DYNAMICAL SYSTEM CHARACTERIZATION OF THE CENTRAL PATH AND ITS VARIANTS — A REVISIT

MOODY T. CHU* AND MATTHEW M. LIN†

Abstract. The notion of central path plays a fundamental role in the development of interior point methods which, in turn, have become important tools for solving various optimization problems. The central path equation is algebraic in nature and is derived from the KKT optimality conditions of a certain logarithmic barrier problem; meanwhile, the primal variable portion of the very same central path can also be cast precisely as the integral curve, known as the affine scaling trajectory, of a certain gradient-type dynamical system. The justification is easy to establish in the context of linear programming. Though expected, the generalization of such a concept to the semi-definite programming is not quite as obvious due to the difficulty of addressing non-commutativity in matrix multiplication. This paper revisits the dynamical system characterization of these flows and addresses the needed details for extension to semi-definite programming by means of a simple notion of operators and specially defined inner product. From a dynamical system point of view, numerical ODE techniques might help to develop some novel interior point methods.

Key words. linear programming, semi-definite programming, central path, interior point method, projected gradient, dynamical system, affine scaling trajectory

AMS subject classifications. 37B35, 37N40, 90C22, 90C51

1. Introduction.

Under the assumption of strong duality, the central path in either the linear or the semi-definite programming paradigm is typically characterized by a coupled algebraic system, called the central path equation, which arises from the optimality conditions for a certain family of parameterized logarithmic barrier problems. The essence in the now well studied and prevalent interior point methods is to approximately track points on the central path as the parameter is decreased to zero. The literature on interior point methods, both in theory and practice, is abundant. We mention, far from being exhaustive, only a few classic reference books [3, 7, 14, 19, 20] in this area.

We envisage the idea that if the central path itself can be characterized as the integral curve of a certain differential system, then the notion of path-following and convergence analysis might be ruminated from the numerical ODE point of view. To save the computational overhead, of course, it would be particularly desirable to describe the curves for the primal and the dual variables without making reference to each other. The purpose of this paper is to revisit this notion in the context of linear programming and to extend the idea to semi-definite programming. The latter task has not been obvious thus far because of the difficulty in addressing the non-commutative nature of matrix operations.

We intend to bring forward two facts in this presentation. Firstly, we argue that, with a specific way of diminishing the barrier parameter to zero, the primal variable portion in the central path is precisely the same as the so called affine scaling trajectory. Such an identification is realizable only if the barrier parameter is varied in an idiosyncratic manner. The unique and continuous way of transforming the barrier parameter is different from the well established strategy in practice of iteratively adapting the parameter to enhance the computational efficiency, but it unifies in theory the concepts of central path and affine scaling trajectory. Secondly, we argue that, not just symbolically, but in a mathematically rigorous justification, the same notion of projected gradient can be generalized from the linear programming to the semi-definite programming. By introducing a suitable operator to accommodate the non-commutativity of matrix multiplication, we are able to characterize the matrix-valued affine scaling trajectory for the semi-definite programming problem in a way that is extremely similar in appearance to that for the linear programming problem.

*Department of Mathematics, North Carolina State University, Raleigh, NC 27695-8205. (chu@math.ncsu.edu).
This research was supported in part by the National Science Foundation under grants DMS-0732299 and DMS-1014666.
†Department of Mathematics, National Chung Cheng University, Chiaya 62012 Taiwan. (mlin@math.ccu.edu.tw).
This research was supported in part by the National Science Foundation under grants DMS-0505880 and DMS-0732299.

2. Preliminaries. To set the background information, we begin with a brief review of the central path for a linear programming (LP) problem in the primal standard form

$$\mathrm{P_L}:\begin{cases}\text{minimize} & \mathbf{c}^\top\mathbf{x},\\ \text{subject to} & A\mathbf{x}=\mathbf{b},\quad \mathbf{x}\geq 0,\end{cases}\tag{2.1}$$

together with its corresponding dual standard form

$$\mathrm{D_L}:\begin{cases}\text{maximize} & \mathbf{b}^\top\mathbf{y},\\ \text{subject to} & A^\top\mathbf{y}+\mathbf{s}=\mathbf{c},\quad \mathbf{s}\geq 0,\end{cases}\tag{2.2}$$

where $\mathbf{c}$, $\mathbf{x}\in\mathbb{R}^n$, $A\in\mathbb{R}^{m\times n}$, $\mathbf{b}\in\mathbb{R}^m$, $\mathbf{s}\in\mathbb{R}^n$ and $m<n$. It is convenient to identify a vector $\mathbf{v}$ with the diagonal matrix $V:=\mathrm{Diag}(\mathbf{v})$ and vice versa. With the introduction of the logarithmic barrier function

$$f(X)=-\ln\det(X)\tag{2.3}$$

for any symmetric and positive semi-definite matrix X, the interior point method reformulates the primal and the dual LP problems as

$$\mathrm{BP_L}(\mu):\begin{cases}\text{minimize} & \mathbf{c}^\top\mathbf{x}+\mu f(X),\\ \text{subject to} & A\mathbf{x}=\mathbf{b},\quad \mathbf{x}\geq 0,\end{cases}\tag{2.4}$$

and

$$\mathrm{DP_L}(\mu):\begin{cases}\text{maximize} & \mathbf{b}^\top\mathbf{y}-\mu f(S),\\ \text{subject to} & A^\top\mathbf{y}+\mathbf{s}=\mathbf{c},\quad \mathbf{s}\geq 0,\end{cases}\tag{2.5}$$

respectively, where again $\mathrm{BP_L}(\mu)$ and $\mathrm{DP_L}(\mu)$ are the Lagrangian dual to each other. By the theory of Lagrange multipliers, we see that the triplets $(\mathbf{x},\mathbf{y},\mathbf{s})$ at the optimal solution to these logarithmic barrier problems are roots of the function $F:\mathbb{R}^n\times\mathbb{R}^m\times\mathbb{R}^n\times\mathbb{R}\rightarrow\mathbb{R}^n\times\mathbb{R}^m\times\mathbb{R}^n$ defined by

$$F(\mathbf{x},\mathbf{y},\mathbf{s},\mu):=\begin{bmatrix}A^\top\mathbf{y}+\mathbf{s}-\mathbf{c}\\ A\mathbf{x}-\mathbf{b}\\ XS-\mu I\end{bmatrix}.\tag{2.6}$$

The equation $F(\mathbf{x},\mathbf{y},\mathbf{s},\mu)=0$, subject to the constraints $\mathbf{x}\geq 0$ and $\mathbf{s}\geq 0$, is known as the *central path equation* (CPE). Assume that A is of full row rank and that feasible sets of (2.1) and (2.2) contain interior feasible points. Then for each positive μ, there is a unique solution $(\mathbf{x}_\mu,\mathbf{y}_\mu,\mathbf{s}_\mu)$ to the CPE. Indeed, these solutions form an analytic path as is guaranteed by the implicit function theorem.

The most effective interior point methods do not follow the central path closely. Instead, they act in a neighborhood of the central path. Take the primal-dual algorithm as an example [15, 19]. The basic strategy is to take "some" Newton steps towards the minimizer of $\mathrm{BP_L}(\mu)$ for each fixed parameter $\mu>0$, intending to approximately solve the CPE until a specified proximity criterion is satisfied. The beauty is that if the parameter μ is decreased to zero in a certain specific way and the iteration repeats, then often one single Newton step is all needed in the computation and convergence in polynomial time can be achieved. Such a path-following mechanism is similar in spirit to but different in practice from tracking the central path by means of some numerical ODE techniques. The latter idea is closer to the notion of affine scaling algorithms which will be explored further in the subsequent discussion.

For the matter of clarity, we stress hereby that the term "trajectory" refers to the integral curve of a certain differential system. This trajectory is to be distinguished from the sequence of iterates generated by any path-following algorithms. Being an analytic curve, the central path is governed by the differential equation

$$\underbrace{\begin{bmatrix} 0_{n\times n} & A^\top & I_n & 0_{n\times 1} \\ A & 0_{m\times m} & 0_{m\times n} & 0_{m\times 1} \\ S & 0_{n\times m} & X & -\mathbf{e}_n \end{bmatrix}}_{\mathfrak{A}(\mathbf{x},\mu)} \begin{bmatrix} \dot{\mathbf{x}} \\ \dot{\mathbf{y}} \\ \dot{\mathbf{s}} \\ \dot{\mu} \end{bmatrix} = 0, \tag{2.7}$$

where the derivative $\dot{} = \frac{d}{dt}$ is taken with respect to some scalar variable t parameterizing the central path and $\mathbf{e}_n \in \mathbb{R}^n$ stands for the column vector of all 1's. Owing to the relationship $XS = \mu I_m$, the coefficient matrix in (2.7) depends on $(\mathbf{x}, \mu)$ only. Let $\mathbf{n}(\mathbf{x}, \mu)$ denote a basis of the null space of $\mathfrak{A}(\mathbf{x}, \mu)$. Without loss of generality, we may normalize the last entry and assume

$$\mathbf{n}(\mathbf{x}, \mu) := \begin{bmatrix} \mathbf{n}_1(\mathbf{x}, \mu) \\ \mathbf{n}_2(\mathbf{x}, \mu) \\ \mathbf{n}_3(\mathbf{x}, \mu) \\ 1 \end{bmatrix}. \tag{2.8}$$

It is then legitimate to claim that the path $(\mathbf{x}(t), \mu(t))$ is characterized by the self-contained dynamical system

$$\left\{ \begin{array}{lcl} \dot{\mathbf{x}} & = & \alpha(t)\mathbf{n}_1(\mathbf{x}, \mu), \\ \dot{\mu} & = & \alpha(t), \end{array} \right. \tag{2.9}$$

where $\alpha(t)$ can be any continuous function such that the corresponding solution $\mu(t)$ stays positive and decays to zero. One such an example is $\alpha(t) = -\mu(t)$, but we are interested in a more special one which will be described later.

Suppose that $\mathbf{x}(0)$ is feasible. It is known that if $\mu(t)$ decreases to zero, then so does the value $\mathbf{c}^\top\mathbf{x}(t)$ decrease to the optimal objective value for the LP [20, Theorem 2.17]. The solution $\mathbf{x}(t)$ to the differential system (2.9), which represents only the primal portion of the "full" central path, therefore is inherently a descent flow on the affine subspace of solutions to $A\mathbf{x} = \mathbf{b}$. We can characterize the vector field on the right side of the dynamical system (2.9) more definitively as we shall derive its closed form below.

2.1. Projected Gradient Flow. For each fixed parameter $\mu > 0$, we describe two kinds of gradient flows that link the central path to the so called affine scaling trajectory.

We first introduce the steepest descent flow. Denote the objection function of the $\mathrm{BP_L}(\mu)$ for each fixed parameter μ by

$$\phi(\mathbf{x}; \mu) = \mathbf{c}^\top\mathbf{x} - \mu \ln \det(X). \tag{2.10}$$

Trivially, we have

$$\nabla\phi(\mathbf{x}; \mu) = \mathbf{c} - \mu X^{-1}\mathbf{e}. \tag{2.11}$$

The flow $\mathbf{x}(t; \mu)$ defined by the negative projected gradient

$$\dot{\mathbf{x}} := -\underbrace{(I - A^\top(AA^\top)^{-1}A)}_{\mathscr{P}_{\mathcal{N}(A)}}(\mathbf{c} - \mu X^{-1}\mathbf{e}), \tag{2.12}$$

where $\mathscr{P}_{\mathcal{N}(A)}$ denotes the projection operator onto the null space $\mathcal{N}(A)$ of A, therefore characterizes the steepest descent movement for $\phi(\mathbf{x};\mu)$ on the feasible set. Note that $\mathscr{P}_{\mathcal{N}(A)}$ is positive semi-definite. Any zero main diagonal entry of $\mathscr{P}_{\mathcal{N}(A)}$ must lie on a zero row and column. It follows that except for the degenerate case, if x_i is sufficiently near zero, then by (2.12) we shall have $\dot{x}_i > 0$. In other words, $\mathbf{x}(t;\mu)$ cannot converge to a boundary point. Rather, $\mathbf{x}(t;\mu)$ converges to a local minimum of $\phi(\mathbf{x};\mu)$ which, by convexity, is the unique $\mathbf{x}_\mu$ on the central path,

Example 1. Consider the case $A = [1, 2]$, $b = 2$, and $\mathbf{c} = [1, 1]^\top$. Then the projected gradient flow is given by

$$\begin{cases} \dot{x}_1 &= -\frac{2}{5}(1 - \frac{2\mu}{x_1} + \frac{\mu}{x_2}), \\ \dot{x}_2 &= \frac{1}{5}(1 - \frac{2\mu}{x_1} + \frac{\mu}{x_2}). \end{cases}$$

Upon the substitution by $x_1 = 2 - 2x_2$, it suffices to consider

$$\dot{x}_2 = \frac{1}{5}\left(1 + \mu\frac{1 - 2x_2}{x_2(1 - x_2)}\right), \quad 0 < x_2 < 1.$$

Suppose $0 < \mu < 1$. Then $\dot{x}_2 \lesseqgtr 0$ if $x_2 \gtreqless \frac{1-2\mu+\sqrt{1+4\mu^2}}{2} = 1 - \mu + \mu^2 + O(\mu^4)$, implying that $x_2(t)$ is being pushed away from the boundary points 0 and 1, and converges to an equilibrium.

We next introduce a scaled gradient flow. Given an interior feasible point $\widehat{\mathbf{x}}$, define the transformation (of scaling)

$$\mathbf{w} := \widehat{X}^{-1}\mathbf{x} \tag{2.13}$$

and reformulate the LP in terms of the variable $\mathbf{w}$. In particular, the objection function for the $\mathrm{BP_L}(\mu)$ becomes

$$\phi(\mathbf{x};\mu) = \theta(\mathbf{w};\mu) := \mathbf{c}^\top \widehat{X}\mathbf{w} - \mu \ln\det(\widehat{X}W). \tag{2.14}$$

Note that the function θ is well defined because $\widehat{X}W$ remains diagonal and positive. The gradient of θ at $\mathbf{e}$ in the $\mathbf{w}$-space is given by

$$\nabla\theta(\mathbf{e}) = \widehat{X}\mathbf{c} - \mu\mathbf{e}. \tag{2.15}$$

Similar to (2.12), the negative projected gradient at $\mathbf{e}$

$$\dot{\mathbf{w}}|_{\mathbf{w}=\mathbf{e}} := -\underbrace{(I - \widehat{X}A^\top(A\widehat{X}^2A^\top)^{-1}A\widehat{X})}_{\mathscr{P}_{\mathcal{N}(A\widehat{X})}}(\widehat{X}\mathbf{c} - \mu\mathbf{e}) \tag{2.16}$$

points to a feasible and steepest descent direction for $\theta(\mathbf{w};\mu)$ at $\mathbf{e}$. Pulling this vector back to the $\mathbf{x}$-space, we obtain a descent direction $\widehat{X}\dot{\mathbf{w}}|_{\mathbf{w}=\mathbf{e}}$, but not necessarily the steepest, for the objection function $\phi(\mathbf{x};\mu)$ at $\widehat{\mathbf{x}}$. Applying this procedure continuously to every interior feasible point, we thus obtain a new dynamical system

$$\dot{\mathbf{x}} := -X\left(I - XA^\top(AX^2A^\top)^{-1}AX\right)(X\mathbf{c} - \mu\mathbf{e}), \tag{2.17}$$

whose solution $\mathbf{x}(t;\mu)$ stays feasible and moves to decrease the objection value $\phi(\mathbf{x};\mu)$.

Example 2. Using the same data as in Example 1, the system (2.17) gives rise to the differential system

$$\begin{cases} \dot{x}_1 &= \frac{-2x_1x_2(x_1x_2 - 2x_2\mu + x_1\mu)}{{x_1}^2 + 4{x_2}^2}, \\ \dot{x}_2 &= \frac{x_1x_2(x_1x_2 - 2x_2\mu + x_1\mu)}{{x_1}^2 + 4{x_2}^2}. \end{cases}$$

Equivalently, after substitution, we have the stand alone system

$$\dot{x}_2 = \frac{(-1+x_2)\,x_2\left(-x_2+{x_2}^2+2x_2\mu-\mu\right)}{1-2x_2+2{x_2}^2}$$

for x_2. It is easy to check that the system has the property that $\dot{x}_2 = 0$ at $x_2 = 0, 1$, but $\dot{x}_2 \lesseqgtr 0$ if $x_2 \gtreqless \frac{1-2\mu+\sqrt{1+4\mu^2}}{2}$. Thus, even though $x_2 = 0, 1$ are equilibria points, they are repellers. Still, $\mathbf{x}(t;\mu)$ converges to an interior equilibrium point which is $\mathbf{x}_\mu$.

2.2. Relating Central Path to Gradient Flow and Affine Scaling Trajectory. The "phase portraits" of the three dynamical systems, (2.9), (2.12), and (2.17) discussed thus far can be depicted in Figure 2.1. We use the vertical axis to represent the barrier parameter μ and the horizontal plane to represent the overly simplified feasible set in $\mathbb{R}^n$ for the primal variable $\mathbf{x}$. For each fixed μ, the green curve represents the trajectory of the projected gradient flow defined by either (2.12) or (2.17). The flow converges to the unique point $\mathbf{x}_\mu$ on the central path defined by (2.7). Based on these projected gradient flows, we consider two possible scenarios when letting μ decrease to zero.

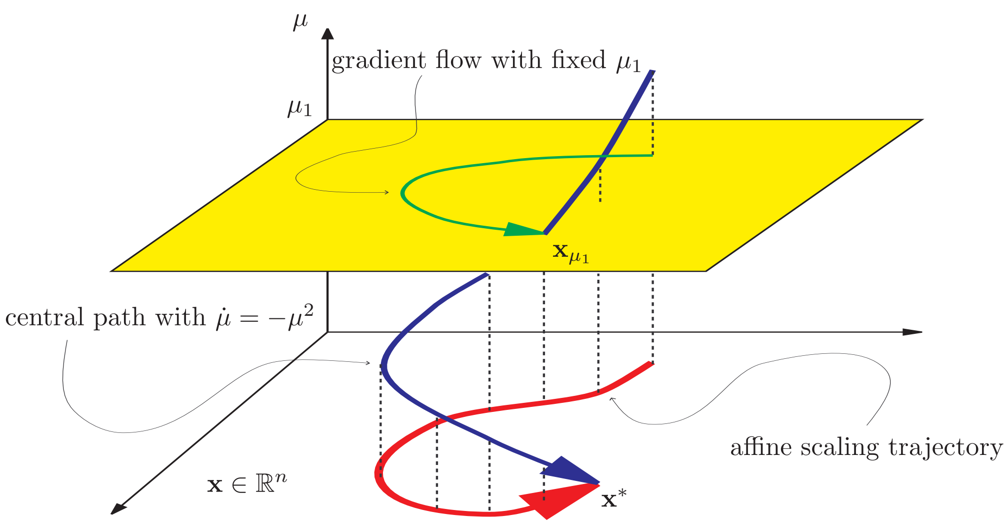

FIG. 2.1. *Central path, gradient flow, and affine scaling trajectory*

In the first scenario, observe that the vector field in either (2.12) or (2.17) is continuous in the parameter μ. By the theory of continuous dependence on initial data and parameters in ODE, we know that

$$\lim_{\mu\to 0} \mathbf{x}(t;\mu) = \mathbf{x}(t;0).$$

The corresponding dynamical systems to (2.12) and (2.17) are reduced, repectively, to

$$\dot{\mathbf{x}} := -(I - A^\top(AA^\top)^{-1}A)\mathbf{c}, \tag{2.18}$$

and

$$\dot{\mathbf{x}} := -X\left(I - XA^\top(AX^2A^\top)^{-1}AX\right)X\mathbf{c}, \tag{2.19}$$

respectively.

The first projected gradient system (2.18) is trivial because its right hand side is a constant vector. Unless $\mathscr{P}_{\mathcal{N}(A)}\mathbf{c} = 0$ to begin with, its solution flow $\mathbf{x}(t;0) = \mathbf{x}_0 - t(I - A^\top(AA^\top)^{-1}A)\mathbf{c}$ moves in a fixed direction until it reaches the boundary, that is, some entries become zero. Even at this optimal point, the projected gradient is not zero.

The solution flow $\mathbf{x}(t;0)$ corresponding to the second system (2.19) is more interesting. It is precisely the so called *affine scaling trajectory* discussed in [5, 6]. It can be argued [5, Lemma 4.2] that, if $\mathscr{P}_{\mathcal{N}(A)}\mathbf{c} \neq 0$, then $\dot{\mathbf{x}} \neq 0$ at every interior feasible point $\mathbf{x}$. The flow therefore has to move asymptotically toward the boundary of the feasible set. Its limiting behavior can be derived from the general framework of weighted primal affine scaling trajectory (w-PAS) discussed in [1]. Though both are descent flows, the fundamental difference between (2.18) and (2.19) is that, starting with an interior feasible point, the former flow will hit the boundary in finite time whereas the latter flow stays in the interior of the feasible set for all t.

In the second scenario, we turn to the task of more specifically characterizing the vector field for the dynamical system (2.7) that sets apart the central path.

THEOREM 2.1. *The triplets $(\mathbf{x}, \mathbf{y}, \mathbf{s})$ in the central path defined by the optimality condition $F(\mathbf{x}, \mathbf{y}, \mathbf{s}, \mu) = 0$ with F given by (2.6) are governed by the dynamical systems,*

$$\dot{\mathbf{x}} = \frac{\dot{\mu}}{\mu^2} X \mathscr{P}_{\mathcal{N}(AX)} X\mathbf{c}, \tag{2.20}$$

$$\dot{\mathbf{y}} = \frac{\dot{\mu}}{\mu}\left(\mathbf{y} - (AX^2A^\top)^{-1}AX^2\mathbf{c}\right), \tag{2.21}$$

$$\dot{\mathbf{s}} = \frac{\dot{\mu}}{\mu} A^\top(AX^2A^\top)^{-1}AX^2\mathbf{s}, \tag{2.22}$$

respectively, where the choice of $\dot{\mu}$ can be arbitrary.

Proof. For the convenience of reference, we rehash (2.7) in its block form

$$A^\top\dot{\mathbf{y}} + \dot{\mathbf{s}} = 0, \tag{2.23}$$

$$A\dot{\mathbf{x}} = 0, \tag{2.24}$$

$$S\dot{\mathbf{x}} + X\dot{\mathbf{s}} = \dot{\mu}\mathbf{e} \tag{2.25}$$

and proceed to solve for the derivatives in closed form.

By (2.23), we know that $\dot{\mathbf{s}}$ is in the range space $\mathcal{R}(A^\top)$ of $A^\top$ and, thus, $X\dot{\mathbf{s}} \perp \mathcal{N}(AX)$. Note also that, by (2.24), $X^{-1}\dot{\mathbf{x}} \in \mathcal{N}(AX)$. Rewrite (2.25) as $\mu X^{-1}\dot{\mathbf{x}} + X\dot{\mathbf{s}} = \frac{\dot{\mu}}{\mu}X\mathbf{s}$ and apply the projector $\mathscr{P}_{\mathcal{N}(AX)}$ to its both sides. By the fact that $(AX)^\top\mathbf{y} + X\mathbf{s} = X\mathbf{c}$ from (2.6), we see that $\mathscr{P}_{\mathcal{N}(AX)}(X\mathbf{s}) = \mathscr{P}_{\mathcal{N}(AX)}(X\mathbf{c})$ and the *primal central path differential system* (2.20) for $\mathbf{x}(t)$ is proved.

Likewise, rewrite (2.25) as $\mu\dot{\mathbf{x}} - X^2A^\top\dot{\mathbf{y}}_= \left(\frac{\dot{\mu}}{\mu}\right)X^2\mathbf{s}$. Upon substitution by (2.20), together with the fact that $AX^2A^\top\mathbf{y} + AX^2\mathbf{s} = AX^2\mathbf{c}$, we obtain the dual central path differential system (2.21) and the slack central path differential system (2.22) for $\mathbf{y}(t)$ and $\mathbf{s}(t)$, respectively.

It can be checked that, with the definition for $\dot{\mathbf{x}}$ and $\dot{\mathbf{s}}$, the equation (2.25) is automatically satisfied with arbitrary $\dot{\mu}$. Starting with any feasible initial value $(\mathbf{x}_0, \mathbf{y}_0, \mathbf{s}_0)$, therefore, the solution flow $(\mathbf{x}(t), \mathbf{y}(t), \mathbf{s}(t))$ stays on the central path for all t. ∎

Because we typically prefer that $\mu(t)$ is driven to zero, of particular interest is the choice

$$\dot{\mu} := -\mu^2. \tag{2.26}$$

COROLLARY 2.2. *If $\mu(t) = \frac{1}{t+\mu_0^{-1}}$, then the dynamical system (2.20) coincides with (2.19). That is, the $\mathbf{x}$ potion of the full central path which is derived from the KKT condition is precisely the same as the affine scaling trajectory which is derived from the notion of projected gradient.*

It is illuminating to represent the relationship characterized in Corollary 2.2 in Figure 2.1 by "projecting" the central path to the horizontal plane to obtain the affine scaling trajectory. It is important to note that the dynamical system (2.19) is autonomous in $\mathbf{x}$ and makes no reference to either the dual variable $\mathbf{y}$ or the slack variable $\mathbf{s}$ at all.

3. Generalization to the SDP. Semi-definite programming (SDP) concerns finding a symmetric matrix to optimize a linear functional subject to linear constraints and the additional condition that the matrix be positive semi-definite. The SDP in the primal standard form is given by

$$\begin{array}{ll} \text{Minimize}_X & \langle C, X\rangle \\ \text{Subject to} & \langle A_i, X\rangle = b_i, \quad i = 1, \ldots, m \\ & X \succeq 0, \end{array} \tag{3.1}$$

where C, $A_1, \ldots, A_m$ are given symmetric matrices in $\mathbb{R}^{n\times n}$, $X \succeq 0$ means that X is positive semi-definite, and $\langle C, X\rangle$ denotes the Frobenius inner product between C and X. Without loss of generality, we assume henceforth that $A_1, \ldots, A_m$ are linear independent.

In recent years, the SDP has emerged as an important tool in mathematical programming. One reason for this to happen is because the notion of SDP is versatile enough to model problems arising in broad discipline areas. Some synoptic discussions on applications ranging from mathematical studies in combinatorial optimization, Boolean and non-convex quadratic programming, min-max eigenvalue problems, and matrix completion problems to engineering practices in nonlinear and time-varying system analysis, controller synthesis, computer-aided control system design, network queueing, optimal statistical model designs, and structural optimization can be found in [7, 15, 18]. Another reason for its popularity is because there is a considerable similarity between the notions of SDP and LP. Many results developed for the LP can be extended "mechanically" to the SDP.

In particular, let $S\mathbb{R}^{n\times n}$ denote the subspace of all $n \times n$ real symmetric matrices and $S\mathbb{R}^{n\times n}_+$ the subset of positive definite matrices in $S\mathbb{R}^{n\times n}$. If we introduce the linear operator $\mathscr{A} : S\mathbb{R}^{n\times n} \to \mathbb{R}^m$ defined by

$$\mathscr{A}X := [\langle A_i, X\rangle]_{i=1}^m, \tag{3.2}$$

then its corresponding adjoint $\mathscr{A}^* : \mathbb{R}^m \to S\mathbb{R}^{n\times n}$ is given by

$$\mathscr{A}^*\mathbf{v} = \sum_{i=1}^m v_i A_i. \tag{3.3}$$

With this notation, we can write the primal and the dual SDP problems as [15]

$$\mathrm{P}_{\mathrm{SDP}} : \begin{cases} \text{minimize} & \langle C, X\rangle, \\ \text{subject to} & \mathscr{A}X = \mathbf{b}, \quad X \succeq 0, \end{cases} \tag{3.4}$$

and

$$\mathrm{D}_{\mathrm{SDP}} : \begin{cases} \text{maximize} & \mathbf{b}^\top \mathbf{y}, \\ \text{subject to} & \mathscr{A}^*\mathbf{y} + S = C, \quad S \succeq 0, \end{cases} \tag{3.5}$$

respectively, whose appearances are now almost identical to those described in (2.1) and (2.2). Much of the dual theory, interior point algorithms, convergence and polynomial time-complexity for LP can be extended to SDP [2, 14]. This generalization thus admits theoretically efficient solution procedures based on what has already been developed for LP problems. A profusion of research results are available in the literature. For example, the book on SDP [18] lists 877 references, while the online bibliography collected by Wolkowicz [17] lists more than a thousand and the number continues growing.

As casual users, we find that the comprehensive treatise in the two books [3, 7] and the two review articles [2, 15] offer quick and useful grasp of this interesting and intensely studied subject. Our goal in this section is to investigate whether the dynamical systems characterized in the preceding sections for the LP can also be generalized to the central path for the SDP.

3.1. Central Path. Analogous to (2.6), it can be shown that the central path equation, i.e., the optimality condition, for an SDP is

$$G(X, \mathbf{y}, S, \mu) := \begin{bmatrix} \mathscr{A}^*\mathbf{y} + S - C \\ \mathscr{A}X - \mathbf{b} \\ S - \mu X^{-1} \end{bmatrix} = 0, \tag{3.6}$$

where X and S are expected to be symmetric and positive semi-definite. Note that the third equation in G is equivalent to $XS = \mu I$, but is conventionally written in this way for the purpose of maintaining its images in the subspace $S\mathbb{R}^{n\times n}$. It is a known fact (see, e.g., [15, Theorem 5.2]) that if both $\mathrm{P}_{\mathrm{SDP}}$ and $\mathrm{D}_{\mathrm{SDP}}$ have strictly feasible points, then corresponding to each positive μ there is a unique solution $(X(\mu), \mathbf{y}(\mu), S(\mu))$ to the equation (3.6). With $\mathscr{A}$ and $\mathscr{A}^*$ denoting operators in mind, we want to solve the linear system

$$\mathscr{A}^*\dot{\mathbf{y}} + \dot{S} = 0, \tag{3.7}$$

$$\mathscr{A}\dot{X} = 0, \tag{3.8}$$

$$\mu X^{-1}\dot{X}X^{-1} + \dot{S} = \dot{\mu}X^{-1}. \tag{3.9}$$

for the dynamical system $(\dot{X}, \dot{\mathbf{y}}, \dot{S}, \dot{\mu})$ of the central path.

We first tackle this problem by the approach suggested in [15]. It suffices to characterize only the primal central path dynamical system.

Lemma 3.1. *The X portion of the central path in the SDP is governed by the differential system*

$$\dot{X} = \frac{\dot{\mu}}{\mu}(I - X\mathscr{A}^*\Omega^{-1}\mathscr{A}X)X, \tag{3.10}$$

where $\Omega = [\omega_{ij}]$ is the $m \times m$ matrix defined by

$$\omega_{ij} := \langle A_i X, X A_j \rangle. \tag{3.11}$$

Proof. Upon substituting $\dot{S} = -\mathscr{A}^*\dot{\mathbf{y}}$ from (3.7) into (3.9), we obtain

$$\dot{X} = \frac{1}{\mu}X(\dot{\mu}X^{-1} + \mathscr{A}^*\dot{\mathbf{y}})X. \tag{3.12}$$

Applying $\mathscr{A}$ to both sides of (3.12), by (3.8), we see that $\dot{\mathbf{y}}$ can be solved from the linear equation

$$\Omega\dot{\mathbf{y}} = -\dot{\mu}\mathscr{A}X. \tag{3.13}$$

Replacing $\dot{\mathbf{y}}$ in (3.12) gives rise to (3.10). ∎

In the above, the operator $X\mathscr{A}^*\Omega^{-1}\mathscr{A}X$ seems symmetric at first glance. However, it really should be read as the $n \times n$ matrix

$$X\mathscr{A}^*\Omega^{-1}\mathscr{A}X = X(\mathscr{A}^*(\Omega^{-1}\mathbf{b})) = \sum_{i=1}^{m}(\Omega^{-1}\mathbf{b})_i X A_i,$$

which generally is not a symmetric matrix at all.

Recall that the LP is a special case of the SDP. This can be seen through the convention that $X = \operatorname{diag}(\mathbf{x})$, $A_i = \operatorname{diag}(\mathbf{a}_i)$ and $A^\top = [\mathbf{a}_1, \ldots, \mathbf{a}_m]$. We then have $\mathscr{A}X = A\mathbf{x}$ and $\mathscr{A}^*\mathbf{y} = \operatorname{diag}(A^\top \mathbf{y})$. It is easy to see that $\Omega = AX^2A^\top$ and that the system (3.10) is reduced to

$$\dot{\mathbf{x}} = \frac{\dot{\mu}}{\mu} X(I - XA^\top(AX^2A)^{-1}AX)\mathbf{e}$$

which is identical to the primal central path system (2.20) (and, hence, the affine scaling system (2.19) via the choice (2.26)) by the fact that $\mathscr{P}_{\mathcal{N}(AX)}(X\mathbf{s}) = \mathscr{P}_{\mathcal{N}(AX)}(\mu\mathbf{e})$. The question now is whether a notion of affine scaling trajectory for the LP can be generalized to the SDP.

Before we move on to derive the matrix-valued affine scaling trajectory, it is important to point out one subtle difference between a continuous trajectory and an associated discrete version of path-following algorithm. The affine scaling algorithm, originally proposed by [8] and rediscovered or modified by others [4, 5, 16], is an iterative method in which the affine scaling direction and the step size for each iteration are selected with respect to the associated Dikin ellipsoid. The affine scaling algorithm, like many other polynomial-time interior point methods, can be naturally extended to the SDP [13], but a task of extending the affine scaling "trajectory" to the SDP (with the same geometric meaning as that for the LP) has not been so obvious in the literature. The hindrance is partially due to the fact that, the underlying SDP variables no longer being diagonal, matrix multiplication is not commutative. Indeed, while it has been proved that the primal central path of the SDP problem, that is, the trajectory determined by (3.10), converges to the analytic center of the optimal face [10], it has also been demonstrated in [13] that both short-step and long-step affine scaling "algorithms" using a certain affine scaling "direction" fail to converge to an optimal point.

If we can prove, as will be done in the following, that the primal central path is in fact the same as the affine scaling trajectory, then this distinction between an affine scaling algorithm and an affine scaling trajectory is significant. An appropriate discretization of the continuous flow might lead to a different numerical algorithm.

3.2. Affine Scaling Trajectory. A clever way to circumvent the difficulty associated with the non-commutativity of matrix multiplication has been proposed in [9]. It relies on the machineries of a specialized Riemannian metric and a group of linear isometries to induce a matrix-valued generalized affine scaling vector field. We find, however, that the derivation of the theory could be obtained via the following approach which is simpler and our resulting dynamical system conforms more favorably than the formula described in [9, Proposition 2.8] to the system already developed for the LP. Our main point is that, when cast under appropriate framework of operators, the notion of gradient flows for the LP carries over to the SDP naturally. We explain our ideas below.

Firstly, we find the diagram in Figure 3.1 instructive in indicating how the operators go back and forth between $S\mathbb{R}^{n\times n}$ and $\mathbb{R}^m$. In particular, the mapping $\mathscr{A}\mathscr{A}^* : \mathbb{R}^m \to \mathbb{R}^m$ can be represented by the $m \times m$ matrix

$$\mathscr{A}\mathscr{A}^* = [\langle A_i, A_j\rangle]. \tag{3.14}$$

The operator

$$\mathscr{P}_{\mathcal{N}(\mathscr{A})} := I - \mathscr{A}^*(\mathscr{A}\mathscr{A}^*)^{-1}\mathscr{A} \tag{3.15}$$

maps $S\mathbb{R}^{n\times n}$ to $S\mathbb{R}^{n\times n}$ and is in fact a projection operator onto the null space $\mathcal{N}(\mathscr{A})$ of $\mathscr{A}$. Thus, with respect to the primal barrier problem associated with $\mathrm{P}_{\mathrm{SDP}}$, the dynamical system

$$\dot{X} := -(I - \mathscr{A}^*(\mathscr{A}\mathscr{A}^*)^{-1}\mathscr{A})(C - \mu X^{-1}) \tag{3.16}$$

defines the projected gradient flow for the objective function

$$\varphi(X;\mu) := \langle C, X\rangle - \mu \ln\det(X). \tag{3.17}$$

$$
\begin{tikzcd}
S\mathbb{R}^{n\times n} \arrow[r, "\mathscr{A}"] \arrow[dr, phantom, "\mathscr{A}^*(\mathscr{A}\mathscr{A}^*)^{-1}\mathscr{A}"] & \mathbb{R}^m \arrow[d, dashed, "(\mathscr{A}\mathscr{A}^*)^{-1}"] & S\mathbb{R}^{n\times n} \arrow[l, "\mathscr{A}"'] \\
S\mathbb{R}^{n\times n} \arrow[u, dashed] & \mathbb{R}^m \arrow[l, "\mathscr{A}^*"] & \mathbb{R}^m \arrow[l, dotted] \arrow[ul, "\mathscr{A}\mathscr{A}^*"] \arrow[u, "\mathscr{A}^*"']
\end{tikzcd}
$$

FIG. 3.1. *Operators $\mathscr{A}$, $\mathscr{A}^*$, $\mathscr{A}\mathscr{A}^*$ and their domains*

Secondly, mimicking (2.13), we define the (scaled) set

$$\mathbb{W} := \widehat{X}^{-1}S\mathbb{R}^{n\times n} = \{\widehat{X}^{-1}Z | Z \in S\mathbb{R}^{n\times n}\} \tag{3.18}$$

for any fixed interior feasible point $\widehat{X}$. For later reference, let $\mathbb{W}_+$ denote the open subset when specifically $Z \in S\mathbb{R}^{n\times n}_+$. By symmetry, we can also write

$$\mathbb{W} = S\mathbb{R}^{n\times n}\widehat{X}. \tag{3.19}$$

Corresponding to any $W \in \mathbb{W}$, there exist unique matrices $\mathscr{L}(W), \mathscr{R}(W) \in S\mathbb{R}^{n\times n}$ such that

$$W = \widehat{X}^{-1}\mathscr{L}(W) = \mathscr{R}(W)\widehat{X}, \tag{3.20}$$

whence $\mathscr{L}(W) = \widehat{X}\mathscr{R}(W)\widehat{X}$. Clearly, the operators $\mathscr{L}$ and $\mathscr{R}$ are isomorphisms from $\mathbb{W}$ to $S\mathbb{R}^{n\times n}$. These operators are used to address the non-commutativity.

Define the map $[\![\cdot,\cdot]\!] : \mathbb{W}\times\mathbb{W} \to \mathbb{R}$ by

$$[\![W_1, W_2]\!] := \langle \mathscr{L}(W_1), \mathscr{R}(W_2)\rangle. \tag{3.21}$$

LEMMA 3.2. *The bilinear map $[\![\cdot,\cdot]\!]$ is an inner product on the subspace $\mathbb{W}$.*

Proof. Observe first that

$$[\![W_2, W_1]\!] = \langle \mathscr{L}(W_2), \mathscr{R}(W_1)\rangle = \langle \widehat{X}\mathscr{R}(W_2)\widehat{X}, \widehat{X}^{-1}\mathscr{L}(W_1)\widehat{X}^{-1}\rangle = [\![W_1, W_2]\!].$$

Observe next that

$$[\![W, W]\!] = \langle \mathscr{L}(W), \mathscr{R}(W)\rangle = \langle \widehat{X}, \mathscr{R}(W)\widehat{X}\mathscr{R}(W)\rangle > 0$$

unless $W = 0$. These properties of conjugate symmetry, linearity, and positive definiteness justify that $[\![\cdot,\cdot]\!]$ is an inner product on the subspace $\mathbb{W}$. $\square$

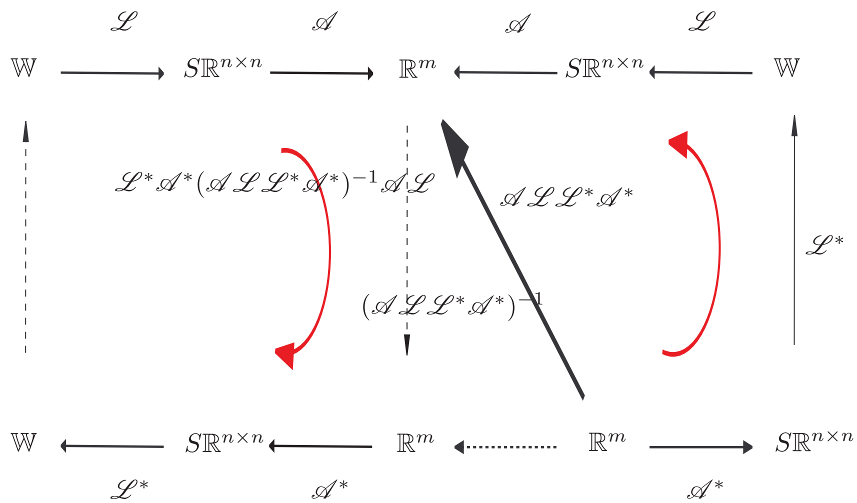

FIG. 3.2. *Operators $\mathscr{A}\mathscr{L}\mathscr{L}^*\mathscr{A}^*$, $\mathscr{L}^*\mathscr{A}^*(\mathscr{A}\mathscr{L}\mathscr{L}^*\mathscr{A}^*)^{-1}\mathscr{A}\mathscr{L}$, and their domains*

LEMMA 3.3. *As $\mathscr{L} : \mathbb{W} \to S\mathbb{R}^{n\times n}$ acts as a left shift operator on $\mathbb{W}$, its adjoint with respect to the inner product $[\![\cdot,\cdot]\!]$ is given by*

$$\mathscr{L}^*(Z) = Z\widehat{X} \tag{3.22}$$

and acts as a right shift operator on $S\mathbb{R}^{n\times n}$.

Proof. Note that domain and range of $\mathscr{L}$ involve different inner products. To define the adjoint operator of $\mathscr{L}$, the relationship

$$\langle \mathscr{L}(W), Z\rangle = \langle \widehat{X}W, Z\rangle = [\![\widehat{X}^{-1}(\widehat{X}W), Z\widehat{X}]\!] = [\![W, Z\widehat{X}]\!]$$

must hold. The adjoint operator $\mathscr{S}^* : S\mathbb{R}^{n\times n} \to \mathbb{W}$ has to be that defined in (3.22). ▯

We now extend the relationships in Figure 3.1 to the diagram in Figure 3.2 to include the role played by the shift operator $\mathscr{L}$ and its adjoint.

LEMMA 3.4. *The mapping $\mathscr{A}\mathscr{L}\mathscr{L}^*\mathscr{A}^* : \mathbb{R}^m \to \mathbb{R}^m$ has a matrix representation*

$$\mathscr{A}\mathscr{L}\mathscr{L}^*\mathscr{A}^* = [\langle A_i\widehat{X}, \widehat{X}A_j\rangle]. \tag{3.23}$$

Proof. Given any vector $\mathbf{z} \in \mathbb{R}^m$, we carry out the operations from the right to the left as follows,

$$\mathscr{A}\mathscr{L}\mathscr{L}^*\mathscr{A}^*(\mathbf{z}) = \mathscr{A}\mathscr{L}\mathscr{L}^*\left(\sum_{i=1}^m z_iA_i\right) = \mathscr{A}\mathscr{L}\left(\sum_{j=1}^m z_jA_j\widehat{X}\right) = \mathscr{A}\left(\sum_{j=1}^m z_j\widehat{X}A_j\widehat{X}\right).$$

Finally, we use the relationship $\langle A_i, \widehat{X}A_j\widehat{X}\rangle = \langle \widehat{X}A_i, A_j\widehat{X}\rangle$ and obtain the matrix representation. ▯

With the change of variable $W = \hat{X}^{-1}X$ for matrices $X \in S\mathbb{R}^{n\times n}_+$, rewrite the objective function in (3.17) as

$$\varphi(X;\mu) := \vartheta(W;\mu) := \langle C, \widehat{X}W\rangle - \mu\ln\det(\widehat{X}W). \tag{3.24}$$

In order that the function $\ln\det(\widehat{X}W)$ makes sense, it is important consider $\vartheta(\cdot,\mu):\mathbb{W}_+\rightarrow\mathbb{R}$. It is easy to see that the "action" of the Fréchet derivative of ϑ on a tangent vector of $\mathbb{W}_+$, i.e., an element $\widehat{X}^{-1}Z\in\mathbb{W}$ with $Z\in S\mathbb{R}^{n\times n}$, assumes the form

$$\nabla\vartheta(W;\mu).(\widehat{X}^{-1}Z)=\langle C-\mu W^{-1}\widehat{X}^{-1},Z\rangle.$$

To retrieve the "gradient" of ϑ over $\mathbb{W}_+$ by the Riesz representation theorem, we have to take into account the inner product imposed upon $\mathbb{W}$. That is, by rewriting the relationships as

$$\nabla\vartheta(W;\mu).(\widehat{X}^{-1}Z)=[\![\nabla\vartheta(W;\mu),\widehat{X}^{-1}Z]\!]=[\![\widehat{X}^{-1}Z,\nabla\vartheta(W;\mu)]\!]=\langle Z,\nabla\vartheta(W;\mu)\widehat{X}^{-1}\rangle, \tag{3.25}$$

we are able to retrieve the gradient of ϑ as

$$\nabla\vartheta(W;\mu)=C\widehat{X}-\mu W^{-1}. \tag{3.26}$$

It is readily seen that

$$\mathscr{P}_{\mathcal{N}(\mathscr{A}\mathscr{L})}=I-\mathscr{L}^*\mathscr{A}^*(\mathscr{A}\mathscr{L}\mathscr{L}^*\mathscr{A}^*)^{-1}\mathscr{A}\mathscr{L} \tag{3.27}$$

is a projection onto the null space $\mathcal{N}(\mathscr{A}\mathscr{L})$ in $\mathbb{W}$. Analogous to (2.16), the steepest descent direction for $\vartheta(W;\mu)$ at $W=I$ is given by

$$\dot{W}|_I=-\mathscr{P}_{\mathcal{N}(\mathscr{A}\mathscr{L})}(C\widehat{X}-\mu I). \tag{3.28}$$

The corresponding pullback $\widehat{X}\dot{W}|_{W=I}$ stands for a descent direction for $\phi(X;\mu)$ at $X=\widehat{X}$. Since $\widehat{X}$ is an arbitrary interior feasible point, we can now describe a descent flow in $S\mathbb{R}^{n\times n}$ for $\phi(X;\mu)$ by

$$\dot{X}=-X(I-\mathscr{L}^*\mathscr{A}^*(\mathscr{A}\mathscr{L}\mathscr{L}^*\mathscr{A}^*)^{-1}\mathscr{A}\mathscr{L})(CX-\mu I). \tag{3.29}$$

Observe the resemblance of (3.29) to (2.17). Observe, in particular, how the diagonal matrix $X=\operatorname{diag}(\mathbf{x})$ is being replaced by $\mathscr{L}$ in the projection operator. We do caution readers, however, that there is a subtle dissimilarity in that matrix multiplications in $S\mathbb{R}^{n\times n}$ generally are not communicative and hence the order of multiplications in (3.29), such as the product CX, cannot be reversed.

By taking $\mu=0$, we obtain the affine scaling trajectory

$$\dot{X}=-X(I-\mathscr{L}^*\mathscr{A}^*(\mathscr{A}\mathscr{L}\mathscr{L}^*\mathscr{A}^*)^{-1}\mathscr{A}\mathscr{L})(CX). \tag{3.30}$$

for the SDP problem in the same way as we did for the LP problem. Again, observe the considerable similarity between (3.30) for the SDP and (2.19) for the LP. Note that the "scaling" for the SDP is no longer done by diagonal matrices such as that in the LP case, but the concept carries over.

Applying the adjoint $\mathscr{L}^*$ to both sides of $\mathscr{A}^*\mathbf{y}+S=C$ and using (3.22), we obtain the equality

$$\mathscr{P}_{\mathcal{N}(\mathscr{A}\mathscr{L})}CX=\mathscr{P}_{\mathcal{N}(\mathscr{A}\mathscr{L})}SX=\mu\mathscr{P}_{\mathcal{N}(\mathscr{A}\mathscr{L})}I.$$

Thus the affine scaling trajectory (3.30) can be expressed via the matrix representation (3.23) as

$$\dot{X}=-\mu X(I-\mathscr{A}^*\Omega^{-1}\mathscr{A}X^2)=-\mu(I-X\mathscr{A}^*\Omega^{-1}\mathscr{A}X)X. \tag{3.31}$$

Comparing the righthand sides of (3.10) and (3.31), we finally have reached our goal in concluding that, same as in the LP, if we choose $\dot{\mu}=-\mu^2$, then the primal central path $X(t)$ derived from the CPE for the SDP is identical to the generalized affine scaling flow derived from the projected gradient. The diagram sketched in Figure 2.1 is still relevant for the SDP.

4. Conclusion. We have used the same argument to justify that the primal central paths, that is, the nonnegative vector $\mathbf{x}(t)$ of the LP problem and the symmetric and positive definite matrix $X(t)$ of the SDP problem, are precisely the integral curves of the differential systems (2.19) and (3.31), respectively. It is interesting to note that the central path equations are algebraic in nature and are derived from the KKT optimality conditions. The affine scaling vector fields, on the other hand, are dynamical in nature and are results of projected gradient flows. Such a distinction dictates the methodologies for path-following [12].

One prevailing idea in the field of optimization is to solve approximately along the central path for the diminishing barrier parameter μ. Different ways to adopt this approach for efficiency have been proposed in the literature and a good many packages of software have been developed. Some known techniques include, for example, the Newton method, predictor-corrector schemes, or the preconditioned conjugate gradient methods [3, 7, 14, 15, 20]. On the other hand, since the central path can be cast as an integral curve, it is also possible to follow the path by concepts from numerical integration. We conclude this paper by briefly sketching the latter idea below with its application to the SDP problems in mind.

In an earlier work [9], a somewhat more complicated mechanism has been employed to derive the affine scaling trajectory for the SDP. The resulting dynamical system is characterized by

$$\dot{X} = -\underbrace{X\left(C - \sum_{i=1}^{m} u_i(X)A_i\right)X}_{D(X)}, \tag{4.1}$$

where the combination coefficients $u_i(X)$ are entries of the vector

$$\mathbf{u}(X) := \Omega^{-1}\begin{bmatrix} \langle A_1X, XC\rangle \\ \vdots \\ \langle A_mX, XC\rangle \end{bmatrix}. \tag{4.2}$$

Indeed, the very same vector field $D(X)$, derived from an essentially different notion of dual estimate, has been selected as an *affine scaling direction* and an iterative scheme of the form

$$X_{k+1} = X_k - \rho_k D(X_k) \tag{4.3}$$

has been employed as an affine scaling algorithm in [13]. The scheme (4.3) can be regarded as a variable-step explicit Euler method which is perhaps the simplest yet efficient maneuver along the course of integrating the differential system (4.1). The step size ρ_k is determined by either the so called short step strategy or the long step strategy to ensure feasibility and to move the iterates to the boundary. Rather surprisingly, however, an interesting example is given in [13] to show that the affine scaling algorithm in the form (4.3) fails to converge to the unique optimal solution by either strategy. Apparently, the step strategies have somehow lost track of the true affine scaling trajectory defined by (4.1).

We should point out that the vector field derived in (4.1) is the same as our result in (3.30). The only difference is that the system (3.30) is a lot more similar in appearance to that in (2.19) and is proved to be identical to the central path (3.10). This equivalence relationship can be seen by carrying out the following sequence of operations,

$$\begin{aligned}\mathscr{L}^*\mathscr{A}^*(\mathscr{A}\mathscr{L}\mathscr{L}^*\mathscr{A}^*)^{-1}\mathscr{A}\mathscr{L}(CX) = \mathscr{L}^*\mathscr{A}^*\Omega^{-1}\mathscr{A}(XCX) &= \mathscr{L}^*\mathscr{A}^*\Omega^{-1}\begin{bmatrix} \langle A_1, XCX\rangle \\ \vdots \\ \langle A_m, XCX\rangle \end{bmatrix} \\ &= \mathscr{L}^*\mathscr{A}^*\mathbf{u}(X) = \mathscr{L}^*(\sum_{i=1}^{m} u_i(X)A_i) = \sum_{i=1}^{m} u_i(X)A_iX.\end{aligned}$$

As we have shown in the preceding section that the affine scaling trajectory is in fact a gradient flow, perhaps what is being missed in the explicit Euler scheme such as (4.3) is the exploitation of this descent property.

Solving the LP or the SDP problems amounts to finding the asymptotically stable limit point of corresponding dynamical systems. So long as this limit point can be found effectively, the precision in tracking the solution flows of either (2.19) or (3.30) (which now is the same as (3.31) and (4.1)) is not essential. The question at hand is how we can take advantage of the fact that $\mathbf{x}(t)$ or $X(t)$ is a descent flow and propose a fast algorithm while preserving the same asymptotically stable limit point as the original system. One possible alternative to the discretization of the affine scaling trajectory is the pseudo-transit method [11]. The idea of this method is to mimic integration to steady-state while managing the time step to move the iteration as rapidly as possible to the Newton method. The method has been proved practical in finding steady-state solutions to time-dependent gradient flows which happen to be an inherent feature of the affine scaling trajectory. Its application to the affine scaling trajectory is currently under investigation.

REFERENCES

[1] I. Adler and R. D. C. Monteiro, *Limiting behavior of the affine scaling continuous trajectories for linear programming problems*, Math. Programming, 50 (1991), pp. 29–51.

[2] F. Alizadeh, *Interior point methods in semidefinite programming with applications to combinatorial optimization*, SIAM J. Optim., 5 (1995), pp. 13–51.

[3] A. Antoniou and W.-S. Lu, *Practical optimization, Algorithms and engineering applications*, Springer, New York, 2007.

[4] E. R. Barnes, *A variation on Karmarkar's algorithm for solving linear programming problems*, Math. Programming, 36 (1986), pp. 174–182.

[5] D. A. Bayer and J. C. Lagarias, *The nonlinear geometry of linear programming. I. Affine and projective scaling trajectories*, Trans. Amer. Math. Soc., 314 (1989), pp. 499–526.

[6] ———, *The nonlinear geometry of linear programming. II. Legendre transform coordinates and central trajectories*, Trans. Amer. Math. Soc., 314 (1989), pp. 527–581.

[7] S. Boyd and L. Vandenberghe, *Convex optimization*, Cambridge University Press, Cambridge, 2004.

[8] I. I. Dikin, *Iterative solution of problems of linear and quadratic programming*, Dokl. Akad. Nauk SSSR, 174 (1967), pp. 747–748.

[9] L. Faybusovich, *On a matrix generalization of affine-scaling vector fields*, SIAM J. Matrix Anal. Appl., 16 (1995), pp. 886–897.

[10] D. Goldfarb and K. Scheinberg, *Interior point trajectories in semidefinite programming*, SIAM J. Optim., 8 (1998), pp. 871–886 (electronic).

[11] C. T. Kelley, L.-Z. Liao, L. Qi, M. T. Chu, J. P. Reese, and C. Winton, *Projected pseudotransient continuation*, SIAM J. Numer. Anal., 46 (2008), pp. 3071–3083.

[12] N. Megiddo, *Pathways to the optimal set in linear programming*, in Progress in mathematical programming (Pacific Grove, CA, 1987), Springer, New York, 1989, pp. 131–158.

[13] M. Muramatsu, *Affine scaling algorithm fails for semidefinite programming*, Math. Programming, 83 (1998), pp. 393–406.

[14] Y. Nesterov and A. Nemirovskii, *Interior-point polynomial algorithms in convex programming*, vol. 13 of SIAM Studies in Applied Mathematics, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1994.

[15] M. J. Todd, *Semidefinite optimization*, Acta Numer., 10 (2001), pp. 515–560.

[16] R. J. Vanderbei and J. C. Lagarias, *I. I. Dikin's convergence result for the affine-scaling algorithm*, in Mathematical developments arising from linear programming (Brunswick, ME, 1988), vol. 114 of Contemp. Math., Amer. Math. Soc., Providence, RI, 1990, pp. 109–119.

[17] H. Wolkowicz, *Bibliography on semidefinite programming*, Department of Combinatorics and Optimization, University of Waterloo. `http://liinwww.ira.uka.de/bibliography/Math/psd.html`.

[18] H. Wolkowicz, R. Saigal, and L. Vandenberghe, eds., *Handbook of semidefinite programming*, International Series in Operations Research & Management Science, 27, Kluwer Academic Publishers, Boston, MA, 2000. Theory, algorithms, and applications.

[19] S. J. Wright, *Primal-dual interior-point methods*, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1997.

[20] Y. Ye, *Interior Point Algorithms, Theory and Analysis*, Wiley-Interscience Series in Discrete Mathematics and Optimization, John Wiley & Sons Inc., New York, 1997.