The Nonlinear Geometry of Linear Programming
I. Affine and Projective Scaling Trajectories

by

D. A. Bayer
Columbia University
New York, New York

J. C. Lagarias
AT&T Bell Laboratories
Murray Hill, New Jersey
(June 24, 1986 revision)

ABSTRACT

This series of papers studies a geometric structure underlying Karmarkar's projective scaling algorithm for solving linear programming problems. A basic feature of the projective scaling algorithm is a vector field depending on the objective function which is defined on the interior of the polytope of feasible solutions of the linear program. The geometric structure we study is the set of trajectories obtained by integrating this vector field, which we call P-trajectories. In order to study P-trajectories we also study a related vector field on the linear programming polytope, which we call the affine scaling vector field, and its associated trajectories, called A-trajectories. The affine scaling vector field is associated to another linear programming algorithm, the affine scaling algorithm. These affine and projective scaling vector fields are each defined for liner programs of a special form, called strict standard form and canonical form, respectively.

This paper defines and presents basic properties of P-trajectories and A-trajectories. It reviews the projective and affine scaling algorithms, defines the projective and affine scaling vector fields, and gives differential equations for P-trajectories and A-trajectories. It presents Karmarkar's interpretation of A-trajectories as steepest descent paths of the objective function $\langle \mathbf{c}, \mathbf{x} \rangle$ with respect to the Riemannian geometry $ds^2 = \sum_{c=1}^{n} \frac{dx_i \, dx_i}{x_i^2}$ defined in the interior of the positive orthant. It establishes a basic relation connecting P-trajectories and A-trajectories, which is that P-trajectories of a Karmarkar canonical form linear program are radial projections of A-trajectories of an associated standard form linear program. As a consequence there is a polynomial time linear programming algorithm using the affine scaling vector field of this associated linear program: this algorithm is essentially Karmarkar's algorithm.

These trajectories will be studied in subsequent papers by a nonlinear change of variables which we call Legendre transform coordinates. It will be shown that both P-trajectories and A-trajectories have two distinct geometric interpretations: parametrized one way they are algebraic curves, while parametrized another way they are geodesics (actually distinguished chords) of a geometry isometric to a Hilbert geometry on a suitable polytope or cone. A summary of the main results of this series of papers is included.

The Nonlinear Geometry of Linear Programming
I. Affine and Projective Scaling Trajectories

by

D. A. Bayer
Columbia University
New York, New York

J. C. Lagarias
AT&T Bell Laboratories
Murray Hill, New Jersey

(June 24, 1986 revision)

(Preliminary draft: May 1, 1986)

1. Introduction

In 1984 Narendra Karmarkar [K] introduced a new linear programming algorithm which is proved to run in polynomial time in the worst case. Computational experiments with this algorithm are very encouraging, suggesting that it will surpass the performance of the simplex algorithm on large linear programs which are sparse in a suitable sense. The basic algorithm has been extended to fractional linear programming [A] and convex quadratic programming [KV].

Karmarkar's algorithm, which we call the *projective scaling algorithm*,* is a piecewise linear algorithm defined in the relative interior of the polytope **P** of feasible solution of a linear programming problem. The algorithm takes a series of (linear) steps, and the step direction is specified by a vector field $\mathbf{v}(\mathbf{x})$ defined at all parts **x** in the relative interior of **P**. This vector field depends on the linear program constraints and on the objective function. The projective scaling algorithms uses projective transformations to compute this vector field (see Section 4.)

Our viewpoint is that the fundamental mathematical object underlying the projective scaling algorithm is the set of trajectories obtained by following this vector field exactly. That is, given a projective scaling vector field $\mathbf{v}(\mathbf{x})$ and an initial point $\mathbf{x}_0$ one obtains (parametrized) curves by integrating the vector field for all initial conditions:

---

* This choice of name is explained in Section 4.

$$\frac{d\mathbf{x}}{dt} = \mathbf{v}(\mathbf{x}) \tag{1.1a}$$

$$\mathbf{x}(0) = \mathbf{x}_0 \ . \tag{1.1b}$$

A P-trajectory (or *projective scaling trajectory*) is an (unparametrized) point set specified by such a curve extended to the full range of t where a solution to the differential equation (1.1) exists.

In this series of papers our first object is to study the P-trajectories, to give several algebraic and geometric characterizations of them, and to prove facts about their behavior. We will show the P-trajectories are interesting in their own right. They have an extremely rich mathematical structure, involving connections to algebraic geometry, differential geometry, partial differential equations, classical mechanics and convexity theory. This structure can be exploited in several ways to give new linear programming algorithms, which we will discuss elsewhere.

Our results concerning P-trajectories are derived using their connection to another set of trajectories, which we call *A-trajectories* (or *affine scaling trajectories*), which are easier to study. Our second object is therefore to give several geometric characterizations of A-trajectories. The A-trajectories arise from integrating a vector field associated to another interior-point linear programming algorithm, which we call the *affine scaling algorithm*.* The affine scaling vector field has been discovered and studied by [B], [VMK], and many others. There is a simple relation between the P-trajectories of a linear programming problem and the A-trajectories associated to an associated linear program that is given in Section 6.

We mention some background and related work. The idea of following trajectories to solve nonlinear equations has a long history, and is a basic methodology in non-linear programming [FM], [GZ]. From this perspective Karmarkar's projective scaling algorithm can be viewed as a *homotopy restart method* using the system of P-trajectories, as was observed by Nazareth [N] (see also [GZ], Sect. 15.4). One method of constructing trajectories is by means of a parameterized family of barrier functions, see [FM] Chapter 5. In this connection it is possible to relate A-trajectories and P-trajectories to trajectories defined using a parameterized family of logarithmic barrier functions. (See equation (2.12) following.) N. Megiddo [M2]

---

* The rationale for this name is given in Section 4.

studies trajectories obtained from other parameterized families of nonlinear optimization problems. The geometric behavior of A-trajectories is being studied by M. Shub [S]. Finally J. Renegar [R] has made use of P-trajectories together with new ideas to construct a new interior-point linear programming algorithm which uses Newton's method to follow the central P-trajectory. Renegar's algorithm runs in polynomial time and requires only $O(\sqrt{n}\ L)$ iterations in the worst case. This improves on Karmarkar's [K] worst-case bound of $O(nL)$ iterations. Surveys of Karmarkar's algorithm and recent developments appear in [H], [M1].

In Section 2 we first summarize the main results of this series of papers, and then summarize the contents of this paper in detail. Section 3 gives a brief description of the affine and projective scaling linear programming algorithms, which is independent of the rest of the paper.

We are indebted to many people for aid during this research. We wish to thank particularly Jim Reeds for conversations on convex analysis and references to Rockafellar's work and Peter Doyle for conversations on Riemannian geometry. We are indebted to Narendra Karmarkar for permission to include his steepest descent interpretation of A-trajectories in Section 5 of this paper.

2. Summary of results

In this section we give an overview of the main results of this series of papers, and then summarize the contents of this paper in more detail.

A. Main results — overview

We give two distinct geometric interpretations of the P-trajectories, corresponding to two different parameterizations of these trajectories. First, in terms of the coordinate system of the linear program, each P-trajectory is a piece of a (real) algebraic curve. The P-trajectory can then be naturally extended to the full (complex) algebraic curve of which it is a part. Viewed algebraically it is then a branched covering of the projective line $\mathbb{P}^1(\mathbb{C})$, while viewed analytically it is a Riemann surface. The objective function value gives a natural parametrization of the P-trajectory. Second, there is a metric $d_H(\cdot\,,\cdot)$ defined on the interior of the polytope $\mathbf{P}$ such that each P-trajectory is an extremal (‘‘geodesic’’) with respect to this metric. The resulting geometry is isometric to Hilbert's projective geometry defined on the interior of a polytope $\mathbf{P}^*$ which is combinatorially dual to $\mathbf{P}$, (Hilbert's geometry is defined in [H], Appendix 2). This geometry is a *chord geometry* in the sense of Busemann [Bu2] and the P-trajectories are the *distinguished chords* in the sense of Busemann-Phadke [BP]. The P-trajectory inherits an obvious parameterization from the metric $d_H(\cdot\,,\,\cdot)$.

Our results about P-trajectories will be proved using their close connection to A-trajectories. Karmarkar's P-trajectories are defined for linear programs of the following special form which we call *canonical form*:

$$\text{minimize } \langle \mathbf{c},\ \mathbf{x} \rangle \tag{2.1a}$$

subject to

$$A\mathbf{x} = \mathbf{0}\ , \tag{2.1b}$$

$$e^T \mathbf{x} = n\ , \tag{2.1c}$$

$$\mathbf{x} \geq 0\ , \tag{2.1d}$$

with side conditions

$$A\mathbf{e} = \mathbf{0}\ . \tag{2.1e}$$

Here $\langle \mathbf{c},\ \mathbf{x} \rangle = \mathbf{c}^T \mathbf{x}$ denotes the usual Euclidean inner product, and $\mathbf{e} = (1,\ 1,\ \ldots\ ,\ 1)^T$. There is a simple relation between P-trajectories of a canonical form linear program (2.1) and A-trajectories of the associated linear programming problem:

$$\text{minimize } \langle \mathbf{c},\ \mathbf{x} \rangle \tag{2.2a}$$

subject to

$$A\mathbf{x} = \mathbf{0}\ , \tag{2.2b}$$

$$\mathbf{x} \geq \mathbf{0}\ , \tag{2.2c}$$

with side conditions

$$A\mathbf{e} = \mathbf{0}\ . \tag{2.2d}$$

The relation is that the radial projection of an A-trajectory onto the hyperplane $\mathbf{e}^T \mathbf{x} = n$ is a P-trajectory. (Theorem 6.1 of this paper.) There is a second relation between P-trajectories and A-trajectories of the linear program (2.1) which we give later in this summary.

The A-trajectories also have several geometric interpretations. First, N. Karmarkar has observed that

A-trajectories of a *standard form* linear program:

$$\text{minimize } \langle \mathbf{c},\ \mathbf{x} \rangle$$

subject to

$$A\mathbf{x} = \mathbf{b}\ ,$$

$$\mathbf{x} \ge \mathbf{0}\ .$$

having a feasible solution $\mathbf{x}$ with all $x_i > 0$ may be interpreted as *steepest descent curves* of $\langle \mathbf{c},\ \mathbf{x} \rangle$ with respect to the Riemannian metric $ds^2 = \sum_{i=1}^{n} \frac{dx_i\, dx_i}{x_i^2}$ defined on the interior of the positive orthant $\mathrm{Int}(\mathbb{R}_+^n) = \{\mathbf{x} : \text{all } x_i > 0\}$. We include a proof of this fact with his permission. Second, there is a metric $d_E(\cdot\ ,\ \cdot)$ defined on the relative interior Rel-Int($\mathbf{P}$) of the polytope $\mathbf{P}$ of feasible solutions such that the affine scaling curves are geodesics with respect to this metric. This metric is isometric to Euclidean geometry restricted to a cone. If $\mathbf{P}$ is a bounded polytope, then it is isometric to Euclidean geometry on $\mathbb{R}^k$ where $k = \dim(\mathbf{P})$. The A-trajectories are algebraic curves with respect to the metric parameter, and this metric parameter is algebraically related to the linear program coordinates, so that the A-trajectories are pieces of (real) algebraic curves in the linear program coordinates. Hence A-trajectories also extend to branched coverings of $P^1(\mathbb{C})$ which are Riemann surfaces. Third, for a linear program in the homogeneous form (1.3) the A-trajectories also have a Hilbert geometry interpretation. The polytope $\mathbf{P}$ of feasible solutions to a homogeneous linear program (2.1) is a cone, and there is a pseudo-metric $\overline{d}_H(\cdot\ ,\ \cdot)$ on Int($\mathbf{P}$) such that the geometry induced by this pseudo-metric is isometric to Hilbert's projective geometry on the dual cone, and the A-trajectories are a set of *distinguished chords*. (A pseudo-metric satisfies the triangle inequality but may have $d_H(\mathbf{x}_1,\ \mathbf{x}_2) = 0$ with $\mathbf{x}_1 \ne \mathbf{x}_2$.)

Our results on P-trajectories and A-trajectories are obtained using a nonlinear change of coordinates. We call the new coordinate system we construct *Legendre transform coordinates*. This name is chosen because the coordinates are constructed using a Legendre transform mapping attached to a logarithmic barrier function, cf. Rockafellar [R2], Section 25. We now describe these Legendre transform coordinates in a special case. Consider a linear program in the following special form, which we call *standard form*:

$$\text{minimize } \langle \mathbf{c},\ \mathbf{x} \rangle \tag{2.3a}$$

subject to

$$A\mathbf{x} = \mathbf{b} \tag{2.3b}$$

$$\mathbf{x} \geq \mathbf{0} \tag{2.3c}$$

with side conditions

$$AA^T \text{ is an invertible matrix .} \tag{2.3d}$$

We say such a linear program has in *strict standard form constraints* if it has a feasible solution $\mathbf{x} = (x_1, \ldots, x_n)$ with all x_i positive. The Legendre transform coordinates are determined by the constraints of the linear program and do not depend on the objective function. Let $\boldsymbol{H}$ denote the set of constraints of a strict standard form linear program, and let $P_{\boldsymbol{H}}$ be its associated polytope of feasible solutions. The relative interior Rel-Int($\mathbf{P}_{\boldsymbol{H}}$) of the polytope of feasible solutions of this linear program then is nonempty and lies in the interior Int($\mathbb{R}^n_+$) of the positive orthant. We consider the logarithmic barrier function f: Int$(\mathbb{R}^n_+) \to \mathbb{R}^n$ defined by

$$f_{\boldsymbol{H}}(\mathbf{x}) = -\sum_{i=1}^{n} \log x_i \ . \tag{2.4}$$

This function has the gradient

$$\nabla f_{\boldsymbol{H}}(\mathbf{x}) = \left[-\frac{1}{x_1}, -\frac{1}{x_2}, \cdots, -\frac{1}{x_n} \right]^T . \tag{2.5}$$

The associated Legendre transform coordinate mapping $\phi_{\boldsymbol{H}}$ maps Rel-Int($\mathbf{P}$) into the subspace

$$A^{\perp} = \{\mathbf{x} : A\mathbf{x} = \mathbf{0}\}$$

of $\mathbb{R}^n$ and is defined by

$$\phi_{\boldsymbol{H}}(\mathbf{x}) = \pi_{A^{\perp}}(\nabla f(\mathbf{x})) \ , \tag{2.6}$$

where $\pi_{A^{\perp}}$ is the orthogonal projection operator onto the subspace $A^{\perp}$. This projection operator is given explicitly by the formula

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

whenever AA^T is invertible. We show as a special case of theorems proved in part II for a strict standard form linear program whose polytope $P_{\boldsymbol{H}}$ of feasible solutions is bounded that Legendre transform mapping

$$\phi_{\boldsymbol{H}} : \text{Rel-Int}(\mathbf{P}_{\boldsymbol{H}}) \to A^{\perp}$$

is a real-analytic diffeomorphism onto all of $A^{\perp}$. In particular it is one-to-one and onto, so there is a unique point $\mathbf{x}_{\boldsymbol{H}}$ in $\boldsymbol{H}$ such that

$$\phi_{\boldsymbol{H}}(\mathbf{x}_{\boldsymbol{H}}) = \mathbf{0} \ , \tag{2.7}$$

and we call this point the *center*. We show in part II that this definition of center coincides with Karmarkar's definition of center when the constraints $\boldsymbol{H}$ are in Karmarkar's canonical form. The center has a geometric interpretation as that point in Rel-Int($P_{\boldsymbol{H}}$) that maximizes the function $w_{\boldsymbol{H}}(\mathbf{x})$ giving the product of the distances of $\mathbf{x}$ to each of the hyperplanes defining the boundary of each inequality constraint. For the constraints $\boldsymbol{H}$ we have

$$w_{\boldsymbol{H}}(\mathbf{x}) = \prod_{i=1}^{n} x_i \ .$$

In part II we show that it is possible to define Legendre transform coordinates $\phi_{\boldsymbol{H}}$ for *any* set $\boldsymbol{H}$ of linear program constraints, including both equality and inequality constraints. In the general case several extra complications appear. These include the facts that the range space of the Legendre transform coordinate mapping in the most general case is the interior of a polyhedral cone, that the mapping may be many-to-one, and that a *center* may not exist. We show that in all cases these coordinates transform contravariantly under an invertible affine transformation $\mathbf{A} : \mathbb{R}^n \to \mathbb{R}^n$. To describe this, let $\mathbf{A}(\mathbf{x}) = L\mathbf{x} + \mathbf{c}$ where $\mathbf{L}$ is an invertible linear map, and let $L^* = (L^T)^{-1}$ denote its dual map. Then the linear program constraints $\boldsymbol{H}$ are carried to the linear program constraints $\mathbf{A}(\boldsymbol{H})$. The contravariance property is expressed in the following commutative diagram:

$$\begin{tikzcd}
\text{Rel-Int}(P_{\boldsymbol{H}}) \arrow[r, "\mathbf{A}"] \arrow[d, "\phi_{\boldsymbol{H}}"] & \text{Rel-Int}(P_{\mathbf{A}(\boldsymbol{H})}) \arrow[d, "\phi_{\mathbf{A}(\boldsymbol{H})}"] \\
A^{\perp} & (L(\mathbf{A}))^{\perp} \ . \arrow[l, "\bar{L}^*"']
\end{tikzcd} \tag{2.8}$$

where $\bar{L}^* = \pi_{A^{\perp}}(L^*)$ is a vector space isomorphism.

The Legendre transform mapping is given by rational functions of the linear program coordinates x_i. This mapping is one-to-one for a strict standard form problem, hence it then has an inverse mapping $\phi_{\boldsymbol{H}}^{-1}$ which is necessarily given by algebraic functions of the Legendre transform coordinate space. The logarithmic barrier function $f_{\boldsymbol{H}}(\mathbf{x})$ can be shown to be strictly convex on Rel-Int($P_{\boldsymbol{H}}$) in this case and by the general theory of convex analysis we can construct its (Fenchel) *conjugate function* $g : A^{\perp} \rightarrow |\mathrm{R}^n$ which is defined for $\mathbf{y} \in \mathbf{A}^{\perp}$ as the solution of the problem:

$$g_{\boldsymbol{H}}(\mathbf{y}) = \sup_{\mathbf{x} \in \mathrm{Rel-Int}(P_H)} (\langle \mathbf{x}, \mathbf{y} \rangle - f(\mathbf{x})) \ , \tag{2.9}$$

(see: [F],[R2] 12.2.2.) Then the Legendre transform duality theorem ([R2], Theorem 26.4) implies that $\phi_{\boldsymbol{H}}^{-1}$ is given by

$$\phi_{\boldsymbol{H}}^{-1}(\mathbf{y}) = \nabla g_{\boldsymbol{H}}(\mathbf{y}) \ . \tag{2.10}$$

At present we cannot directly use this explicit formula due to our lack of knowledge how to compute the Fenchel conjugate function except in special cases.

The Legendre transform mapping originally arose as a tool in studying ordinary and partial differential equations, cf. [CH], Vol. II, pp. 32-39. In particular it is used to convert the Lagrangian formulation of a classical mechanical system to the Hamiltonian formulation (see [A], pp. 59-65, [Ln]). This connection is not accidental — the second author will show elsewhere there is an interpretation of A-trajectories arising from a new family of completely integrable Hamiltonian dynamical systems [L].

The utility of Legendre transform coordinates is established in part III of this series of papers, where it is shown that it maps the set of A-trajectories of a strict standard form linear program with bounded feasible polyhedron to the complete set of parallel straight lines with slope

$$\mathbf{c}' = \pi_{A^{\perp}}(\mathbf{c}) \ . \tag{2.11}$$

The A-trajectories of a strict standard form linear program having an unbounded polyhedron of feasible solutions are mapped to a family of parallel half-lines or line segments having the same slope $\mathbf{c}'$. Consequently each A-trajectory is an inverse image of part of a straight line under the Legendre transform mapping $\phi_{\boldsymbol{H}}$. Since this mapping is a rational map each A-trajectory of a strict standard form linear program must be part of a real algebraic curve. Then since each P-trajectory of a canonical form linear program (2.1) is rationally related to an A-trajectory of the strict standard form linear program (2.2) it must also be a part of a real algebraic curve.

We distinguish the particular P-trajectory of a canonical form linear program (2.1) which passes though the center $\mathbf{x}_{\boldsymbol{H}}$ of its constraint set (2.1b)-(2.1d) and call it the *central P-trajectory with objective function* $\langle \mathbf{c}, \mathbf{x} \rangle$. We define the *central A-trajectory* of a strict standard form linear program having a center with objective function $\langle \mathbf{c},\ \mathbf{x} \rangle$ analogously. We prove in part III for a canonical form linear program (2.1) that the central P-trajectory and central A-trajectory with objective function $\langle \mathbf{c},\ \mathbf{x} \rangle$ coincide. This is a second relation between P-trajectories and A-trajectories. In particular it implies that the central P-trajectory is a straight line in Legendre transform coordinates.

The central P-trajectory (which is the central A-trajectory) plays a fundamental role in Karmarkar's algorithm. In part III we give a number of other geometric characterizations of this trajectory, the most interesting of which is that it is the *locus of centers* of the linear programs obtained from the given standard form linear program by adding the extra equality constraint,

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

where λ ranges over the possible values of the objective function in Rel-Int($P_{\boldsymbol{H}}$). Another related interpretation of the central P-trajectory of a standard form problem (2.3) is that it is described by the solution $\mathbf{x}(\mu) = \mathbf{x}(\mu\ ;\ \phi)$ of a family of non-linear fixed-point problems parametrized by μ. This family is given by:

$$\text{minimize } \phi(\langle \mathbf{c},\ \mathbf{x} \rangle) - \mu \sum_{i=1}^{n} \log x_i \tag{2.12a}$$

subject to

$$A\mathbf{x} = \mathbf{b}\ , \tag{2.12b}$$

$$\mathbf{x} > \mathbf{0}\ , \tag{2.12c}$$

where $\phi(t) : \mathbb{R} \to \mathbb{R}$ is any one-one onto monotonic increasing function and $-\infty < \mu < \infty$. This representation describes the central P-trajectory as the set of solutions to a parametrized family of logarithmic barrier functions.

We also analyze the behavior of non-central P-trajectories. In part III we prove that every P-trajectory lies in a plane in Legendre transform coordinates. For a non-central P-trajectory this plane is determined the line given by the central P-trajectory (for the same objective function) together with any point on the given non-central P-trajectory. Any noncentral P-trajectory is *not* a straight line in Legendre transform coordinates. If the objective function $\langle \mathbf{c},\ \mathbf{x} \rangle$ is *normalized* in Karmarkar's sense, i.e. it takes the value $\mathbf{0}$ at the optimal solution of a canonical form linear program, then the non-central P-trajectories in Legendre transform coordinates for $\langle \mathbf{c},\ \mathbf{x} \rangle$ asymptotically approach the central P-trajectory as $\mathbf{x}$ approaches the optimal point.

Any noncentral P-trajectory of a canonical form linear program can be mapped to a central P-trajectory (in a different Legendre transform coordinate system) through a suitable projective transformation which transforms the linear program constraints $\boldsymbol{H}$ to a new set of linear program constraints $\boldsymbol{H}'$ which are also in canonical form. This follows immediately from Karmarkar's observation that a projective transformation exists taking an arbitrary point in Rel-Int($P_{\boldsymbol{H}}$) to the center of the transformed polytope.

In part I of this paper we define P-trajectories for canonical form linear programs and A-trajectories for strict standard form linear programs. Indeed the original definitions in terms of the projective scaling vector fields and affine scaling vector fields only make sense in this context, see Section 4 of this paper. In part II we define Legendre transform coordinates for any set of linear programming constraints. In part III we then take the characterization of A-trajectories as straight lines in Legendre transform coordinates with slope $\mathbf{c}'$ given in (2.11) as a *definition* of A-trajectories valid for all linear programs. In part III we also use the relation between P-trajectories of the canonical form problem (2.1) and A-trajectories of the homogeneous standard form problem (2.2) to give a *definition* of P-trajectories valid for all linear programs. With these definitions, we prove that A-trajectories are preserved by invertible affine transformations of variables, and that P-trajectories are preserved by a (slightly restricted) set of projective transformations, which includes all invertible affine transformations.

In part III we also use Legendre transform coordinates to compute the power series expansion of the central P-trajectory. These power series coefficients assume a very simple form which is easy to compute. This leads to the possibility of practical linear programming algorithms based on power series approaches. This will be discussed elsewhere [BKL].

Now that we have established that P-trajectories and A-trajectories are parts of real algebraic curves, we can define them *outside the polytopes* Rel-Int($P_{\boldsymbol{H}}$) *on which they were originally defined*. These algebraic curves extend into other cells determined by the arrangement of hyperplanes obtained from the inequality constraints of the linear program by regarding them as equality constraints. It turns out that each extended A-trajectory (resp. P-trajectory) visits a cell of the arrangement at most once, and that in each cell it visits an extended A-trajectory (resp. P-trajectory) is an A-trajectory (resp. P-trajectory) for a linear program having that cell as its set of feasible solutions. These linear programs are obtained from the original linear program by reversing a suitable subset of the inequality constraints.

In part IV we use Legendre transform coordinates to show that P-trajectories are ''geodesics'' of a metric geometry isometric to Hilbert's geometry on the interior of the dual polytope, as well as an analogous result for A-trajectories for homogeneous standard form linear programs.

B. Results of this paper

In this paper we define and present basic properties of P-trajectories and A-trajectories. In section 3 we briefly review the projective and affine scaling algorithms, in order to provide background and perspective on later developments. In section 4 we derive the affine and projective scaling vector fields, and then obtain differential equations for A-trajectories and P-trajectories. The affine scaling vector field is calculated using an affine rescaling of coordinates, and the projective scaling vector field is calculated using a projective rescaling of coordinates. (This motivates our choice of names for these algorithms.) In order to apply these rescaling transformations the linear programs must be of special forms: *strict standard form* for the affine scaling algorithm, and the *canonical form* (2.1) for the projective scaling algorithm. Consequently A-trajectories are defined in part I only for standard form problems and P-trajectories only for canonical form problems. (In part III of this series of papers we will extend the definition of A-trajectory and P-trajectory to other linear programs.) A connection with fractional linear programming is also made in Section 4.

In section 5 we give Karmarkar's geometric interpretation of A-trajectories for standard form linear programs as steepest descent curves with respect to the Riemannian metric $ds^2 = \sum_{i=1}^{n} \frac{dx_i \, dx_i}{x_i^2}$. This Riemannian metric has a rather special property: it is invariant under projective transformations taking the positive orthant $\mathrm{Int}(\mathbb{R}_+^n)$ into itself. The results of this section are not used elsewhere in these papers.

In Section 6 we derive a fundamental relation between P-trajectories and A-trajectories, which is that the P-trajectories of the canonical form linear program (2.1) are radial projections of the associated homogeneous strict standard form linear program obtained by dropping the inhomogeneous constraint $\langle \mathbf{e} , \mathbf{x} \rangle = n$ from (2.1). In particular these P-trajectories and A-trajectories are algebraically related.

In the final section 7 a simple consequence of this relation is made. It is that a polynomial time linear programming algorithm for a canonical form linear program results from following the affine scaling vector field of the associated homogeneous standard form problem, which is:

$$\text{minimize } \langle \mathbf{c},\ \mathbf{x} \rangle \tag{2.13a}$$

subject to

$$A\mathbf{x} = \mathbf{0} \tag{2.13b}$$

$$\mathbf{x} \geq \mathbf{0} \tag{2.13c}$$

with side conditions

$$A\mathbf{e} = \mathbf{0} \tag{2.13d}$$

$$AA^T \text{ is invertible .} \tag{2.13e}$$

The piecewise linear steps of the resulting ‘‘affine scaling’’ algorithm then radially project onto the piecewise linear steps of Karmarkar’s projective scaling algorithm, so this ‘‘affine scaling’’ algorithm is essentially Karmarkar’s projective scaling algorithm. We mention it because it is an example of a provably polynomial time linear programming algorithm based on the affine scaling vector field. A final observation is that this ‘‘affine scaling’’ algorithm is not solving the linear program (2.13), but rather is solving the fractional linear program with objective function $\dfrac{\langle \mathbf{c},\ \mathbf{x} \rangle}{\langle \mathbf{e},\ \mathbf{x} \rangle}$ subject to the homogeneous standard form constraints (2.13b)-(2.13e). The results of Section 7 are perhaps best viewed as an interpretation of Karmarkar’s projective scaling algorithm as an ‘‘affine scaling’’ algorithm for a particular fractional linear programming problem. In this connection see [A].

3. Affine and projective scaling algorithms

In this section we briefly summarize Karmarkar’s projective scaling algorithm [K] and the affine scaling algorithm, described in [B] and [VMF]. We start with Karmarkar’s algorithm. Karmarkar’s *projective scaling algorithm* is a piecewise linear algorithm which proceeds in steps through the relative interior of the polytope of feasible solutions to the linear programming problem. It has the following main features: an *initial starting point*, a choice of *step direction*, a choice of *step size* at each step, and a *stopping rule*.

The initial starting point is supplied by the fact that the algorithm is defined only for linear programming problems whose constraints are of a special form, which we call (*Karmarkar) canonical form*, which comes with a particular initial feasible starting point which Karmarkar calls the *center*. Karmarkar's algorithm also requires that the objective function $\mathbf{z} = \langle \mathbf{c}, \mathbf{x} \rangle$ satisfy the special restriction that its value at the optimum point of the linear program is zero. We call such an objective function a *normalized objective function*. In order to obtain a general linear programming algorithm, Karmarkar [K, Section 5] shows how any linear programming problem may be converted to an associated linear programming problem in canonical form which has a normalized objective function. This conversion is done by combining the primal and dual problems, then adding slack variables and an artificial variable, and as a last step using a projective transformation. An optimal solution of the original linear programming problem can be easily recovered from an optimal solution of the associated linear program constructed in this way. The *step direction* is supplied by a vector field defined on the relative interior Rel-Int(P) of the polytope of feasible solutions of a canonical form linear program. Karmarkar's vector field depends on both the constraints and the objective function. It can be defined for any objective function on a canonical form problem, whether or not this objective function is normalized. However Karmarkar only proves good convergence properties for the piecewise linear algorithm he obtains using a normalized objective function. Karmarkar's vector field is defined implicitly in his paper [K], in which projective transformations serve as a means for its calculation. This is described in Section 4.

The *step size* in Karmarkar's algorithm is computed using an auxiliary function $g \colon \text{Rel-Int}(P) \to \mathbb{R}$ which he calls a *potential function*. In fact $g \colon (\mathbb{R}^n)^+ \to \mathbb{R}$ is defined by

$$g(\mathbf{x}) = n \log \mathbf{c}^T \mathbf{x} - \sum_{i=1}^{n} \log x_i \ .$$

It depends on the objective function $\mathbf{c}^T \mathbf{x}$ and approaches $-\infty$ at the optimal point on the boundary ∂P of the polytope P of feasible solutions, and approaches $+\infty$ at all other boundary points. It is related to the objective function by the inequality

$$g(\mathbf{x}) \geq n \log (\mathbf{c}^T \mathbf{x}) \ . \tag{3.1}$$

If $\mathbf{x}_j$ is the starting point of the j^{th} step and $\mathbb{R}_+ \mathbf{v}$ the step direction, then the step size is taken to arrive at that point $\mathbf{x}_{j+1}$ on the ray $\mathbf{x} + \mathbb{R}_+ \mathbf{v}$ which minimizes $g(\mathbf{x})$ on this ray. If $\mathbf{x}_{j+1}$ is not an optimal point, then $\mathbf{x}_{j+1}$ remains in Rel-Int(P). Karmarkar proves that

$$g(\mathbf{x}_{j+1}) \le g(\mathbf{x}_j) - \frac{1}{5} \qquad (3.2)$$

provided that $\mathbf{c}^T\mathbf{x}$ is a normalized objective function. Finally, the *stopping rule* is related to the input data and to the bound (3.2) on the potential function. If (3.2) fails to hold at any step, the original L.P. was infeasible or unbounded. If we start at the center $\mathbf{x}_0 = \mathbf{e}$ then

$$g(\mathbf{x}_0) = n \log \mathbf{c}^T\mathbf{x}_0 \ .$$

With (2.1) and (2.2) this implies for a normalized objective function that

$$\frac{\mathbf{c}^T\mathbf{x}_j}{\mathbf{c}^T\mathbf{x}_0} \le e^{-\frac{1}{5}j} \ .$$

It is known that there is a bound L easily computable from the input data of a canonical form linear program with normalized objective function such that

$$\mathbf{c}^T\mathbf{w} \ge 2^{-L}$$

for any non-optimal vertex $\mathbf{w}$ of the polytope. When $e^{-\frac{1}{5}j} \le 2^{-L}$ the algorithm is stopped, and one locates a vertex $\mathbf{w}$ of $\mathbf{P}$ with

$$\mathbf{c}^T\mathbf{w} \le \mathbf{c}^T\mathbf{x}_j \ , \qquad (3.4)$$

which is then guaranteed to be optimal. In practice one does not wait until the bound $e^{-\frac{1}{5}j} \le 2^{-L}$ is reached; instead every few iterates one derives a solution $\mathbf{w}$ to (3.4) and checks whether or not it is optimal.

The *affine scaling algorithm* is similar to the projective scaling algorithm. It differs in the following respects. The input linear program is one required to have constraints of a special form which we call *strict standard form constraints*. This form is less restricted than (Karmarkar) canonical form. It is described in detail in Section 4. The *step direction* is calculated using a different scaling transformation based on an affine change of variable; this justifies calling this algorithm the affine scaling algorithm. There are a number of different proposals for calculating the *step size*, one of which is to go a fixed fraction (say 95%) of the way to the boundary along the ray specified by the step direction. The *stopping rule* is the same as in Karmarkar's algorithm. The affine scaling algorithm using the fixed fraction step size has been proved (in both [B] and [VMF]) to converge to an optimum solution under suitable nondegeneracy conditions. The affine scaling algorithm has not been proved to run in polynomial time in the worst case, and this may well not be true.

In Section 7 we show that a particular special case of the affine scaling algorithm does give a provably polynomial time algorithm for linear programming. This occurs, however, because the resulting algorithm is essentially identical to Karmarkar's projective scaling algorithm.

4. Affine and projective scaling vector fields and differential equations

In this section we review the derivation of the affine and projective scaling vector fields as obtained by rescaling the coordinates of the position orthant $\mathbb{R}^n_+$.

A. Affine scaling vector field

We define the affine scaling vector field for linear programs of a special form which we call *strict standard form*. A *standard form* linear program is:

$$\text{minimize } \langle \mathbf{c}, \mathbf{x} \rangle \tag{4.1}$$

subject to

$$A\mathbf{x} = \mathbf{b} \tag{4.2a}$$

$$\mathbf{x} \geq \mathbf{0} \tag{4.2b}$$

with side condition

$$AA^T \text{ is invertible .} \tag{4.2c}$$

The invertibility condition (4.2c) guarantees that the projection operator $\pi_{A^\perp}$ which projects $\mathbb{R}^n$ onto the subspace $A^\perp = \{\mathbf{x} : A\mathbf{x} = \mathbf{0}\}$ is given by

$$\pi_A = I - A^T(AA^T)^{-1}A \ . \tag{4.3}$$

We define *standard form constraints* to be the constraint conditions (4.2). We say that a set of linear programming constraints is in *strict standard form if it is a set of standard form constraints and it has a feasible solution* $\mathbf{x} = (x_1, \ldots, x_n)$ *such that all* $x_i > 0$. *The notion of strict standard form constraints* $\boldsymbol{H}$ *is a mathematical convenience introduced to make it easy to describe Rel-Int(*$P_{\boldsymbol{H}}$*), which is then* $P_{\boldsymbol{H}} \cap \mathbb{R}^n_+$*, and to be able to give explicit formulae for the effect of affine scaling transformations (and for Legendre transform coordinates (2.6)). Note that any standard form linear program can be converted to one that is in strict standard form by dropping all variables* x_i *that are identically zero on* $P_{\boldsymbol{H}}$*. A homogeneous strict standard form problem* is a linear program having strict standard form constraints in which $\mathbf{b} = \mathbf{0}$, and its constraints are *homogeneous strict standard form constraints*.

In defining the affine scaling vector field we first consider a strict standard form linear program having the point $\mathbf{e} = (1, 1, \ldots, 1)^T$ as a feasible point. We define the *affine scaling direction* $\mathbf{v}_A(e\,;\,\mathbf{c})$ *at the point* $\mathbf{e}$ *to be the steepest descent direction for* $\langle \mathbf{c}, \mathbf{x} \rangle$ *at* $\mathbf{x}_0 = \mathbf{e}$, *subject to the constraint* $A\mathbf{x} = \mathbf{b}$, *so that*

$$\mathbf{v}_A(\mathbf{x}, \mathbf{c}) = -\pi_{A^\perp}(\mathbf{c}) \; . \tag{4.4}$$

This may be obtained by Lagrange multipliers as a solution to the constrained minimization problem:

$$\text{minimize } \langle \mathbf{c}, \mathbf{x} \rangle - \langle \mathbf{c}, \mathbf{e} \rangle \tag{4.5a}$$

subject to

$$\langle \mathbf{x} - \mathbf{e}, \mathbf{x} - \mathbf{e} \rangle = \varepsilon \; , \tag{4.5b}$$

$$A\mathbf{x} = \mathbf{b} \; , \tag{4.5c}$$

for any $\varepsilon > 0$.

Now we define the affine scaling vector field $v(\mathbf{d}\,;\,\mathbf{c})$ for an arbitrary strict standard form linear program at an arbitrary feasible point $\mathbf{d} = (d_1, \ldots, d_n)$ in

$$\text{Int}\,(\mathbb{R}^n_+) = \{\mathbf{x} : \text{all } x_i > 0\} \; .$$

Let $D = \text{diag}(d_1, \ldots, d_n)$ to be the diagonal matrix corresponding to $\mathbf{d}$, so that $\mathbf{d} = D\mathbf{e}$. We introduce new coordinates by the affine (scaling) transformation

$$\mathbf{y} = \Phi_D(\mathbf{x}) = D^{-1}\mathbf{x}$$

with inverse transformation

$$\Phi_D^{-1}(\mathbf{y}) = D\mathbf{y} = \mathbf{x} \; .$$

Under this change of variables the standard form program (4.1)-(4.2) becomes the following standard form program.

$$\text{minimize } \langle D\mathbf{c}, \mathbf{x} \rangle \tag{4.6}$$

subject to

$$AD\mathbf{y} = \mathbf{b} \tag{4.7a}$$

$$\mathbf{y} \geq \mathbf{0} \tag{4.7b}$$

with side condition

$$AD^2A^T \text{ is invertible } . \tag{4.7c}$$

Furthermore $\Phi_D(\mathbf{d}) = \mathbf{e}$. By definition the affine scaling direction for this problem is — $\pi_{(AD)^\perp}(D\mathbf{c})$, and we define the affine scaling vector $\mathbf{v}_A(\mathbf{d}\ ;\ \mathbf{c})$ as the pullback by Φ_D^{-1} of this vector, which yields

$$\mathbf{v}_A(\mathbf{d}\ ;\ \mathbf{c}) = -D\pi_{(AD)^\perp}(D\mathbf{c}) \ .$$

$$= -D(I - DA^T(AD^2A^T)^{-1}AD)D\mathbf{c} \tag{4.8}$$

We check that the affine scaling vector depends only on the component $\pi_{A^\perp}(\mathbf{c})$ of $\mathbf{c}$ in the $A^\perp$ direction, and summarize the discussion so far as a lemma.

Lemma 4.1. The affine scaling vector field for a standard form problem (4.1)-(4.2) having a feasible solution $\mathbf{x} = (x_1, \ldots, x_n)$ *with all* $x_i > 0$ *is*

$$\mathbf{v}_A(\mathbf{d}\ ;\ \mathbf{c}) = -D\pi_{(AD)^\perp}(D\mathbf{c}) \ . \tag{4.9}$$

In addition

$$\mathbf{v}_A(\mathbf{d}\ ;\ \mathbf{c}) = \mathbf{v}(\mathbf{d};\ \pi_{A^\perp}(\mathbf{c})) \ . \tag{4.10}$$

Proof. The formula (4.9) is just (4.8). Using

$$\pi_A(\mathbf{c}) = A^T(AA^T)^{-1}A\mathbf{c} = A^T\boldsymbol{\lambda} \ ,$$

since direct substitution in (4.9) yields

$$\mathbf{v}_A(\mathbf{d}\ ;\ \pi_A(\mathbf{c})) = -D^2A^T\boldsymbol{\lambda} + D^2A^T(AD^2A^T)^{-1}AD^2A^T\boldsymbol{\lambda} = \mathbf{0} \ ,$$

which proves (4.10). ■

B. Projective scaling vector field

We define the projective scaling vector field for linear programs in the following form, which we call *canonical form*:

$$\text{minimize } \langle \mathbf{c},\ \mathbf{x} \rangle \tag{4.11}$$

subject to

$$A\mathbf{x} = \mathbf{0} \ , \tag{4.12a}$$

$$\mathbf{e}^T\mathbf{x} = n \ , \tag{4.12b}$$

$$\mathbf{x} \geq \mathbf{0} \ , \tag{4.12c}$$

with side conditions

$$A\mathbf{e} = \mathbf{0}, \tag{4.12d}$$

$$AA^T \text{ is invertible} . \tag{4.12e}$$

Note that a canonical form problem is always in strict standard form. We define *canonical form constraints* to be constraints satisfying (4.12).

The projective scaling vector field is more naturally associated with a *canonical form fractional linear program*, which is:

$$\text{minimize } \frac{\langle \mathbf{c}, \mathbf{x} \rangle}{\langle \mathbf{b}, \mathbf{x} \rangle} \tag{4.13}$$

subject to

$$A\mathbf{x} = \mathbf{0} , \tag{4.14a}$$

$$\langle \mathbf{e} , \mathbf{x} \rangle = n , \tag{4.14b}$$

$$\mathbf{x} \geq 0 , \tag{4.14c}$$

with side conditions

$$A\mathbf{e} = \mathbf{0} , \tag{4.14d}$$

$$AA^T \text{ is invertible} . \tag{4.14e}$$

where the denominator $\mathbf{b} \geq \mathbf{0}$ is nonnegative, and is scaled so that $\langle \mathbf{b}, \mathbf{e} \rangle = 1$. The condition (4.14d) guarantees that $\mathbf{e}$ is a feasible solution to this fractional linear program.

We define the *fractional projective scaling vector* $\mathbf{v}_{FP}(\mathbf{e} ; \mathbf{c})$ of a canonical form fractional linear program at $\mathbf{e}$ to be the steepest descent direction of the numerator $\langle \mathbf{c}, \mathbf{x} \rangle$ of the fractional linear objective function; subject to the constraints $A\mathbf{x} = 0$ and $\mathbf{e}^T \mathbf{x} = 1$, which is

$$\mathbf{v}_{FP}(\mathbf{e} ; c) = -\pi \begin{bmatrix} A \\ e^T \end{bmatrix}^{\perp} (\mathbf{c}) . \tag{4.15}$$

The fact that this definition does not take into account the denominator $\langle \mathbf{b} , \mathbf{x} \rangle$ of the FLP objective function may seem rather surprising. After defining the projective scaling vector field we will show however that it gives a reasonable search direction for minimizing a *normalized* objective function.

We obtain the projective scaling direction for a canonical form linear program (4.11)-(4.12) by *identifying* it with the fractional linear program having objective function $\dfrac{\langle \mathbf{c}, \mathbf{x} \rangle}{\langle \mathbf{e} , \mathbf{x} \rangle}$. Observe that this FLP objective function is just the LP objective function $\langle \mathbf{c}, \mathbf{x} \rangle$ everywhere on the constraint set in view of the constraint $\langle \mathbf{e},\mathbf{x} \rangle = n$. We define the projective scaling vector $\mathbf{v}_P(\mathbf{e}\ ;\ \mathbf{c})$ to be $\mathbf{v}_P(\mathbf{e}\ ;\ \mathbf{c})$, so that

$$\mathbf{v}_P(\mathbf{e}\ ;\ \mathbf{c}) = -\pi \begin{bmatrix} A \\ \mathbf{e}^T \end{bmatrix}^{\perp} (\mathbf{c}) \ . \tag{4.16}$$

Now we define the projective scaling vector field $\mathbf{v}_{\mathbf{P}}(\mathbf{d}\ ;\ \mathbf{c})$ for a canonical form problem at an arbitrary feasible point $\mathbf{d}$ in Rel-Int$(S_{n-1}) = \{\mathbf{x} : \langle \mathbf{e},\mathbf{x} \rangle = n \text{ and } \mathbf{x} > 0\}$. We define new variables using the projective (scaling) transformation

$$\mathbf{y} = \tilde{\Phi}_D(\mathbf{x}) = n \frac{D^{-1}\mathbf{x}}{\mathbf{e}^T D^{-1}\mathbf{x}} \ . \tag{4.17}$$

with inverse transformation

$$\tilde{\Phi}_D^{-1}(\mathbf{y}) = n\ \frac{D\mathbf{y}}{\mathbf{e}^T D\mathbf{y}} = \mathbf{x} \ . \tag{4.18}$$

Under this change of variables the canonical form fractional linear program (4.13)-(4.14) with objective function $\dfrac{\langle \mathbf{c},\ \mathbf{x} \rangle}{\langle \mathbf{e},\mathbf{x} \rangle}$ becomes the following canonical form fractional linear program.

$$\text{minimize}\ \frac{\langle D\mathbf{c},\ \mathbf{y} \rangle}{\langle D\mathbf{e},\ \mathbf{y} \rangle} \tag{4.19}$$

subject to

$$AD\mathbf{y} = \mathbf{0}\ , \tag{4.20a}$$

$$\langle \mathbf{e},\ \mathbf{y} \rangle = n\ , \tag{4.20b}$$

$$\mathbf{y} \ge \mathbf{0}\ , \tag{4.20c}$$

with side conditions

$$D\mathbf{e} = 0\ , \tag{4.20d}$$

$$AD^2A^T \text{ is invertible}\ . \tag{4.20e}$$

Note that the denominator $\langle D\mathbf{e},\ \mathbf{y} \rangle$ is scaled so that $\langle D\mathbf{e},\ \mathbf{e} \rangle = 1$. Furthermore $\tilde{\Phi}_D(\mathbf{d}) = \mathbf{e}$. By definition the (fractional) projective scaling direction for this point is

$$\mathbf{v}_{FP}(\mathbf{e};\ D\mathbf{c}) = -\ \pi \begin{bmatrix} AD \\ \mathbf{e}^T \end{bmatrix}^{\perp} (D\mathbf{c}) \ . \tag{4.21}$$

We define the projective scaling vector $\mathbf{v}_P(\mathbf{d};\ \mathbf{c})$ to be the pullback under $\tilde{\Phi}_D^{-1}$ of this vector, i.e.

$$\mathbf{v}_P(\mathbf{d}\ ;\ \mathbf{c}) = (\tilde{\Phi}_D^{-1})_*\ (\mathbf{v}_{FP}(\mathbf{e}\ ;\ D\mathbf{c})) \tag{4.22}$$

Now $\tilde{\Phi}_D^{-1}$ is a non-linear map, and a computation gives the formula

$$\left[\tilde{\Phi}_D^{-1} \right]_* (\mathbf{w}) = D\mathbf{w} - \frac{1}{n} \langle D\mathbf{e}, \mathbf{w} \rangle \, D\mathbf{e} \ .$$

The last three formulae combine to yield

$$\mathbf{v}_P(\mathbf{d} \ ; \mathbf{c}) = -D\pi \left[\frac{AD}{e^T} \right]^{\perp} (D\mathbf{c}) + \frac{1}{n} \langle D\mathbf{e}, \ \pi \left[\frac{AD}{\mathbf{e}^T} \right]^{\perp} (D\mathbf{c}) \rangle D\mathbf{e} \ . \tag{4.23}$$

One motivation for this definition of the projective scaling direction is that it gives a ‘‘good’’ direction for fractional linear programs having a normalized objective function. To show this we use observations of Anstreicher [A]. We define a *normalized objective function* of an FLP to be one whose value at the optimum point is zero. This property depends only on the numerator $\langle \mathbf{c}, \ \mathbf{x} \rangle$ of the FLP objective function. The property of being normalized is preserved by the projective change of variable $\mathbf{y} = \tilde{\Phi}_D(\mathbf{x}) = \dfrac{nD^{-1}\mathbf{x}}{\mathbf{e}^T D^{-1} \mathbf{x}}$. In fact the FLP (4.13)-(4.14) is normalized if and only if the transformed FLP (4.19)-(4.20) is normalized. Now consider the FLP (4.13)-(4.14) with an arbitrary objective function. Let $\mathbf{x}^*$ denote the optimal solution vector of a fractional linear program of form (4.13)-(4.14), and let $z^* = \dfrac{\langle \mathbf{c}, \ \mathbf{x}^* \rangle}{\langle \mathbf{b}, \ \mathbf{x}^* \rangle}$ be the optimal objective function value. Define the auxiliary linear program with objective function

$$\text{minimize } \langle \mathbf{c}, \ \mathbf{x} \rangle - z^* \langle \mathbf{b}, \mathbf{x} \rangle \ .$$

and the same constraints (4.14) as the FLP. The point $\mathbf{x}^*$ is easily checked to be an optimal solution of this auxiliary linear program, using the fact that $\dfrac{\langle \mathbf{c}, \ \mathbf{x} \rangle}{\langle \mathbf{b}, \mathbf{x} \rangle} \geq z^*$ for all feasible $\mathbf{x}$. In the special case that $z^* = 0$ which arises from a normalized FLP, the steepest descent direction for this auxiliary linear program is just the fractional projective scaling direction (4.15). Since normalization is preserved under projective transformation $\mathbf{y} = \tilde{\Phi}_D(\mathbf{x})$ this leads to the definition (4.23) of the projective scaling direction $\mathbf{v}_P(\mathbf{d} \ ; \mathbf{c})$ for a canonical form linear program with a normalized objective function.

This discussion provides no justification for the claim that the projective scaling direction $\mathbf{v}_P(\mathbf{d} \ ; \mathbf{c})$ given by (4.15) is an interesting search direction for minimizing a general objective function. In fact the direction specified by $\mathbf{v}_P(\mathbf{d} \ ; \mathbf{c})$ in the general case does have a reasonable consequence, as follows: it leads to the simple relationship between affine scaling trajectories and projective scaling trajectories given in Theorem 6.1.

Now we obtain a simplified formula for the projective scaling direction $\mathbf{v}_P(\mathbf{d}\,;\,\mathbf{c})$, and also show that it depends only on the component $\pi_{A^\perp}(\mathbf{c})$ of $\mathbf{c}$ in the $A^\perp$ direction. We summarize the facts in the following Lemma.

Lemma 4.2. The projective scaling vector field for a canonical form linear program (4.11)-(4.12) is given by

$$v_P(\mathbf{d}\,;\,\mathbf{c}) = -D\pi_{(AD)^\perp}(D\mathbf{c}) + \frac{1}{n}\langle D\mathbf{e},\, \pi_{(AD)^\perp}(D\mathbf{c})\rangle D\mathbf{e}\ . \tag{4.24}$$

In addition

$$\mathbf{v}_P(\mathbf{d}\,;\,\mathbf{c}) = \mathbf{v}_P(\mathbf{d};\, \pi_{A^\perp}(\mathbf{c}))\ . \tag{4.25}$$

Before giving the proof we remark that $\mathbf{v}_P(\mathbf{d}\,;\,\mathbf{c}) \neq \mathbf{v}_P(\mathbf{d}\,;\,\pi_{\left[\frac{A}{e}\right]^\perp}(\mathbf{c}))$ in general.

Proof. By construction $\mathbf{v}_P(\mathbf{d}\,;\,\mathbf{c})$ lies in $\left[\frac{A}{\mathbf{e}^T}\right]^\perp$. To see that $\mathbf{v}_P(\mathbf{d}\,;\,\mathbf{c})$ lies in $(\mathbf{e}^T)^\perp$, we compute by (4.23) that

$$\begin{aligned}\langle \mathbf{e}\,,\, \mathbf{v}_P(\mathbf{d}\,;\,\mathbf{c})\rangle &= \langle D\mathbf{e}\,, \pi_{\left[\frac{AD}{\mathbf{e}^T}\right]^\perp}(D\mathbf{c})\rangle + \frac{1}{n}\langle D\mathbf{e},\, \pi_{\left[\frac{AD}{\mathbf{e}^T}\right]^\perp}(D\mathbf{c})\rangle\,\langle D\mathbf{c}\,,\,\mathbf{e}\rangle \\ &= 0\ .\end{aligned}$$

Now we simplify (4.23) by observing that the feasibility of $\mathbf{d}$ gives

$$AD\mathbf{e} = A\mathbf{d} = \mathbf{0}\ .$$

Hence the projections $\pi_{(AD)^\perp}$ and $\pi_{(e^T)^\perp}$ commute with each other and

$$\pi_{\left[\frac{AD}{\mathbf{e}^T}\right]^\perp} = \pi_{(\mathbf{e}^T)^\perp}\pi_{(AD)^\perp}\ .$$

Next we observe that $\pi_{(\mathbf{e}^T)^\perp} = I - \frac{1}{n}J$ where $J = \mathbf{e}\mathbf{e}^T$ is the matrix with all entries one, and that $J\mathbf{w} = \langle\, \mathbf{e}\,,\,\mathbf{w}\rangle\mathbf{e}$ for all vectors $\mathbf{w}$. Applying these facts to (4.23) we obtain

$$v_P(\mathbf{d}\,;\,\mathbf{c}) = -D\pi_{(\mathbf{e}^T)^\perp}(\pi_{(AD)^\perp}(D\mathbf{c})) + \lambda D\mathbf{e}$$

$$= -D\pi_{(AD)^\perp}(D\mathbf{c}) + \frac{1}{n}DJ\pi_{(AD)^\perp}(Dc) + \lambda D\mathbf{e}$$

$$= -D\pi_{(AD)^\perp}(D\mathbf{c}) + \mu D\mathbf{e} \tag{4.26}$$

where λ and μ are scalars and

$$\mu = \frac{1}{n}\,\langle D\mathbf{e}\,,\,\pi_{\begin{bmatrix} AD \\ e^T \end{bmatrix}^\perp}(D\mathbf{c})\rangle + \frac{1}{n}\langle \mathbf{e},\,\pi_{(AD)^\perp}(D\mathbf{c})\rangle\;. \tag{4.27}$$

Multiplying (4.26) by $\mathbf{e}^T$, and using the identity $\langle \mathbf{e}\,,\,\mathbf{v}_P(\mathbf{d}\,;\,\mathbf{c})\rangle = 0$ we derive an alternate expression for μ which is

$$\mu = \frac{1}{n}\,\langle D\mathbf{e}\,,\,\pi_{(AD)^\perp}(D\mathbf{c})\rangle\;,$$

and this proves (4.24).

To prove the remaining formula, start from

$$\pi_A(\mathbf{c}) = A^T(AA^T)^{-1}\,A\mathbf{c} = A^T\boldsymbol{\lambda}\;.$$

where we define $\boldsymbol{\lambda} = (AA^T)^{-1}A\mathbf{c}$. Then

$$\pi_{AD^\perp}(D\pi_A(\mathbf{c})) = -(I - DA^T(AD^2A^T)^{-1}AD)DA^T\boldsymbol{\lambda}$$

$$= \mathbf{0}\;.$$

Substituting this in (4.24) yields

$$\mathbf{v}_P(d;\,\pi_{A^T}(\mathbf{c})) = \mathbf{0}\;.$$

Since $\mathbf{c} = \pi_{A^\perp}(\mathbf{c}) + \pi_A(\mathbf{c})$ the formula (4.25) follows. ■

The projective scaling vector field $v_P(\mathbf{d}\,;\,\mathbf{c})$ depends on the component of $\mathbf{c}$ in the $\mathbf{e}$-direction. The requirement in Karmarkar's algorithm that objective function be *normalized* so that $\langle \mathbf{c}\,,\,\mathbf{x}_{\text{opt}}\rangle = \mathbf{0}$ specifies the component of $\mathbf{c}$ in the $\mathbf{e}$-direction and removes this ambiguity.

Lemma 4.3 *Given a canonical form linear program* $\boldsymbol{H}$ *and an objective function* $\mathbf{c}$ *there is a unique normalized objective function* $\mathbf{c}_N$ *such that*

(i) $\mathbf{c}_N$ *lies in* $A^\perp$.

(ii) $\pi_{\begin{bmatrix} A \\ e^T \end{bmatrix}^\perp}(\mathbf{c}) = \pi_{\begin{bmatrix} A \\ e^T \end{bmatrix}}(\mathbf{c}_N) = \pi_{e^T}(\mathbf{c}_N)$.

If $\mathbf{c}' = \pi_{\begin{bmatrix} A \\ e^T \end{bmatrix}}(\mathbf{c})$ *then* c_N *is given by*

$$\mathbf{c}_N = \mathbf{c}' - \frac{1}{n} \langle \mathbf{c}', \mathbf{x}_{\mathrm{opt}} \rangle \, \mathbf{e} \tag{4.28}$$

Proof. The condition $A\mathbf{e} = \mathbf{0}$ implies that $A^{\perp} = \begin{bmatrix} A \\ e^T \end{bmatrix}^{\perp} \oplus \mathbb{R} \langle \mathbf{e} \rangle$. Hence the conditions (i) and (ii) imply that any normalized objective function satisfying (i) and (ii) has

$$\mathbf{c}_N = \mathbf{c}' + \mu \mathbf{e} \ .$$

for some scalar μ. The normalization condition gives

$$\langle \mathbf{c}_N , \mathbf{x}_{\mathrm{opt}} \rangle = \langle \mathbf{c}', \mathbf{x}_{\mathrm{opt}} \rangle - \mu \langle \mathbf{e}, \mathbf{x}_{\mathrm{opt}} \rangle = 0 \ .$$

Since a canonical form problem has $\langle \mathbf{e}, \mathbf{x} \rangle = n$ we have $\langle \mathbf{e}, \mathbf{x}_{\mathrm{opt}} \rangle = n$ so

$$\mu = \frac{1}{n} \langle \mathbf{c}', \mathbf{x}_{\mathrm{opt}} \rangle \ .$$

is unique. ■

C. Affine and Projective Scaling Differential Equations

The affine and projective scaling trajectories are found by integrating the affine and projective scaling vector fields, respectively. Now we give definitions.

For the affine scaling case, consider a strict strict standard form problem

$$\text{minimize } \langle \mathbf{c} , \mathbf{x} \rangle$$

subject to

$$A\mathbf{x} = \mathbf{b}$$

$$\mathbf{x} \ge \mathbf{0}$$

having a feasible solution $x = (x_1, \ldots , x_n)$ *with* $x_i > 0$. *In that case the relative interior Rel-Int (***P***) of the polytope* **P** *of feasible solutions is*

$$\mathrm{Rel-Int}(\mathbf{P}) = \{ \mathbf{x} : A\mathbf{x} = \mathbf{b} \text{ and } \mathbf{x} > 0 \} \ . \tag{4.29}$$

Suppose that $\mathbf{x}_0$ *is in Rel-Int(***P***). We define the A-trajectory* $T_A(\mathbf{x}_0 \, ; A, \mathbf{b}, \ \mathbf{c})$ *containing* $\mathbf{x}_0$ to be the point set given by the integral curve $\mathbf{x}(\mathbf{t})$ of the *affine scaling differential equation:*

$$\frac{d\mathbf{x}}{dt} = -X \, \pi_{(AX)^{\perp}} (X\mathbf{c}) \ , \tag{4.30a}$$

$$\mathbf{x}(0) = \mathbf{x}_0 \ , \tag{4.30b}$$

in which $X = X(t)$ is the diagonal matrix with diagonal elements $x_1 \ (t), \ldots , x_n(t)$, so that $\mathbf{x}(t) = X(t)\mathbf{e}$.

This differential equation is obtained from the affine scaling vector field as defined in Lemma 4.1, together with the initial value $\mathbf{x}_0$. The integral curve $\mathbf{x}(t)$ is defined for a range $t_1(\mathbf{x}_0\ ; A,\, \mathbf{c}) < t < t_2(\mathbf{x}_0\ ; A,\, \mathbf{c})$ which is chosen to be the maximum interval on which the solution exists. (Here $t_1 = -\infty$ and $t_2 = +\infty$ are allowable values. It turns out that finite values of t_1 or t_2 may occur, cf. see equation (5.13).) An A-trajectory $T(\mathbf{x}_0 ; A,\, \mathbf{b},\, \mathbf{c})$ lies in Rel-Int(**P**) because the vector field in (4.30) is defined only for $\mathbf{x}(t)$ in Rel-Int(**P**).

For the projective scaling case, consider a canonical form problem (4.11)-(4.12). In this case

$$\text{Rel-Int}(\mathbf{P}) = \{ \mathbf{x} : A\mathbf{x} = \mathbf{0},\ \mathbf{e}^T\mathbf{x} = 1 \text{ and } \mathbf{x} > \mathbf{0} \} .$$

Suppose that $\mathbf{x}_0$ is in Rel-Int(**P**). We define the *P-trajectory* $T_P \left[\mathbf{x}_0\ ; A\ ,\ \ \mathbf{c} \right]$ *containing* $\mathbf{x}_0$ to be the point set given by the integral curve $\mathbf{x}(t)$ of the *projective scaling differential equation:*

$$\frac{d\mathbf{x}}{dt} = -X\ \pi_{(AX)^\perp}\ (X\mathbf{c}) + \langle X\mathbf{e}\ ,\ \pi_{(AX)^\perp}(X\mathbf{c}) \rangle\ X\mathbf{e} \tag{4.31a}$$

$$\mathbf{x}(0) = \mathbf{x}_0\ . \tag{4.31b}$$

This differential equation is obtained from the projective scaling vector field as defined in Lemma 4.2, together with the initial value $\mathbf{x}_0$.

We have defined the A-trajectories and P-trajectories as point sets. The solutions to the differential equations (4.30) and (4.31) specify these point sets as parametrized curves. An arbitrary scaling of the vector fields by an everywhere positive function $\rho(\mathbf{x},\, t)$ leads to a differential equation whose solution will give the same trajectories with a different parametrization. Conversely, a reparametrization of the curve by a variable $u = \psi(t)$ with $\psi'(t) > 0$ for all t leads to a similar differential equation with a rescaled vector field with $\rho(\mathbf{x},\, t) = \psi'(t)$. If $\mathbf{y}(t) = \mathbf{x}(\psi(t))$ and $\mathbf{y}(\mathbf{0}) = \mathbf{x}_\mathbf{0}$ and $\mathbf{x}(t)$ satisfies the affine scaling differential equation then $\mathbf{y}(t)$ satisfies:

$$\frac{d\mathbf{y}}{dt} = -\psi'(t)\ \ Y\ \pi_{(AY)^\perp}(Y\mathbf{c}) \tag{4.32a}$$

$$\mathbf{y}(0) = \mathbf{x}_0\ .$$

If $\mathbf{x}(t)$ satisfies the projective scaling differential equation instead then $\mathbf{y}(t)$ satisfies:

$$\frac{d\mathbf{y}}{d\mathbf{t}} = -\psi'(t)\ [Y\ \pi_{(AY)^\perp}(Y\mathbf{c}) - \langle Y\mathbf{e}\ ,\ \pi_{(AY)^\perp}(Y\mathbf{c})\rangle Y\mathbf{e}] \tag{4.33a}$$

$$\mathbf{y}(0) = \mathbf{x}_0\ . \tag{4.33b}$$

In part III we will give explicit parametrized forms for the A-trajectories and P-trajectories which allow us to characterize their geometric behavior.

5. The affine scaling vector field as a steepest descent vector field

In this section we present Karmarkar's observation that the affine scaling vector field of a strict standard form linear program is a steepest descent vector field of the objective function $\langle \mathbf{c},\ \mathbf{x}\rangle$ with respect to a particular Riemannian metric ds^2 defined on the relative interior of the polytope of feasible solutions of the linear program.

We first review the definition of steepest descent with respect to a Riemannian metric. Let

$$ds^2 = \sum_{i=1}^{n} \sum_{j=1}^{n} g_{ij}(x)\, dx_i\, dx_j \tag{5.1}$$

be a Riemannian metric defined on an open subset Ω of $\mathbb{R}^n$, i.e. we require that the matrix

$$G(\mathbf{x}) = [g_{ij}(\mathbf{x})] \tag{5.2}$$

be a positive-definite symmetric matrix on Ω. Let

$$f(\mathbf{x}) : \Omega \to \mathbb{R} \tag{5.3}$$

be a differentiable function. The *differential* $df_{\mathbf{x}}$ at $\mathbf{x}$ is a linear map on the tangent space $\mathbb{R}^n$ at $\mathbf{x}$,

$$df_{\mathbf{x}} : (\mathbb{R}^n) \to \mathbb{R} \tag{5.4}$$

given by

$$f(\mathbf{x} + \varepsilon\mathbf{v}) = f(\mathbf{x}) + \varepsilon\, df_{\mathbf{x}}(\mathbf{v}) + O(\varepsilon^2) \tag{5.5}$$

as $\varepsilon \to 0$ and $\mathbf{v} \in \mathbb{R}^n$. The Riemannian metric ds^2 permits us to define the gradient vector field $\nabla_G f : \Omega \to \mathbb{R}^n$ with respect to $G(\mathbf{x})$ by $\nabla_G f(\mathbf{x})$ is that direction such that f increases most steeply with respect to ds^2 at $\mathbf{x}$. This is the direction of the minimum of $f(\mathbf{x})$ on an infinitesimal unit ball of ds^2 (which is an ellipsoid) centered at $\mathbf{x}$. Formally

$$\nabla_G f(\mathbf{x}) = G(\mathbf{x})^{-1} \begin{bmatrix} df_{\mathbf{x}} \left[\dfrac{\partial}{\partial x_1} \right] \\ \vdots \\ df_{\mathbf{x}} \left[\dfrac{\partial}{\partial x_n} \right] \end{bmatrix} . \tag{5.6}$$

Note that if ds^2 is the Euclidean metric

$$ds^2 = \sum_{i=1}^{n} dx_i dx_i ,$$

then $\nabla_G f$ is the usual gradient ∇f. (See [Fl], p. 43.)

There is an analogous definition for the gradient vector field $\nabla_G f|_F$ of a function f restricted to a k-dimensional flat F in $\mathbb{R}^n$. Let the flat F be $\mathbf{x}_0 + V$ where V is a n-k-dimensional subspace of $\mathbb{R}^n$ given by

$$V = \{ \mathbf{x} : A\mathbf{x} = \mathbf{0} \} ,$$

in which A is a $k \times n$ matrix of full row rank k. Geometrically the steepest descent direction $\nabla_G f(x_0)|_F$ is that direction in F that maximizes $f(\mathbf{x})$ on an infinitesimal unit ball centered at $\mathbf{x}_0$ of the metric $ds^2|_F$ restricted to F. A computation with Lagrange multipliers given in Appendix A shows that

$$\nabla_G f(\mathbf{x}_0)|_F = (G^{-1} - G^{-1}A^T(AG^{-1}A^T)^{-1}AG^{-1}) \begin{bmatrix} Df_{\mathbf{x}_0} \left[\dfrac{\partial}{\partial x_1} \right] \\ \vdots \\ Df_{\mathbf{x}_0} \left[\dfrac{\partial}{\partial x_n} \right] \end{bmatrix} . \tag{5.7}$$

where ds^2 has coefficient matrix $G = G(\mathbf{x}_0)$ at $\mathbf{x}_0$.

Now we consider a linear programming problem given in strict standard form:

$$\text{minimize } \langle \mathbf{c}, \mathbf{x} \rangle \tag{5.8}$$

subject to

$$A\mathbf{x} = \mathbf{b} \tag{5.8a}$$

$$\mathbf{x} \geq \mathbf{0} \tag{5.8b}$$

with side conditions

$$AA^T \text{ is nonsingular .} \tag{5.8c}$$

having a feasible solution $\mathbf{x}$ with all $x_i > 0$. Karmarkar's steepest descent interpretation of the affine scaling vector field is as follows.

Theorem 5.1. (Karmarkar) *The affine scaling vector field* $\mathbf{v}_A(\mathbf{c}\,;\,\mathbf{d})$ *of a strict standard form problem (5.8) is the steepest descent vector* $-\nabla_G(\langle c\,,\,\mathbf{x}_0\rangle)|_F$ *at* $\mathbf{x}_0 = \mathbf{d}$ *with respect to the Riemannian metric obtained by restricting the metric*

$$ds^2 = \sum_{i=1}^{n} \frac{dx_i\, dx_i}{x_i^2} \tag{5.9}$$

defined on $\mathrm{Int}(\mathbb{R}_+^n)$ *to the flat* $F = \{\mathbf{x} : A\mathbf{x} = \mathbf{b}\}$.

Before proving this result (which is a simple computation) we discuss the metric (5.9). This is a very special Riemannian metric. It may be characterized as the unique Riemannian metric (up to a positive constant factor) on $\mathrm{Int}(\mathbb{R}_+^n)$ which is invariant under the *scaling transformations* $\Phi_D : \mathbb{R}_+^n \to \mathbb{R}_+^n$ given by

$$x_i \to d_i x_i \quad \text{for} \quad 1 \le i \le n\,, \tag{5.10}$$

with all $d_i > 0$ and $D = \mathrm{diag}\,(d_1, \dots, d_n)$, and under the *inverting transformations*

$$I_i((x_1, \dots, x_i, \dots, x_n)) = (x_1, \dots, \frac{1}{x_i}, \dots, x_n) \tag{5.11}$$

for $1 \le i \le n$ and under all permutations $\sigma((x_1, \dots, x_n)) = (x_{\sigma(1)}, \dots, x_{\sigma(n)})$. The geometry induced by ds^2 on $\mathrm{Int}(\mathbb{R}_+^n)$ is isometric to Euclidean geometry on $\mathbb{R}^n$ under the change of variables $y_i = \log x_i$ for $1 \le i \le n$. All these facts are proved in Appendix B.

Proof of Theorem 5.1. The metric $ds^2 = \sum_{i=1}^{n} \frac{(dx_i)^2}{x_i^2}$ induces a unique Riemannian metric $ds^2|_F$ on the region

$$\mathrm{Rel-Int}(\mathbf{P}) = \{\mathbf{x} : A\mathbf{x} = \mathbf{b} \;\text{ and }\; \mathbf{x} > 0\}\,.$$

inside the flat $F = \{\mathbf{x} : A\mathbf{x} = \mathbf{b}\}$. The matrix $\overline{G}(\mathbf{x})$ associated to ds^2 is the diagonal matrix

$$\overline{G}(x) = \mathrm{diag}\left[\frac{1}{x_1^2}, \dots, \frac{1}{x_n^2}\right] = X^{-2}\,,$$

where $X = \mathrm{diag}(x_1, \dots, x_n)$. Using the definition (5.7) applied to the function $\Rightarrow_c(\mathbf{x}) = \langle \mathbf{c}, \mathbf{x}\rangle$. we obtain

$$\nabla_{\overline{G}}(\Rightarrow_c(\mathbf{x}))|_F = X(I - XA(AX^2A^T)^{-1}AX)X\mathbf{c}\,.$$

The right side of this equation is $-\mathbf{v}_A(\mathbf{x}\,;\,\mathbf{c})$ by Lemma 3.1. ■

We now show by an example that these steepest descent curves are not geodesics of the metric $ds^2|_F$ even in the simplest case. Consider the strict standard form problem with no equality constraints:

$$\text{minimize } \langle \mathbf{c},\, \mathbf{x} \rangle$$

subject to

$$\mathbf{x} \geq \mathbf{0} \ .$$

The affine scaling differential equation (4.30) becomes in this case

$$\frac{d\mathbf{x}}{dt} = -X^2 \mathbf{c} \tag{5.12a}$$

$$\mathbf{x}(0) = (d_1, \ldots, d_n) \ . \tag{5.12b}$$

This is a decoupled set of Riccati equations

$$\frac{dx_i}{dt} = -x_i^2 c_i \ ,$$

$$x_i(0) = d_i \ .$$

for $1 \leq i \leq n$. Using the change of variables $y_i = \dfrac{1}{x_i}$ we easily find that

$$\frac{dy_i}{dt} = c_i \ ,$$

$$y_i(0) = \frac{1}{d_i} \ ,$$

for $1 \leq i \leq n$. From this we obtain

$$\mathbf{x}(t) = \left[\frac{1}{\frac{1}{d_1} + c_1 t} \ , \ \cdots \ , \ \frac{1}{\frac{1}{d_n} + c_n t} \right] . \tag{5.13}$$

This trajectory is defined for $t_1 < t < t_2$ where

$$t_1 = \max\left\{ \frac{c_i}{d_i} : c_i > 0 \right\} \tag{5.14a}$$

$$t_2 = \min\left\{ \frac{c_i}{d_i} : c_i < 0 \right\} \tag{5.14b}$$

with the convention that $t_1 = -\infty$ if all $c_i \leq 0$ and $t_2 = \infty$ if all $c_i \geq 0$. The geodesic curves of

$$ds^2 = \sum_{i=1}^{n} \frac{dx_i \, dx_i}{x_i^2}$$

are explicitly evaluated in Appendix B to be

$$\boldsymbol{\gamma}(t) = (e^{a_1 t + b_1}, \ldots, e^{a_n t + b_n}) ,$$

where $\sum_{c=1}^{n} a_i^2 = 1$ for $-\infty < t < \infty$. It is easy to see these do not coincide with the curves (5.13) for $n \geq 2$, since $\mathbf{x}(t)$ is a rational curve while $\boldsymbol{\gamma}(t)$ satisfies no algebraic dependencies among its coordinates in general.

6. Relations between *P*-trajectories and *A*-trajectories

There is a simple relationship between the P-trajectories of the canonical form linear program:

$$\text{minimize } \langle \mathbf{c}, \mathbf{x} \rangle \tag{6.1a}$$

subject to

$$A\mathbf{x} = \mathbf{0} \tag{6.1b}$$

$$\langle \mathbf{e}, \mathbf{x} \rangle = n \tag{6.1c}$$

$$\mathbf{x} \geq \mathbf{0} \tag{6.1d}$$

with side conditions

$$A\mathbf{e} = \mathbf{0} \tag{6.1e}$$

$$AA^T \text{ is invertible .} \tag{6.1f}$$

and the A-trajectories of the associated homogeneous strict standard form linear program:

$$\text{minimize } \langle \mathbf{c}, \mathbf{x} \rangle \tag{6.2a}$$

subject to

$$A\mathbf{x} = \mathbf{0} \tag{6.2b}$$

$$\mathbf{x} \geq \mathbf{0} \tag{6.2c}$$

with side conditions

$$A\mathbf{e} = \mathbf{0} \tag{6.2d}$$

$$AA^T \text{ is invertible .} \tag{6.2e}$$

It is as follows.

Theorem 6.1. If $T_A(\mathbf{x}_0\ ; A, \mathbf{c})$ is an A-trajectory of the homogeneous strict standard form problem (6.2) then its radial projection

$$T = \left\{ \frac{n\mathbf{x}}{\langle \mathbf{e}\ , \mathbf{x} \rangle} : \mathbf{x} \in\ T_A(\mathbf{x}_0\ ; A, \mathbf{0}\ ,\ \mathbf{c}) \right\} \tag{6.3}$$

is a P-trajectory of the associated canonical form linear program, which is given by

$$T = T_P \left[\frac{n\mathbf{x}_0}{\langle \mathbf{e}\ , \mathbf{x}_0 \rangle}\ ; [A]\ ,\ \mathbf{c} \right] . \tag{6.4}$$

Proof. Geometrically the radial projection produces the radial component in the projective scaling vector field evident on comparing Lemmas 4.1 and 4.2. The trajectory $T_A(\mathbf{x}_0\ ; A, \mathbf{0}, \mathbf{c})$ is parametrized by a solution $\mathbf{x}(t)$ of the differential equation

$$\frac{d\mathbf{x}}{dt} = -X\pi_{(AX)^\perp}(X\mathbf{c}) . \tag{6.5}$$

$$\mathbf{x}(0) = \mathbf{x}_0 .$$

Now define

$$\mathbf{y}(t) = \frac{n\mathbf{x}(t)}{\langle \mathbf{e}\ , \mathbf{x}(t) \rangle} .$$

We verify directly that $\mathbf{y}(t)$ satisfies a (scaled) version of the projective scaling differential equation.

Let $Y(t) = \text{diag}\ (y_1(t)\ , \ldots\ , y_n(t))$ and note that $Y(t) = n\langle e\ , \mathbf{x}(t) \rangle^{-1} X(t)$ so that

$$X\pi_{(AX)^\perp}(X\mathbf{c}) = n^{-2} \langle \mathbf{e}, \mathbf{x}(t) \rangle^2\ Y\pi_{(AY)^\perp}(Y\mathbf{c}) .$$

Using this fact and $Y\mathbf{e} = n\langle \mathbf{e}\ , \mathbf{x}(t) \rangle^{-1}\mathbf{x}$ we obtain

$$\frac{d\mathbf{y}}{dt} = n\langle \mathbf{e}, \mathbf{x}(t) \rangle^{-1}\ \frac{d\mathbf{x}}{dt} - n^2 \langle \mathbf{e}, \mathbf{x}(t) \rangle^{-2} \langle \mathbf{e}\ , \frac{d\mathbf{x}}{dt} \rangle\ \mathbf{x}$$

$$= -n\langle \mathbf{e}\ , \mathbf{x}(t) \rangle^{-1}\ (n^{-2} \langle \mathbf{e}, \mathbf{x}(t) \rangle^2\ Y\pi_{(AY)^\perp}\ (Y\mathbf{c}) - n^{-3} \langle \mathbf{e}, \mathbf{x}(t) \rangle^2 \langle \mathbf{e}, Y\pi_{(AY)^\perp}\ (Y\mathbf{c} \rangle Y\mathbf{e})$$

$$= \frac{1}{n} \langle \mathbf{e}, \mathbf{x}(t) \rangle\ (Y\pi_{(AY)^\perp}\ (Y\mathbf{c}) - \langle Y\mathbf{e}, \pi_{(AY)^\perp}(Y\mathbf{c}) \rangle Y\mathbf{e}$$

$$= \frac{1}{n} \langle \mathbf{e}\ , \mathbf{x}(t) \rangle\ \mathbf{v_P}(\mathbf{y}\ ; \mathbf{c}) .$$

Since $\psi'(t; \mathbf{x}_0) = \frac{1}{n}\langle \mathbf{e}\ , \mathbf{x}(t) \rangle > 0$ since $\mathbf{x}(t) \in \text{Int}(\mathbb{R}^n_+)$ this is a version of the projective scaling differential equation (4.33). This proves (6.4) holds. ■

As an example we apply Theorem 6.4 to the canonical form linear program with no extra equality constraints:

$$\text{minimize } \langle \mathbf{c},\ \mathbf{x} \rangle$$

subject to

$$\mathbf{e}^T \mathbf{x} = n ,$$

$$\mathbf{x} \geq 0 .$$

The feasible solutions to this problem form a regular simplex S_{n-1}. In this case the associated homogeneous standard form problem has no equality constraints:

$$\text{minimize } \langle \mathbf{c},\ \mathbf{x} \rangle$$

subject to

$$\mathbf{x} \geq \mathbf{0} .$$

Using the formula (5.13) parametrizing the affine scaling trajectories:

$$T_A(\mathbf{d}\ ;\ \phi,\ \phi,\ \mathbf{c}) = \left\{ \left[\frac{1}{\dfrac{1}{d_1} + c_1 t} , \ldots , \frac{1}{\dfrac{1}{d_n} + c_n t} \right] : t_1 < t < t_2 \right\}$$

we find that if $\mathbf{d}$ lies in $\mathrm{Int}(S_n)$ then the projective scaling trajectory given by Theorem 6.4 is

$$T_P(\mathbf{d}\ ;\ \phi\ ,\ \mathbf{c}) = \left\{ \frac{n}{\sum_{i=1}^{n} \left[\dfrac{1}{d_i} + c_i t \right]^{-1}} \left[\frac{1}{\dfrac{1}{d_1} + c_1 t} , \ldots , \frac{1}{\dfrac{1}{d_n} + c_n t} \right] : t_1 < t < t_2 \right\} .$$

where t_1 and t_2 are given by (5.14). Notice that both $T_A(\mathbf{d}\ ;\ \phi,\ \phi,\ \mathbf{c})$ and $T_{\mathbf{P}}(\mathbf{d}\ ;\ \phi,\ \phi\ ,\ \mathbf{c})$ are rational curves in this example.

Since any canonical form problem is automatically a standard form problem, both P-trajectories and A-trajectories are defined for a canonical form problem. In general an A-trajectory is not a P-trajectory and vice-versa. However the A-trajectories and P-trajectories through the point $\mathbf{e}$ do coincide, and we have the relation:

$$T_P\left[\mathbf{e}\,;\,[A]\,,\,\mathbf{c}\right] \;=\; T_A\left[\mathbf{e}\,;\,\begin{bmatrix} A \\ e^T \end{bmatrix},\;\begin{bmatrix} \mathbf{0} \\ 1 \end{bmatrix},\;\mathbf{c}\right]\;. \tag{6.6}$$

This is proved in [BL3]. We call the point **e** the *center* (as does Karmarkar) and we call the trajectories (6.6) *central trajectories*.

7. The homogeneous affine scaling algorithm

Consider the homogeneous standard form linear program:

$$\text{minimize } \langle \mathbf{c},\, \mathbf{x} \rangle \tag{7.1a}$$

subject to

$$A\mathbf{x} \;=\; \mathbf{0}\;, \tag{7.1b}$$

$$\mathbf{x} \;\geq\; \mathbf{0}\;, \tag{7.1c}$$

with side conditions

$$A\mathbf{e} \;=\; \mathbf{0}\;, \tag{7.1d}$$

$$AA^T \text{ is invertible }. \tag{7.1e}$$

We define the *homogeneous affine scaling algorithm* to be a piecewise linear algorithm in which the *starting value* is given by $\mathbf{x}_0 \;=\; \mathbf{e}$, the *step direction* is specified by the affine scaling vector field associated to (7.1) and the *step size* is chosen to minimize Karmarkar's ‘‘potential function’’

$$g(\mathbf{x}) \;=\; \sum_{i=1}^{n} \log \left[\frac{\langle \mathbf{c},\, \mathbf{x} \rangle}{x_i} \right] \tag{7.2}$$

along the line segment inside the feasible solution polytope specified by the step direction. Let $\mathbf{x}_0, \ldots, \mathbf{x}_n$ denote the resulting sequence of interior points obtained using this algorithm. Consider the associated canonical form problem:

$$\text{minimize } \langle \mathbf{c},\, \mathbf{x} \rangle \tag{7.3a}$$

subject to

$$A\mathbf{x} \;=\; \mathbf{0}\;, \tag{7.3b}$$

$$\langle \mathbf{e}\,,\, \mathbf{x} \rangle \;=\; n\;, \tag{7.3c}$$

$$\mathbf{x} \;\geq\; \mathbf{0}\;, \tag{7.3d}$$

with side conditions

$$A\mathbf{e} = 0 \,, \tag{7.3e}$$

$$AA^T \text{ is invertible} \,. \tag{7.3f}$$

We have the following result.

Theorem 7.1. If $\{\mathbf{x}^{(k)} : 0 \le k < \infty\}$ *are the homogeneous affine scaling algorithm iterates associated to the linear program (7.1) and if* $\mathbf{y}_i^{(k)}$ *are defined by*

$$\mathbf{y}_i = \frac{n\mathbf{x}^{(k)}}{\langle \mathbf{e} \,, \mathbf{x}^{(k)} \rangle} \,, \tag{7.4}$$

then $\{\mathbf{x}^{(k)} : 0 \le k < \infty\}$ *are the projective scaling algorithm iterates of the canonical form problem (7.2).*

Proof. We observe that Karmarkar's ''potential function'' is constant on rays through the origin:

$$g(\lambda\mathbf{x}) = g(\mathbf{x}) \quad \text{if} \quad \lambda > 0 \,. \tag{7.5}$$

Now we prove the theorem by induction on the iteration number *k*. It is true by definition for $k = 0$. If it is true for a given *k*, then the proof of Theorem 6.1 showed that the non-radial component of the affine scaling vector field agrees with the projective scaling vector field. Hence the radial projection of the homogeneous affine scaling step direction line segment inside $\mathbb{R}_+^n$ is the projective scaling step direction line segment inside $\mathbb{R}_+^n$. Since Karmarkar's potential function is constant on rays, the step size criterion for the homogeneous affine scaling algorithm causes (7.4) to hold for $k + 1$, completing the induction step. ■

Theorem 7.1 gives an interpretation of Karmarkar's projective scaling algorithm as a polynomial time linear programming algorithm using an affine scaling vector field. The homogeneous affine scaling algorithm can alternatively be regarded as an algorithm solving the fractional linear program with objective function

$$\text{minimize } \frac{\langle \mathbf{c}, \mathbf{x} \rangle}{\langle \mathbf{e} \,, \mathbf{x} \rangle} \,,$$

subject to the standard form constraints (7.1b)-(7.1e). If Karmarkar's stopping rule is used one obtains a polynomial time algorithm for solving this fractional linear program.

Appendix A. Steepest descent direction with respect to a Riemannian metric

We compute the steepest descent direction $\nabla_G f(\mathbf{x}_0)|_F$ of a function $f(\mathbf{x})$ defined on a flat $F = \mathbf{x}_0 + \{\mathbf{x} : A\mathbf{x} = \mathbf{0}\}$ with respect to a Riemannian metric $ds^2 = \sum_{i=1}^{n} \sum_{j=1}^{n} g_{ij}(x)\ dx_i\, dx_j$ at $\mathbf{x}_0$. We may suppose without loss of generality that $\mathbf{x}_0 = \mathbf{0}$, and set $G = [g_{ij}(\mathbf{0})]$.

The steepest descent direction is found by maximizing the linear functional

$$\langle df_0\ , \mathbf{v} \rangle = \left[df_{\mathbf{0}} \left[\frac{\partial}{\partial x_1} \right] , \ldots , df_{\mathbf{0}} \left[\frac{\partial}{\partial x_n} \right] \right] \mathbf{v} \tag{A.1}$$

on the ellipsoid

$$\sum_{i=1}^{n} \sum_{j=1}^{n} g_{ij} v_i v_j = \varepsilon^2 , \tag{A.2}$$

subject to the constraints

$$A\mathbf{v} = 0 . \tag{A.3}$$

Note that the direction obtained is independent of ε. We define $\mathbf{d} \equiv \left[df_{\mathbf{0}} \left[\frac{\partial}{\partial x_1} \right] , \ldots , df_{\mathbf{0}} \left[\frac{\partial}{\partial x_n} \right] \right]^T$, and set this problem up as a Lagrange multiplier problem. We wish to find a stationary point of

$$L = \mathbf{d}^T \mathbf{v} - \lambda^T A \mathbf{v} - \mu(\mathbf{v}^T G \mathbf{v} - \varepsilon^2) . \tag{A.4}$$

The stationarity conditions are

$$\frac{\partial L}{\partial \mathbf{v}} = \mathbf{d} - A^T \boldsymbol{\lambda} - \mu(G + G^T)\mathbf{v} = \mathbf{0} , \tag{A.5}$$

$$\frac{\partial L}{\partial \boldsymbol{\lambda}} = -A\mathbf{v} = \mathbf{0} , \tag{A.6}$$

$$\frac{\partial L}{\partial \mu} = \mathbf{v}^T G \mathbf{v} = \varepsilon^2 . \tag{A.7}$$

Using (A.5) and $G = G^T$ we find that

$$\mathbf{v} = \frac{1}{2\mu} G^{-1} (\mathbf{d} - A^T \boldsymbol{\lambda}) . \tag{A.8}$$

Substituting this into (A.6) yields

$$A G^{-1} A^T \boldsymbol{\lambda} = A G^{-1} \mathbf{d} .$$

Hence

$$\boldsymbol{\lambda} = (AG^{-1}A^T)^{-1}AG^{-1}\mathbf{d}\ . \tag{A.9}$$

Substituting this into (A.8) yields

$$\mathbf{v} = \frac{1}{2\mu}\ (G^{-1} - G^{-1}A^T(AG^{-1}A^T)^{-1}AG^{-1})\ \mathbf{d}\ . \tag{A.10}$$

Now we check that the tangent vector

$$\mathbf{w} = (G^{-1} - G^{-1}A^T(AG^{-1}A^T)^{-1}AG^{-1})\,\mathbf{d} \tag{A.11}$$

points in the maximizing direction. This corresponds to taking $\mu > 0$ in (A.10). To show this, we show that the linear functional $\mathbf{d}^T\mathbf{v}$ is nonnegative at $\mathbf{v} = \mathbf{w}$. Now recall that any positive definite symmetric matrix M has a unique positive definite symmetric square root $M^{1/2}$. Using this fact on G we have

$$\mathbf{v}^T\mathbf{w} = \mathbf{d}^TG^{-1}\mathbf{d} - \mathbf{d}^TG^{-1}A^T(AG^{-1}A^T)^{-1}AG^{-1}\mathbf{d}$$

$$= (\mathbf{d}G^{-1/2})(I - G^{-1/2}A^T(AG^{-1}A^T)^{-1}AG)^{-1/2}\mathbf{d}$$

Now $\pi_W = I - G^{-1/2}A^T(AG^{-1}A^T)^{-1}AG^{-1/2}$ is a projection operator onto the subspace $W = \{\mathbf{x} : AG^{-1/2}\mathbf{x} = \mathbf{0}\}$. and (A.12) gives

$$\mathbf{d}^T\mathbf{w} = (G^{-1/2}\mathbf{d})^T\ \pi_W(G^{-1/2}\mathbf{d})$$

$$= \|\pi_W(G^{-1/2}\mathbf{d})\|^2 \geq \mathbf{0}\ ,$$

where $\|\cdot\|$denotes the Euclidean norm. There are two degenerate cases where $\mathbf{d}^T\mathbf{w} = 0$. The first is where $\mathbf{d} = \mathbf{0}$, which corresponds to $\mathbf{0}$ being a stationary point of f, and the second is where $\mathbf{d} \neq \mathbf{0}$ but $\mathbf{d}^T\mathbf{w} = \mathbf{0}$, in which case the linear functional $\langle df_{\mathbf{0}}\ , \mathbf{v}\rangle = \mathbf{d}^T\mathbf{v}$ is constant on the flat F.

The vector (A.11) is the gradient vector field with respect to G. We obtain the analogue of a *unit gradient field* by using the Lagrange multiplier μ to scale the length of $\mathbf{v}$. Substituting (A.10) into (A.7) yields

$$4\mu^2\varepsilon = \mathbf{d}^TG^{-1}\mathbf{d} - \mathbf{d}^TG^{-1}A^T(AG^{-1}A^T)^{-1}AG^{-1}\mathbf{d}$$

Hence

$$\mu = \frac{\pm 1}{2\varepsilon}\ (\mathbf{d}^TG^{-1}\mathbf{d} - \mathbf{d}^TG^{-1}A^T(AG^{-1}A^T)^{-1}AG^{-1}\mathbf{d})^{1/2}$$

We obtain

$$\lim_{\varepsilon\to 0}\ \frac{1}{\varepsilon\mathbf{v}} = \theta(G,\ \mathbf{d})(G^{-1} - G^{-1}A^T(AG^{-1}A^T)^{-1}AG^{-1})\,\mathbf{d} \tag{A.12}$$

where $\theta(G,\ \mathbf{d})$ is the scaling factor

$$\theta(G, \mathbf{d}) = (\mathbf{d}^T G^{-1} \mathbf{d} - \mathbf{d}^T G^{-1} A^T (AGA^T)^{-1} AG^{-1} \mathbf{d})^{-1/2}$$

Here $\theta(G, \mathbf{d})$ measures the length of the tangent vector **w** with respect to the metric ds^2. (As a check, note that for the Euclidean metric and $F = \mathbb{R}^n$ the formula (A.11) for **w** gives the ordinary gradient and (A.12) gives the unit gradient.)

Appendix B. Invariant Riemannian metrics on the positive orthant $\mathbb{R}^n_+$.

We consider Riemannian metrics such that

$$ds^2 = \sum_{i=1}^{n} \sum_{j=1}^{n} g_{ij}(\mathbf{x})\, dx_i\, dx_j \tag{B.1}$$

has $g_{ij}(\mathbf{x}) = g_{ji}(\mathbf{x})$ and all functions $g_{ij}(\mathbf{x})$ are defined on the interior Int($\mathbb{R}^n_+$) of the positive orthant $\mathbb{R}^n_+$, i.e., if $\mathbf{x} = (x_1, \dots, x_n)^T$ then

$$Int(\mathbb{R}^n_+) = \{\mathbf{x} : x_i > 0 \ \text{ for } \ 1 \le i \le n\} .$$

Let $D = \operatorname{diag}(d_1, \dots, d_n)$ where all $d_i > 0$, and let

$$\Phi_D(\mathbf{x}) = D\mathbf{x} . \tag{B.2}$$

and let G^n_+ denote the (Lie) group of positive scaling transformations

$$G^n_+ = \{\Phi_D : d_i > 0 \ \text{ for } \ 1 \le i \le n\} . \tag{B.3}$$

Then G^n_+ acts transitively on $\mathbb{R}^n_+$.

Theorem B.1. The Riemannian metrics defined on $Int(\mathbb{R}^n_+)$ that are invariant under G^n_+ are exactly the metrics

$$ds^2 = \sum_{i=1}^{n} \frac{c_{ij}}{x_i x_j}\, dx_i\, dx_j \tag{B.4}$$

where $C = [c_{ij}]$ is a positive definite symmetric matrix.

Proof. To study a general metric on $\mathbb{R}^n_+$ we use the map $L \colon (\mathbb{R}^n_+) \to \mathbb{R}^n$ given by

$$L(\mathbf{x}) = (\log x_1, \dots, \log x_n) ,$$

i.e. the new coordinates are $y_i = \log x_i$. Then (B.1) in the new coordinates is

$$ds^2 = \sum_{i=1}^{n} \sum_{j=1}^{n} \bar{g}_{ij}(\mathbf{y})\, e^{y_1 + y_j}\, dy_i\, dy_j \tag{B.5}$$

where

$$\bar{g}_{ij}(\mathbf{y}) = g_{ij}(e^{y_i}, \dots, e^{y_n}) . \tag{B.6}$$

Under this transformation the group G^n_+ becomes the group T_n of translations on $\mathbb{R}^n$, i.e.

$$L(\Phi_D(\mathbf{x})) = L(\mathbf{x}) + L(D\mathbf{e})$$

where

$$L(D\mathbf{e}) = (\log d_1, \ldots, \log d_n) .$$

Now a translation-invariant Riemannian metric on $|\mathrm{R}^n$ is specified by its infinitesimal unit ball at the origin, i.e. it is a constant metric

$$ds^2 = \sum_{i=1}^{n} \sum_{j=1}^{n} c_{ij} dy_i dy_j , \tag{B.7}$$

where $C = [c_{ij}]$ is a fixed positive definite symmetric matrix. Substituting this in (B.5) yields $\bar{g}_{ij}(y) = c_{ij} e^{-(y_i + y_j)}$, so that by (B.6) we have

$$g_{ij}(\mathbf{x}) = \frac{c_{ij}}{x_i x_j} .$$

Since the metrics (B.4) are all invariant under G_+^n we have proved Theorem B.1. ■

Theorem B.2. The only Riemannian metrics defined on $\mathrm{Int}(|\mathrm{R}_+^n)$ *that are invariant under* G_+^n *and under all inversions*

$$I_k((x_1, \ldots, x_{k-1}, x_k, x_{k+1}, \ldots, x_n)) = (x_1, \ldots, x_{k-1}, \frac{1}{x_k}, x_{k+1}, \ldots, x_n)$$

for $1 \le k \le n$ *is*

$$ds^2 = \sum_{i=1}^{n} c_i \frac{dx_i dx_i}{x_i^2} . \tag{B.8}$$

where all $c_i > 0$. *The only such metrics that are invariant under these transformations and also under all permutations* $\sigma_j(x_1, \ldots, x_n) = (x_{\sigma(1)}, \ldots, x_{\sigma n)})$ *are those of the form*

$$ds^2 = c \sum_{i=1}^{n} \frac{dx_i dx_i}{x_i^2} \tag{B.9}$$

where $c > 0$.

Proof. By Theorem A-1 we may assume that the metric has the form

$$ds^2 = \sum \frac{c_{ij}}{x_i x_j} dx_i dx_j . \tag{B.10}$$

Now let $y_k = \frac{1}{x_k}$, $y_j = x_j$ for $j \ne k$. Then we compute

$$ds^2 = \sum \bar{g}_{ij}(y) dy_i dy_j$$

where

$$\bar{g}_{ij}(\mathbf{y}) = \frac{c_{ij}}{x_i x_j} \frac{\partial x_i}{\partial y_i} \frac{\partial x_j}{\partial y_j} .$$

In particular for $j \neq k$ we have

$$\bar{g}_{kj}(\mathbf{y}) = \frac{c_{jk} y_k}{y_j} \left[-\frac{1}{y_k^2} \right] = -\frac{c_{kj}}{y_j y_k} .$$

By the invariance hypothesis we must have

$$\bar{g}_{kj}(y) = \frac{c_{kj}}{y_j y_k}$$

This implies that

$$c_{ij} = 0 \quad \text{if} \quad i \neq j$$

and (B.8) follows with $c_i = c_{ii}$.

For the second part if $y_i = x_{\sigma(i)}$ then a computation gives

$$ds^2 = \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{c_{\sigma(i),\sigma(j)}}{y_i y_j} dy_i dy_j$$

hence comparison with (B.10) gives

$$c_{ii} = c_{\sigma(i)\,,\,\sigma(i)}$$

for all permutations σ and (B.9) follows. ■

Theorem B.3. The geodesics of

$$ds^2 = \sum_{i=1}^{n} \frac{dx_i dx_i}{x_i^2} \tag{B.11}$$

in $\mathrm{Int}(\mathrm{|R}_+^n)$ *are exactly the curves*

$$\gamma(t) = \gamma_{\mathbf{a},\mathbf{b}}(t) = (e^{a_1 t + b_1}, e^{a_2 t + b_2}, \ldots, e^{a_n t + b_n}), \; -\infty < t < \infty . \tag{B.12}$$

where $\|\mathbf{a}\|^2 = a_1^2 + a_2^2 + \ldots + a_n^2 = 1$.

Proof. The mapping $L : \mathrm{|R}_+^n \to \mathrm{|R}^n$

$$L(\mathbf{x}) = (\log x_1, \ldots, \log x_n) \tag{B.13}$$

takes the metric (B.11) to the Euclidean metric

$$ds^2 = \sum_{c=1}^{n} dy_i dy_i$$

on $\mathrm{|R}^n$. The geodesics of ds^2 are clearly

$$\gamma(t) = (a_1 t + b_1, \ldots, a_n t + b_n)$$

where $\mathbf{a} = (a_1, \ldots, a_n)$ has $\|a_i\|^2 = 1$. The formula (A.12) follows using the inverse map

$$L^{-1}(\mathbf{y}) = (e^{y_1}, \ldots, e^{y_n}) \ . \ \blacksquare$$

The metric $\sum \frac{dx_i \, dx_i}{x_i^2}$ has Gaussian curvature 0 at every point of $\mathbb{R}^2_+$, i.e. it is a *flat metric*. This follows since the transformation (B.12) does not change Gaussian curvature, and the Euclidean metric is flat.

REFERENCES

[A] K. Anstreicher, A monotonic projective algorithm for fractional linear programming, preprint (1985).

[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, preprint (1985).

[BKL] D. Bayer, N. Karmarkar, J. C. Lagarias, paper in preparation.

[BL2] D. Bayer and J. C. Lagarias, The nonlinear geometry of linear programming II. Legendre transform coordinates, preprint.

[BL3] D. Bayer and J. C. Lagarias, The non-linear geometry of linear programming III. Central trajectories, in preparation.

[BM] S. Bochner and W. T. Martin, *Several Complex Variables*, Princeton U. Press, Princeton, New Jersey 1948.

[BP] H. Busemann, and B. B. Phadke, Beltrami's theorem in the large, Pacific J. Math. **115** (1984), 299-315.

[Bu] H. Busemann, *The geometry of geodesics*, Academic Press, New York 1955.

[Bu2] H. Busemann, Spaces with distinguished shortest joints, in: *A Spectrum of Mathematics*, Auckland 1971, 108-120.

[CH] R. Courant and D. Hilbert, *Methods of Mathematical Physics*, Vol. I, II Wiley, New York 1962.

[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*, John Wiley and Sons, New York 1968.

[Fl] W. H. Fleming, *Functions of Several Variables*, Addison-Wesley, Reading, Mass. 1965.

[GZ] C. B. Garcia and W. I. Zangnill, *Pathways to Solutions, Fixed Points and Equilibria*, Prentice-Hall (Englewood Cliffs, N.J.) 1981.

[H] D. Hilbert, *Grundlagen der Geometry*, 7th Ed., Leipzig 1930. (English Translation: *Foundations of Geometry*).

[Ho] J. Hooker, The projective linear programming algorithm, Interfaces, 1986, to appear.

[K] N. Karmarkar, A new polynomial time algorithm for linear programming, Combinatorica *4* (1984) 373-395.

[KV] S. Kapoor and P. M. Vaidya, Fast algorithms for convex quadratic programming and multicommodity flows, Proc. 18th ACM Symp. on Theory of Computing, 1986 (to appear).

[L4] J. C. Lagarias, The non-linear geometry of linear programming IV. Hilbert geometry, in preparation.

[L] J. C. Lagarias, paper in preparation.

[Ln] C. Lanczos, *The Variational Principles of Mechanics*, Univ. of Toronto Press, Toronto 1949.

[M] N. Megiddo, On the complexity of linear programming, in: *Advances in Economic Theory* (1985), (T. Bewley, Ed.), Cambridge Univ. Press, 1986, to appear.

[M2] N. Megiddo, Pathways to the optimal set in linear programming, preprint 1986.

[N] J. L. Nazareth, Karmarkar's method and homotopies with restarts, preprint 1985.

[Re] J. Renegar, paper in preparation.

[R1] R. T. Rockafellar, Conjugates and Legendre Transforms of Convex Functions, Canad. J. Math. *19* (1967) 200-205.

[R2] R. T. Rockafellar, *Convex Analysis*, Princeton U. Press, 1970.

[Sh] M. Shub, paper in preparation.

[SW] J. Stoer and C. Witzgall, *Convexity and Optimization in Finite Dimensions* I, Springer-Verlag, New York 1970.

[VMF] R. J. Vanderbei, M. J. Meketon and B. A. Freedman, A modification of Karmarkar’s linear programming algorithm, preprint (1985).