimal solutions of these two linear programs). Hence they exhibit *forward* and *backward scattering*. Is there a scattering theory interpretation of what this dynamical system is doing in the \mathbf{q} -variable space?

APPENDIX A. ABSTRACT LEGENDRE TRANSFORM COORDINATES

The Legendre transform coordinate mapping $\phi_H(\mathbf{x})$ has a coordinate-free version, which we call the *abstract Legendre transform coordinate mapping* and denote $\hat{\phi}_H(\mathbf{x})$. Consider a set H of inequality constraints

$$\langle \mathbf{a}_j, \mathbf{x} \rangle \geq b_j, \quad 1 \leq j \leq m,$$

having the polytope of feasible solutions P_H and feasible direction subspace D_H . Let H_n denote the set of nonsingular constraints, i.e., these constraints that do not hold with equality in P_H , which by renumbering constraints if

necessary may be taken to be

$$\langle \mathbf{a}_j, \mathbf{x} \rangle \geq b_j, \quad 1 \leq j \leq m^*.$$

The logarithmic barrier function

$$f_{H_n}(\mathbf{x}) = - \sum_{j=1}^m \log(\langle \mathbf{a}_j, \mathbf{x} \rangle - b_j)$$

has differential $df_{H_n}: \text{Int}(P_{H_n}) \rightarrow (\mathbf{R}^n)^*$ defined by

$$f_{H_n}(\mathbf{x} + \varepsilon \mathbf{v}) = f_{H_n}(\mathbf{x}) + \varepsilon \langle (df_{H_n})_{\mathbf{x}}, \mathbf{v} \rangle + O(\varepsilon^2),$$

as $\varepsilon \rightarrow 0$. Let D_H^p be the subspace polar to D_H in $(\mathbf{R}^n)^*$, defined by

$$D_H^p = \{y \in (\mathbf{R}^n)^*: \langle y, \mathbf{x} \rangle = 0 \text{ for all } \mathbf{x} \in D_H\},$$

and let $\bar{\pi}_{D_H^p}: (\mathbf{R}^n)^* \rightarrow (\mathbf{R}^n)^*/D_H^p$ be the quotient mapping. The *abstract Legendre transform coordinate mapping* $\hat{\phi}_H(\mathbf{x})$ is defined by

$$\hat{\phi}_H(\mathbf{x}) = \bar{\pi}_{D_H^p}((df_{H_n})_{\mathbf{x}}).$$

The coordinate-free analogues of the results of §3 are as follows. Here we regard the constraint coefficients \mathbf{a}_j as elements of $(\mathbf{R}^n)^*$.

Theorem A.1. *The abstract Legendre transform coordinate mapping $\hat{\phi}_H(\mathbf{x})$ has domain $\text{Rel-Int}(P_H)$ and range $\text{Rel-Int}(\bar{C}_H)$ where $\bar{C}_H = \bar{\pi}_{D_H^p}(\bar{C}_{H_n})$ and $\bar{C}_{H_n} = \mathbf{R}^+[-\mathbf{a}_1, \dots, -\mathbf{a}_{m^*}]$ in $(\mathbf{R}^n)^*$. If H is of full rank then $\dim(\bar{C}_H) = \dim(P_H)$ and $\hat{\phi}_H$ is a one-to-one mapping.*

Let $\mathbf{J}(\mathbf{x}) = L\mathbf{x} + \mathbf{m}$ be an affine mapping for \mathbf{R}^n to \mathbf{R}^k . Define its *adjoint mapping* \mathbf{J}^* to be the linear mapping $(\mathbf{R}^k)^* \rightarrow (\mathbf{R}^n)^*$ given by

$$\langle \mathbf{J}^*(\mathbf{y}), \mathbf{v} \rangle = \langle \mathbf{y}, \mathbf{J}(\mathbf{v}) - \mathbf{J}(\mathbf{0}) \rangle.$$

Let H be a set of constraints with feasible solution set P_H having M_H as its affine hull. Define $\hat{\phi}_{J(H)}$ to be $\hat{\phi}_H$ where \tilde{H} is any set of constraints that agrees with the images of the constraints of H on the flat $\mathbf{J}(M_H)$. It can be checked that $\hat{\phi}_{J(H)}$ is well defined independent of the choice of \tilde{H} .

Theorem A.2. *Let $\mathbf{J}(\mathbf{x}) = L\mathbf{x} + \mathbf{m}$ be any affine mapping from \mathbf{R}^n to \mathbf{R}^k which is one-to-one on M_H , the affine hull of P_H . Then $D_{J(H)}^p \subseteq \ker(\bar{\pi}_{D_H^p} \circ \mathbf{J}^*)$ so that the induced map $\bar{\pi}_{D_H^p} \circ \mathbf{J}^*: (\mathbf{R}^k)^*/D_{J(H)}^p \rightarrow (\mathbf{R}^n)^*/D_H^p$ is well defined, and the following diagram commutes:*

$$\begin{array}{ccc} \text{Rel-Int}(P_H) & \xrightarrow{\mathbf{J}} & \text{Rel-Int}(P_{J(H)}) \\ \hat{\phi}_H \downarrow & & \downarrow \hat{\phi}_{J(H)} \\ \text{Rel-Int}(\bar{C}_H) & \xleftarrow{\bar{\pi}_{D_H^p} \circ \mathbf{J}^*} & \text{Rel-Int}(\bar{C}_{J(H)}). \end{array}$$

REFERENCES

- [AKRV] I. Adler, N. Karmarkar, M. Resende, and G. Veiga, *An implementation of Karmarkar's algorithm for linear programming*, preprint, Univ. of California, Berkeley, May 1986.
- [A] K. Anstreicher, *A monotonic projective algorithm for fractional linear programming*, *Algorithmica* **1** (1986), 483–498.
- [Ar] V. I. Arnold, *Mathematical methods of classical mechanics*, Springer-Verlag, New York, 1978.
- [B] E. R. Barnes, *A variation on Karmarkar's algorithm for solving linear programming problems*, *Math. Programming* **36** (1986), 174–182.
- [BL1] D. A. Bayer and J. C. Lagarias, *The nonlinear geometry of linear programming. I. Affine and projective scaling trajectories*, *Trans. Amer. Math. Soc.* (to appear).
- [BL] —, *Karmarkar's linear programming algorithm and Newton's method*, *Math. Programming* (to appear).
- [BM] S. Bochner and W. T. Martin, *Several complex variables*, Princeton Univ. Press, Princeton, N. J., 1948.
- [BT] B. Bernstein and R. A. Toupin, *Some properties of the Hessian of a strictly convex function*, *J. Reine Angew. Math.* **210** (1962), 65–72.
- [CH] R. Courant and D. Hilbert, *Methods of mathematical physics*, Vols. I, II, Wiley, New York, 1962.
- [D1] I. I. Dikin, *Iterative solution of problems of linear and quadratic programming*, *Dokl. Akad. Nauk SSSR* **174** (1967), 747–748. (English transl., *Soviet Math. Dokl.* **8** (1967), 674–675.)
- [D2] —, *About the convergence of an iterative process*, *Controllable Systems IM IK SO AN SSSR* 1974, No. 12, 54–60. (Russian)
- [F] W. Fenchel, *On conjugate convex functions*, *Canad. J. Math.* **1** (1949), 73–77.
- [FM] A. V. Fiacco and G. W. McCormick, *Nonlinear programming: Sequential unconstrained minimization techniques*, Wiley, New York, 1968.
- [F1] W. H. Fleming, *Functions of several variables*, Addison-Wesley, Reading, Mass., 1965.
- [GZ] C. B. Garcia and W. I. Zangwill, *Pathways to solutions, fixed points and equilibria*, Prentice-Hall, Englewood Cliffs, N. J., 1981.
- [Go] C. Gonzaga, *An algorithm for solving linear programming problems in $O(n^3L)$ operations*, *Progress in Mathematical Programming, Interior-Point and Related Methods* (N. Megiddo, Ed.), Springer-Verlag, New York, 1989, pp. 1–28.
- [GMN] J. Guckenheimer, J. Moser, and S. E. Newhouse, *Dynamical systems*, C.I.M.E. Lectures, Birkhäuser, Boston, Mass., 1980.
- [H] D. Hilbert, *Grundlagen der Geometrie*, 7th ed., Leipzig, 1930. (English transl., *Foundations of geometry*.)
- [HK] K. Hoffman and R. Kunze, *Linear algebra*, Prentice-Hall, Englewood Cliffs, N. J., 1961.
- [Ho] J. Hooker, *Karmarkar's linear programming algorithm*, *Interfaces* **16** (1986), 75–90.
- [K] N. Karmarkar, *A new polynomial time algorithm for linear programming*, *Combinatorica* **4** (1984), 373–395.
- [KLSW] N. Karmarkar, J. C. Lagarias, L. Slutsman and P. Wang, *Power-series variants of Karmarkar-type algorithms*, *A.T.&T. Tech. J.* **68** (to appear).
- [KV] S. Kapoor and P. M. Vaidya, *First algorithms for convex quadratic programming and multi-commodity flows*, *Proc. 18th ACM Sympos. on Theory of Computing*, 1986, pp. 147–159.
- [KMY] M. Kojima, S. Mizuno, and A. Yoshise, *A primal-dual interior point algorithm for linear programming*, *Progress in Mathematical Programming, Interior-Point and Related Methods* (N. Megiddo, Ed.), Springer-Verlag, New York, 1989, pp. 29–48.
- [L3] J. C. Lagarias, *The non-linear geometry of linear programming. III, Projective Legendre transform coordinates and Hilbert geometry*, *Trans. Amer. Math. Soc.* (to appear).
- [Ln] C. Lanczos, *The variational principles of mechanics*, Univ. of Toronto Press, Toronto, 1949.

- [M] N. Megiddo, *On the complexity of linear programming*, Advances in Economic Theory (T. Bewley, Ed.), Cambridge Univ. Press, 1986.
- [M2] —, *Pathways to the optimal set in linear programming*, Progress in Mathematical Programming, Interior-Point and Related Methods (N. Megiddo, Ed.), Springer-Verlag, New York, 1989, pp. 131–158.
- [MS] N. Megiddo and M. Shub, *Boundary behavior of interior point algorithms in linear programming*, preprint, IBM Research Report RJ 5319, 1986.
- [Mo] J. Moser, *Various aspects of integrable Hamiltonian systems*, pp. 233–290 in [GMN].
- [N] J. L. Nazareth, *Homotopy techniques in linear programming*, Algorithmica **1** (1986), 528–535.
- [Os] M. R. Osborne, *Dual barrier functions with superfast rates of convergence for the linear programming problem*, J. Austral. Math. Soc. (Ser. B) **29** (1987), 39–58.
- [Re] J. Renegar, *A polynomial-time algorithm, based on Newton's method, for linear programming*, Math. Programming **40** (1988), 59–94.
- [R1] R. T. Rockafellar, *Conjugates and Legendre transforms of convex functions*, Canad. J. Math. **19** (1967), 200–205.
- [R2] —, *Convex analysis*, Princeton Univ. Press, Princeton, N. J., 1970.
- [Sol] Gy. Sonnevend, *An "analytical centre" for polyhedrons and new classes of global algorithms for linear (smooth, convex) programming*, Proc. 12th IFIP Conf. System Modelling, Budapest, 1985, Lecture Notes in Computer Science, 1986.
- [So2] —, *A new method for solving a set of linear (convex) inequalities and its application for identification and optimization*, Proc. Sympos. on Dynamic Modelling, IFAC-IFORS, Budapest, June 1986.
- [SW] J. Stoer and C. Witzgall, *Convexity and optimization in finite dimensions. I*, Springer-Verlag, New York, 1970.
- [TB] M. Todd and B. Burrell, *An extension of Karmarkar's algorithm for linear programming using dual variables*, Algorithmica **1** (1986), 409–424.
- [V] P. Vaidya, *An algorithm for linear programming which requires $O(((m+n)n^2 + (m+n)^{1.5}n)L)$ arithmetic operations*, Proc. 19th ACM Sympos. on Theory of Computing, 1987, pp. 29–38.
- [VMF] R. J. Vanderbei, M. J. Meketon and B. A. Freedman, *A modification of Karmarkar's linear programming algorithm*, Algorithmica **1** (1986), 395–407.

DEPARTMENT OF MATHEMATICS, COLUMBIA UNIVERSITY, NEW YORK NEW YORK, 10027

AT & T BELL LABORATORIES, MURRAY HILL, NEW JERSEY 07974