TUTORIALS IN OPERATIONS RESEARCH
INFORMS 2009

informs

© 2009 INFORMS | ISBN 978-1-877640-24-7
DOI 10.1287/educ.1090.0061

Twenty-Five Years of Interior Point Methods

Tamás Terlaky
Department of Industrial and Systems Engineering, Lehigh University,
Bethlehem, Pennsylvania 18015, terlaky@lehigh.edu

Abstract The 1984 paper of Narendra K. Karmarkar [N. K. Karmarkar. A new polynomial-time algorithm for linear programming. *Combinatorica* 4:373–395, 1984] launched the age of interior point methods (IPMs). Hundreds of polynomial time IPMs were designed in the past quarter century. Interior point software implementations have challenged simplex method implementations and frequently surpassed their performance. All aspects of linear optimization have had to be revisited. Duality theory and the significance and applications of a strictly complementary solution have been explored, and novel concepts of sensitivity analysis have been introduced. We show the limitations of IPMs and present some conjectures and open problems. Polynomial time IPMs have been generalized to smooth convex optimization problems, and IPMs have been successfully implemented to solve general nonlinear optimization problems as well. New problem classes, such as second-order conic and semidefinite optimization problems, are now efficiently solvable by IPMs. Novel paradigms, such as the Ben-Tal–Nemirovski [A. Ben-Tal and A. Nemirovski. *Lectures on Modern Convex Optimization: Analysis, Algorithms, and Engineering Applications*. MPS-SIAM Series on Optimization, SIAM, Philadelphia, 2001] robust optimization methodology, opened never-seen opportunities to solve large important classes of optimization problems, including truss-topology design and robust radiation therapy treatment. As IPMs have spread to all optimization areas, the theory and practice of optimization has changed forever. Twenty-five years after the publication of Karmarkar's path-breaking paper, this tutorial attempts to give a glimpse of the seminal results induced by the "interior point revolution"—as Margaret H. Wright [M. H. Wright. The interior point revolution in optimization: History, recent developments, and lasting consequences. *Bulletin of the American Mathematical Society* 42(1):39–56, 2004] coined the term.

Keywords interior point methods; Newton's method; complexity; sensitivity analysis; linear optimization; Klee-Minty cube; Hirsch conjecture; nonlinear optimization

1. Introduction

Since the path-breaking paper of Karmarkar [48], a tremendous amount of research has been devoted to the study of interior point methods (IPMs). By now, IPMs are among the most efficient methods for solving linear and also wide classes of general convex optimization problems. Hundreds of polynomial time IPM variants were developed for linear optimization (LO). Interior point software implementations have challenged simplex method implementations and frequently surpassed their performance (Bixby [13]). All aspects of linear optimization had to be revisited. Not only the research on algorithms exploded, but research on duality theory (Roos et al. [75], Terlaky [82]) and implementation strategies (Gondzio and Terlaky [38], Andersen et al. [7]) got new impulse. The significance and applications of a strictly complementary solution were explored (Greenberg [40]), and novel concepts of sensitivity analysis were introduced (Jansen et al. [46], Koltai and Terlaky [52], Ghaffari Hadigheh et al. [35], Roos et al. [75]).

A decade of turbulent research developed a good understanding of the fundamentals of IPMs. From 1994 on, several books were published that summarize and explore different aspects of IPMs. den Hertog [26] gives a thorough survey of primal and dual path-following IPMs for linear and structured convex optimization problems. The seminal work of Nesterov and Nemirovski [64] provides the most general framework for polynomial time IPMs for smooth convex optimization problems. Jansen [44] discusses primal-dual target-following algorithms for linear optimization and complementarity problems. Saigal's [76] book deals with affine scaling algorithms. The volume edited by Terlaky [81] contains 13 survey papers that cover most of the key aspects of IPMs, their extensions, and some applications. Wright [94] concentrates on primal-dual infeasible IPMs, numerical issues, and local and asymptotic convergence properties. The book of Ye [96] is a rich source of polynomial IPMs not only for LO, but for convex optimization problems as well. It extends the IPM theory to derive bounds and approximations for classes of nonconvex optimization problems as well. Finally, Roos et al. [75] present a thorough treatment of the IPM-based theory—duality, initialization, complexity, sensitivity analysis, and implementation strategies—and large classes of IPMs for LO.

The best complexity bound to date, $O(n^3L)$ arithmetic operations,[1] was already reached in 1988 (Gonzaga [39], Renegar [71], Roos [73]). The following years enriched theory and extended the class of polynomial time solvable problems, and high-performance software tools were developed, but the best complexity result for LO remained the same. Research started to explore the worst-case behavior of IPMs to derive lower bounds for the number of iterations (Todd and Ye [83], Vavasis and Ye [89]) and explore extreme behavior of the central path (Deza et al. [31]). Deza et al. [27] demonstrated that the $O(\sqrt{n}\log((x^0)^T s^0/\epsilon))$ iteration complexity bound is tight. In subsequent work, Deza et al. [30, 29] drew intriguing analogies between worst- and average-case behaviors of edge paths followed by simplex algorithms, and the central path followed by IPMs.

In the following sections, while building the duality theory of LO based on IPM concepts, without going into much detail, the major concepts and results of the standard central-path-following IPMs are summarized.

IPMs not only provide polynomial time algorithms and a powerful methodology for LO, but they are generalized to solve large classes of optimization problems. IPMs were generalized to solve *conic linear*—in particular, *semidefinite* and *second-order conic*—optimization problems (Alizadeh [2], Alizadeh and Goldfarb [4], Nesterov and Nemirovski [64]). The IPM methodology allowed to solve large classes of novel engineering and control problems (Ben-Tal and Nemirovski [12], Boyd and Vandenberghe [16], Pólik and Terlaky [68]), as well as solve, or approximately solve, combinatorial optimization problems (Alizadeh [2, 3], Goemans and Williamson [36]). A rich theory of IPMs for smooth nonlinear optimization (NLO) problems was developed from the early days of IPMs (Nesterov and Nemirovski [64], Renegar [72]). A survey of the NLO results, mostly focusing on the conic case, can be found in Pólik and Terlaky [69]. Although the basic concepts in the extensions and generalizations of IPMs to more general problem classes are similar to the case of LO, some details, most importantly, the analysis and the results, are somewhat—and sometimes quite—different.

1.1. Structure of this Paper

This paper is structured as follows. First, in §2.1 we briefly review the general LO problem in canonical form and discuss how Goldman and Tucker's (Goldman and Tucker [37], Tucker [85]) self-dual and homogeneous model is derived. In §2.2, the Goldman-Tucker theorem, i.e., the existence of a strictly complementary solution for the skew-symmetric self-dual model, is proved. Here, such fundamental IPM objects as the existence of an interior solution, the central path, the Newton step, and the analytic center of polytopes are introduced.

[1] Here, L denotes the binary length of the input date. In reporting complexity results it is frequently replaced by certain condition numbers of the problem, or in IPMs, by the logarithm of the ratio of the initial and terminal duality gap. The $O(n^3L)$ arithmetic operations bound comes from an $O(\sqrt{n}L)$ iterations complexity bound, where the cost of one iteration is $O(n^{2.5})$ arithmetic operations resulting from a partial update scheme when solving the Newton system.

We show that the central path converges to a strictly complementary solution and that an exact strictly complementary solution for LO, or a certificate of infeasibility, can be obtained after a finite number of iterations. Our theoretical development is summarized in §2.3. Then, in §2.4 a general scheme of IPMs is presented with some concrete realizations. In §3 we exploit the effect of redundancy on the central path and show that by systematically adding exponentially many redundant constraints, the central path may visit all the vertices of the Klee-Minty cube in the same order as simplex methods do. Most notably, these constructions allow us to demonstrate that the iteration complexity bound of central-path-following IPMs is tight. In §4 we draw analogies between the diameter and the curvature of polytopes and formulate some intriguing conjectures that are inspired by the Hirsch conjecture. In §5 we briefly review the results regarding the extensions of IPMs to convex conic and nonlinear optimization and make some notes on available software, both commercial and open source. Finally, some conclusions are presented in §6.

1.2. Notation

In this paper, $\mathbb{R}^n_+$ denotes the set of nonnegative vectors in $\mathbb{R}^n$. Throughout, we use $\|\cdot\|_p$ ($p \in \{1, 2, \infty\}$) to denote the p-norm on $\mathbb{R}^n$, with $\|\cdot\|$ denoting the Euclidean norm $\|\cdot\|_2$. The identity matrix is denoted by I, and e denotes the vector that has all of its components equal to one. Given an n-dimensional vector x, the matrix X denotes the $n \times n$ diagonal matrix whose diagonal entries are the coordinates x_j of x. If $x,\ s \in \mathbb{R}^n$, then $x^T s$ denotes the dot product of the two vectors. Furthermore, xs and x^α for $\alpha \in \mathbb{R}$ and $\max\{x, y\}$ denote the vectors resulting from coordinatewise operations. For any matrix $A \in \mathbb{R}^{m\times n}$, A_j denotes the jth column of A. Furthermore,

$$\pi(A) := \prod_{j=1}^{n} \|A_j\|.$$

For any index set $J \subseteq \{1, 2, \ldots, n\}$, $|J|$ denotes the cardinality of J, and $A_J \in \mathbb{R}^{m\times|J|}$ the submatrix of A whose columns are indexed by the elements in J. Moreover, if $K \subseteq \{1, 2, \ldots, m\}$, $A_{KJ} \in \mathbb{R}^{|K|\times|J|}$ is the submatrix of A_J whose rows are indexed by the elements in K.

Vectors are assumed to be column vectors. The (vertical) concatenation of two vectors (or matrices of appropriate size) u and v is denoted by $(u; v)$, whereas the horizontal concatenation is (u, v).

2. IPMs for LO

After years of intensive research, a deep understanding of IPMs has been developed. There are easy-to-understand, simple variants of polynomial IPMs. The self-dual embedding strategy (Jansen et al. [45], Roos et al. [75], Terlaky [82], Ye et al. [97]) provides an elegant solution for the initialization problem of IPMs. It is also possible to build not only the complete duality theory of Roos et al. [75] of LO but to perform sensitivity analysis (Jansen [44], Jansen et al. [46], Monteiro and Mehrotra [60], Roos et al. [75]) on the basis of IPMs. We also demonstrate that IPMs not only converge to an optimal solution (if it exists) but after a finite number of iterations also allow a strongly polynomial rounding procedure (Mehrotra and Ye [59], Roos et al. [75]) to generate exact solutions. This all requires only the knowledge of elementary calculus and can be taught not only at a graduate level but at an advanced undergraduate level as well. Our aim is to present such an approach, based on the one presented in Roos et al. [75] and Terlaky [82].

2.1. The Linear Optimization Problem: Duality

This section is based on Terlaky [82] and Part I and the appendix in Roos et al. [75]. In this section we give a short introduction to some basic concepts of path-following IPMs and briefly indicate how the duality theory of LO may be built in an elementary way, purely on fundamental concepts of IPMs.

We consider the general LO problem (P) and its dual (D) in canonical form:

$$\min\{c^T x\colon Ax \geq b,\, x \geq 0\}, \tag{P}$$

$$\max\{b^T y\colon A^T y \leq c,\, y \geq 0\}, \tag{D}$$

where A is an $m \times k$ matrix, $b, y \in \mathbb{R}^m$, and $c, x \in \mathbb{R}^k$. It is well known that by using only elementary transformations, any given LO problem can easily be transformed into a "minimal" canonical form. These transformations can be summarized as follows:

- Introduce slack variables to get equations (if a variable has both lower and upper bounds, then one of these bounds is transformed to an inequality constraint);
- Shift the variables with the lower or upper bound so that the respective bound becomes 0 and, if needed, replace the variable by its negative;
- Eliminate free variables; and
- Use Gaussian elimination to transform the problem into a form where all equations have a singleton column (i.e., choose a basis and multiply the equations by the inverse basis), whereas dependent constraints are eliminated.

The weak duality theorem sets the stage for establishing the fundamental duality relationship between the primal problem (P) and its dual (D), and it is easily proved for LO problems given in the canonical form.

Theorem 1 (Weak Duality for Linear Optimization). *Let us assume that $x \in \mathbb{R}^k$ and $y \in \mathbb{R}^m$ are feasible solutions for the primal problem* (P) *and dual problem* (D), *respectively. Then one has*

$$c^T x \geq b^T y,$$

where equality holds if and only if[2]

(i) $x_i(c - A^T y)_i = 0$ *for all* $i = 1, \ldots, k$; *and*
(ii) $y_j(Ax - b)_j = 0$ *for all* $j = 1, \ldots, m$.

Proof. Using primal and dual feasibility of x and y, we may write

$$(c - A^T y)^T x \geq 0 \quad \text{and} \quad y^T(Ax - b) \geq 0,$$

with equality if and only if (i), respectively (ii), holds. Summing up these two inequalities we have the desired inequality

$$0 \leq (c - A^T y)^T x + y^T(Ax - b) = c^T x - b^T y.$$

The theorem is proved. □

The weak duality theorem implies the following sufficient optimality condition.

Corollary 1. *Let a primal and dual feasible solution $x \in \mathbb{R}^k$ and $y \in \mathbb{R}^m$ with $c^T x = b^T y$ be given. Then x is an optimal solution of the primal problem* (P), *and y is an optimal solution of the dual problem* (D). □

Although the weak duality theorem provides a sufficient condition to check optimality of a feasible solution pair, it does not prove that for every feasible primal-dual LO problem pair an optimal pair of solutions with zero duality gap exists. This result is the so-called

[2] These conditions are in general referred to as the *complementarity conditions*. Using the coordinatewise notation we may write $x(c - A^T y) = 0$ and $y(Ax - b) = 0$. By the weak duality theorem, complementarity and feasibility imply optimality.

strong duality theorem that we prove in the next sections by using only simple calculus and fundamental concepts of IPMs.

Because, optimistically, we are looking for optimal solutions with zero duality gap, we utilize the results of the weak duality theorem (Theorem 1). Thus, we need to find a solution of the inequality system formed by the primal and the dual feasibility constraints and by requiring that the dual objective is at least as large as the primal one. For this system we already know that any solution of this system gives both a primal and a dual feasible solution with equal objective values; thus, by Corollary 1, these solutions are optimal. By introducing slack variables, the inequality system is as follows:

$$\begin{aligned} Ax - z = b, &\quad x \geq 0,\, z \geq 0; \\ A^T y + s = c, &\quad y \geq 0,\, s \geq 0; \\ b^T y - c^T x - \rho = 0, &\quad \rho \geq 0. \end{aligned}$$

By homogenizing the inequality system, the *Goldman-Tucker model* (Goldman and Tucker [37], Tucker [85]) is obtained.

$$\begin{aligned} Ax - \tau b - z = 0, &\quad x \geq 0,\, z \geq 0; \\ -A^T y + \tau c - s = 0, &\quad y \geq 0,\, s \geq 0; \\ b^T y - c^T x - \rho = 0, &\quad \tau \geq 0,\, \rho \geq 0. \end{aligned} \tag{1}$$

It is easy to verify that for any solution (y, x, τ, z, s, ρ) of the Goldman-Tucker system (1), $\tau\rho > 0$ cannot hold. Indeed, if $\tau\rho$ were positive, then the we would have

$$0 < \tau\rho = \tau b^T y - \tau c^T x = x^T A y - z^T y - y^T A^T x - s^T x = -z^T y - s^T x \leq 0,$$

yielding a contradiction.

Furthermore, any optimal pair (x, y) of problems (P) and (D) with zero duality gap is a solution of the Goldman-Tucker system with $\tau = 1$ and $\rho = 0$. It is also clear that the homogeneous Goldman-Tucker system admits the trivial zero solution, but that has no value for solving (P) and (D). We are looking for nontrivial solutions with $\tau > 0$, because those give primal and dual optimal solutions $(x/\tau, y/\tau)$ with zero duality gap, as $\tau > 0$ implies $\rho = 0$.

On the other hand, if the Goldman-Tucker system admits a nontrivial feasible solution $(\bar{y}, \bar{x}, \bar{\tau}, \bar{z}, \bar{s}, \bar{\rho})$ with $\bar{\tau} = 0$ and $\bar{\rho} > 0$, then we may conclude that either (P) or (D), or both (P) and (D) are infeasible. Indeed, $\bar{\tau} = 0$ implies that $A\bar{x} \geq 0$ and $A^T\bar{y} \leq 0$. Furthermore, if $\bar{\rho} > 0$, then we have either $b^T\bar{y} > 0$ or $c^T\bar{x} < 0$, or both. If $b^T\bar{y} > 0$, then by assuming that there is a feasible solution $x \geq 0$ of (P), we have

$$0 < b^T\bar{y} \leq x^T A^T \bar{y} \leq 0,$$

which is a contradiction. Thus, (P) must be infeasible. Similarly, if $c^T\bar{x} < 0$, then by assuming that there is a dual feasible solution $y \geq 0$ for (D), we have

$$0 > c^T\bar{x} \geq y^T A\bar{x} \geq 0,$$

which is a contradiction. Thus, if $c^T\bar{x} > 0$, then (D) must be infeasible.

Summarizing these results, we have the following theorem.

Theorem 2. *Let a primal dual pair* (P) *and* (D) *of* LO *problems be given. The following statements hold for the solutions of the Goldman-Tucker system* (1).

(1) *Any optimal pair* (x, y) *of* (P) *and* (D) *with zero duality gap is a solution of the corresponding Goldman-Tucker system with* $\tau = 1$.

(2) *If (y, x, τ, z, s, ρ) is a solution of the Goldman-Tucker system, then either $\tau = 0$ or $\rho = 0$, or both; i.e., $\tau\rho > 0$ cannot happen.*

(3) *Any solution (y, x, τ, z, s, ρ) of the Goldman-Tucker system, where $\tau > 0$ and $\rho = 0$, gives a primal and dual optimal pair $(x/\tau, y/\tau)$ with zero duality gap.*

(4) *If the Goldman-Tucker system admits a feasible solution $(\bar{y}, \bar{x}, \bar{\tau}, \bar{z}, \bar{s}, \bar{\rho})$ with $\bar{\tau} = 0$ and $\bar{\rho} > 0$, then we may conclude that either* (P) *or* (D) *or both of them are infeasible.* □

As will be shown in this section, the interior point approach and, ultimately, IPMs lead us to a solution of the Goldman-Tucker system where either $\tau > 0$ or $\rho > 0$. Thus, the IPM approach ensures that $\tau + \rho > 0$. This is a significant advantage of IPMs over simplex methods.

To establish the desired duality results, we give a simple form of the Goldman-Tucker model and simplify our notations. Observe that the Goldman-Tucker system can be written in the following compact form:

$$Mu \geq 0, \quad u \geq 0, \quad w(u) = Mu, \tag{2}$$

where

$$u = \begin{pmatrix} y \\ x \\ \tau \end{pmatrix}, \quad w(u) = \begin{pmatrix} z \\ s \\ \rho \end{pmatrix}, \quad \text{and} \quad M = \begin{pmatrix} 0 & A & -b \\ -A^T & 0 & c \\ b^T & -c^T & 0 \end{pmatrix}$$

is a skew-symmetric $n \times n$ matrix; i.e., $M^T = -M$ and $n = m + k + 1$. The Goldman-Tucker theorem (Goldman and Tucker [37], Roos et al. [75], Tucker [85]) states that system (2) admits a *strictly complementary* solution that is a feasible solution for which $uw(u) = 0$ and $u + w(u) > 0$. This is the key result to prove the strong duality theorem. An elementary proof is given in the next subsection.

Theorem 3 (Goldman-Tucker). *System (2) has a strictly complementary feasible solution, i.e., a solution for which $u + w(u) > 0$.*

This theorem ensures that either Case (3) or Case (4) of Theorem 2 must occur when one solves the Goldman-Tucker system of LO. This is in fact the strong duality theorem of LO.

Theorem 4 (Strong Duality for Linear Optimization). *Let a primal and dual* LO *problem be given. Exactly one of the following statements holds:*

- *Both* (P) *and* (D) *are feasible, and there are optimal solutions x^* and y^* such that*

$$c^T x^* = b^T y^*.$$

- *Either problem* (P) *or* (D) *or both of them are infeasible.*

Proof. Theorem 3 implies that the Goldman-Tucker system of the LO problem admits a strictly complementary solution. Thus, in such a solution, if $\tau > 0$ then statement (3) of Theorem 2 implies the existence of an optimal pair with zero duality gap. On the other hand, when $\tau = 0$, then $\rho > 0$; thus statement (4) of Theorem 2 proves that either (P) or (D) or both of them are infeasible. □

Our next task is to sketch an elementary constructive proof of Theorem 3 that completes the duality theory for LO.

2.2. The Skew-Symmetric Self-Dual Embedding Model

2.2.1. Interior Point Solution for the Goldman-Tucker Model. Recall that because of the second statement of Theorem 2, system (2) cannot have a solution with $\tau > 0$ and $\rho > 0$, that is, it cannot have a solution where all the variables are strictly positive. The existence of a solution with all variables strictly positive is called the *interior point condition*[3] (*IPC*). To prove the Goldman-Tucker theorem (Theorem 3), one need modify problem (2) so that the resulting equivalent problem satisfies the IPC.

Self-Dual Embedding of* (2) *with Interior Point Solution. Problem (2) is a homogeneous feasibility problem, and we transform it into an equivalent problem that satisfies the IPC. This happens by embedding the problem and defining an appropriate nonnegative vector q that will be both the right-hand side and the objective vector in the larger LO problem.

Let us take $u = w(u) = e$. These vectors are positive, but they do not satisfy the equality conditions in (2). Let us further define the error vector r obtained this way by

$$r := e - Me, \quad \text{and let} \quad \lambda := n + 1.$$

Then we have

$$\begin{pmatrix} M & r \\ -r^T & 0 \end{pmatrix} \begin{pmatrix} e \\ 1 \end{pmatrix} + \begin{pmatrix} 0 \\ \lambda \end{pmatrix} = \begin{pmatrix} Me + r \\ -r^T e + \lambda \end{pmatrix} = \begin{pmatrix} e \\ 1 \end{pmatrix}.$$

Hence, the following problem

$$\min\left\{ \lambda\vartheta\colon -\begin{pmatrix} M & r \\ -r^T & 0 \end{pmatrix} \begin{pmatrix} u \\ \vartheta \end{pmatrix} + \begin{pmatrix} w \\ \nu \end{pmatrix} = \begin{pmatrix} 0 \\ \lambda \end{pmatrix};\ \begin{pmatrix} u \\ \vartheta \end{pmatrix}, \begin{pmatrix} w \\ \nu \end{pmatrix} \geq 0 \right\} \tag{$\overline{\text{SP}}$}$$

satisfies the IPC because the all-one vector is feasible for this problem. This problem is a self-dual skew-symmetric LO problem, because by introducing the notation

$$\bar{M} = \begin{pmatrix} M & r \\ -r^T & 0 \end{pmatrix}, \quad \bar{u} = \begin{pmatrix} u \\ \vartheta \end{pmatrix}, \quad \text{and} \quad \bar{q} = \begin{pmatrix} 0 \\ \lambda \end{pmatrix},$$

one can write problem ($\overline{\text{SP}}$) as

$$\min\{\bar{q}^T \bar{u}\colon \bar{M}\bar{u} \geq -\bar{q},\ \bar{x} \geq 0\}. \tag{SPe}$$

We claim that finding a strictly complementary solution to (2) is equivalent to finding a strictly complementary optimal solution to problem (SPe). This claim is valid because (SPe) satisfies the IPC, and thus, as we will see, its optimal value is zero. Moreover, it admits a strictly complementary optimal solution. Because the objective function is just a constant multiple of ϑ, this variable must be zero in any optimal solution (see also Lemma 1). This observation implies the claimed result.

Conclusion. Every LO problem can be embedded in a self-dual problem ($\overline{\text{SP}}$) of the form of (SPe). This can be done in such a way that $\bar{u} = e$ is feasible for ($\overline{\text{SP}}$) and $\bar{w}(\bar{u}) = e$. Having a strictly complementary solution of (SPe), we either find an optimal solution of the embedded LO problem, or we can conclude that the LO problem does not have an optimal solution.

After constructing the embedding problem, we continue to study our skew-symmetric self-dual problem (SPe) by utilizing that the IPC holds for this LO problem.

2.2.2. Basic Properties of the Skew-Symmetric Self-Dual Model. For ease of discussion, let us simplify our notation. Let the embedding problem (SPe) of our skew-symmetric model (2) be given as

$$\min\{q^T u\colon Mu \geq -q,\ u \geq 0\}, \tag{SP}$$

[3] In general terms, the interior point condition requires a feasible solution for which all inequalities hold strictly.

where the matrix $M \in \mathbb{R}^{n\times n}$ is *skew symmetric*, $0 \le q \in \mathbb{R}^n$, and problem (SP) satisfies the IPC; i.e., there exists an $u > 0$ with $Mu > 0$. The set of feasible solutions of (SP) is denoted by

$$\mathrm{SP} := \{u\colon\, u \ge 0,\, Mu \ge -q\}.$$

By using the assumption that the coefficient matrix M is skew symmetric and the right-hand-side vector $-q$ is the negative of the objective coefficient vector, one easily verifies that the dual of (SP) is equivalent to (SP) itself; i.e., problem (SP) is *self-dual*. Because of the self-dual property the following result is obvious.

Lemma 1. *The optimal value of* (SP) *is zero, and* (SP) *admits the zero vector* $u = 0$ *as a feasible and optimal solution.*

Given $(u, w(u))$, where $w(u) = Mu + q$, we may write

$$0 \le q^T u = u^T(w(u) - Mu) = u^T w(u) = e^T(uw(u));$$

i.e., for any optimal solution, $e^T(uw(u)) = 0$, implying that the vectors u and $w(u)$ are complementary. For further use, the *optimal set* of (SP) is denoted by

$$\mathrm{SP}^* := \{u\colon\, u \ge 0,\, Mu + q = w(u) \ge 0,\, uw(u) = 0\}.$$

Optimal solutions of the self-dual model are complementary in the general sense; i.e., they are not only complementary with respect to (w.r.t.) their own slack vector, but complementary w.r.t. the slack vector for any other optimal solution as well.

Lemma 2. *Let u and v be feasible for* (SP)*. Then, u and v are optimal if and only if*

$$uw(v) = vw(u) = uw(u) = vw(v) = 0.$$

Proof. Because M is skew symmetric, we have $(u - v)^T M(u - v) = 0$, which implies that $(u - v)^T(w(u) - w(v)) = 0$. Hence, $u^T w(v) + v^T w(u) = u^T w(u) + v^T w(v)$, and this vanishes if and only if u and v are optimal. □

All of the results so far, including to find a trivial optimal solution, are straightforward for (SP). The only nontrivial result that we need to prove is the existence of a strictly complementary solution. Under the IPC, this result can be proved in an elementary way.

2.2.3. The Central Path and the Optimal Partition of (SP). Let $u \in \mathrm{SP}$ and $w = w(u)$ be a *feasible pair*. Because of self-duality, the *duality gap* for this pair is twice the value

$$q^T u = u^T w;$$

however, for the sake of simplicity, the quantity $q^T u = u^T w$ itself will be referred to as *the duality gap*. The following result shows that the IPC not only implies the boundedness of the level sets, but the converse is also true. This result is presented without proof; for a proof the reader is referred to Roos et al. [75] and Terlaky [82].

Lemma 3. *Let* (SP) *be feasible. Then, the following statements are equivalent:*

(i) *The interior point condition is satisfied;*
(ii) *The level sets of $u^T w$ are bounded; and*
(iii) *The optimal set* SP^* *of* (SP) *is bounded.*

The *central path* plays a crucial role in path-following IPMs. Next we define the central path (Fiacco and McCormick [32], Frisch [33], Megiddo [55], Sonnevend [77]) of (SP).

Definition 1. Let the IPC be satisfied. The set of solutions

$$\{(u(\mu), w(u(\mu)))\colon\, Mu + q = w,\, uw = \mu e,\, u > 0 \text{ for some } \mu > 0\}$$

is called the *central path* of (SP).

If no confusion is possible, instead of $w(u(\mu))$ the notation $w(\mu)$ will be used. Now we are ready to present our main theorem. This, in fact, establishes the existence of the central path. For an elementary proof using only calculus, see Terlaky [82].

Theorem 5. *The next statements are equivalent.*

(i) (SP) *satisfies the interior point condition.*
(ii) *For each* $0 < \mu \in \mathbb{R}$ *there exists* $(u(\mu), w(\mu)) > 0$ *such that*

$$\begin{aligned} Mu + q &= w, \\ uw &= \mu e. \end{aligned}$$

(iii) *For each* $0 < p \in \mathbb{R}^n$ *there exists* $(u, w) > 0$ *such that*

$$\begin{aligned} Mu + q &= w, \\ uw &= p. \end{aligned}$$

Moreover, the solutions of these systems are unique.

The Newton Step. In proving the existence of a solution for the systems in (ii) and (iii), the main tool is a careful analysis of the Newton step (Terlaky [82]) when applied to the nonlinear systems in (iii).[4]

Let a vector $(u, w) > 0$ with $w = Mu + q$ be given. For a particular $0 < p \in \mathbb{R}^n$, one wants to find the displacement $(\Delta u, \Delta w)$ that solves

$$\begin{aligned} M(u + \Delta u) + q &= w + \Delta w, \\ (u + \Delta u)(w + \Delta w) &= p. \end{aligned}$$

Using feasibility of the actual solution, this system reduces to

$$\begin{aligned} M\Delta u &= \Delta w, \\ u\Delta w + w\Delta u + \Delta u \Delta w &= p - uw. \end{aligned}$$

If in this nonlinear equation system, we neglect the second-order term $\Delta u \Delta w$; then the *Newton equation*

$$\begin{aligned} M\Delta u &= \Delta w, \\ u\Delta w + w\Delta u &= p - uw \end{aligned} \tag{3}$$

is obtained. This is a linear equation system, and the reader easily verifies that the *Newton direction* Δu is the solution of the nonsingular system of equations[5]

$$(M + U^{-1}W)\Delta u = u^{-1}p - w.$$

When we perform a step in the Newton search direction with step length α, for the new solutions $(u^+, w^+) = (u + \alpha\Delta u, w + \alpha\Delta w)$ we have

$$\begin{aligned} u^+ w^+ &= (u + \alpha\Delta u)(w + \alpha\Delta w) = uw + \alpha(u\Delta w + w\Delta u) + \alpha^2 \Delta u \Delta w \\ &= uw + \alpha(p - uw) + \alpha^2 \Delta u \Delta w. \end{aligned}$$

[4] Observe that no preliminary knowledge on any variants of Newton's method is assumed. The Newton step is defined and analyzed for the particular situation. The reader is referred to Terlaky [82] for complete details of the proof.

[5] Nonsingularity follows from the fact that the sum of a skew-symmetric—thus, positive semidefinite—and a positive definite matrix is positive definite. Although it is not advised to use for numerical computations, the Newton direction can be expressed as $\Delta u = (M + U^{-1}W)^{-1}(u^{-1}p - w)$.

The next important concept is the notion of the optimal partition. The partition $(\mathbf{B},\mathbf{N})$ of the index set $\{1,\dots,n\}$ given by

$$\mathbf{B} := \{i\colon u_i > 0, \text{ for some } u \in \mathrm{SP}^*\},$$
$$\mathbf{N} := \{i\colon w(u)_i > 0, \text{ for some } u \in \mathrm{SP}^*\}$$

is called the *optimal partition*. By Lemma 2, the sets $\mathbf{B}$ and $\mathbf{N}$ are disjoint. Our main result says that the central path converges to a strictly complementary optimal solution, and this result proves that $\mathbf{B}\cup\mathbf{N} = \{1,\dots,n\}$. When this result is established, the Goldman-Tucker theorem (Theorem 3) for the general LO problem is proved, because we use the embedding method presented in §2.2.1.

Theorem 6. *If the IPC holds, then there exists an optimal solution u^* and $w^* = w(u^*)$ of problem (SP) such that $u^*w^* = 0$, $u^*_{\mathbf{B}} > 0$, $w^*_{\mathbf{N}} > 0$, and $u^* + w^* > 0$.*

As we mentioned earlier, this result is powerful enough to prove the strong duality theorem of LO in the strong form, including strict complementarity, i.e., the Goldman-Tucker theorem (Theorem 3) for (SP) and thus for (P) and (D).

Our next step is to prove that the accumulation point $u(\mu) \to u^*$ as $\mu \searrow 0$ is unique and strictly complementary.

2.2.4. Convergence to the Analytic Center. In this subsection we show that the central path has only one limit point; i.e., it converges to a unique point; the so-called analytic center (Sonnevend [77]) of the optimal set SP^*.

Definition 2. Let $\bar u \in \mathrm{SP}^*$, $\bar w = w(\bar u)$ maximize the product

$$\prod_{i\in\mathbf{B}} u_i \prod_{i\in\mathbf{N}} w_i$$

over $u \in \mathrm{SP}^*$. Then, $\bar u$ is called the *analytic center* of SP^*.

It is easy to verify that the analytic center is unique. Let us assume to the contrary that there are two different vectors $\bar u \neq \tilde u$ with $\bar u, \tilde u \in \mathrm{SP}^*$ that satisfy the definition of the analytic center; i.e.,

$$\vartheta^* = \prod_{i\in\mathbf{B}} \bar u_i \prod_{i\in\mathbf{N}} \bar w_i = \prod_{i\in\mathbf{B}} \tilde u_i \prod_{i\in\mathbf{N}} \tilde w_i = \max_{u\in\mathrm{SP}^*} \prod_{i\in\mathbf{B}} u_i \prod_{i\in\mathbf{N}} w_i.$$

Let us define $u^* = (\bar u + \tilde u)/2$. Then we have

$$\begin{aligned}
\prod_{i\in\mathbf{B}} u_i^* \prod_{i\in\mathbf{N}} w_i^* &= \prod_{i\in\mathbf{B}} \frac{1}{2}(\bar u_i + \tilde u_i) \prod_{i\in\mathbf{N}} (\bar w_i + \tilde w_i)\\
&= \prod_{i\in\mathbf{B}} \frac{1}{2}\left(\sqrt{\frac{\bar u_i}{\tilde u_i}} + \sqrt{\frac{\tilde u_i}{\bar u_i}}\right) \prod_{i\in\mathbf{N}} \frac{1}{2}\left(\sqrt{\frac{\bar w_i}{\tilde w_i}} + \sqrt{\frac{\tilde w_i}{\bar w_i}}\right) \sqrt{\prod_{i\in\mathbf{B}} \bar u_i \prod_{i\in\mathbf{N}} \bar w_i \prod_{i\in\mathbf{B}} \tilde u_i \prod_{i\in\mathbf{N}} \tilde w_i}\\
&> \prod_{i\in\mathbf{B}} \bar u_i \prod_{i\in\mathbf{N}} \bar w_i = \vartheta^*,
\end{aligned}$$

which shows that $\bar u$ is not the analytic center. Here, the last inequality follows from the classical inequality $\alpha + 1/\alpha \geq 2$ if $\alpha \in \mathbb{R}_+$ and strict inequality holds when $\alpha \neq 1$.

Theorem 7. *The limit point u^* of the central path is the analytic center of* SP^*.

Proof. Let $(\bar u, \bar w)$ be an optimal solution. Then, from the orthogonality relation $(\bar u - u^*)^T(\bar w - w^*) = 0$, we derive

$$\sum_{i\in\mathbf{B}} \frac{\bar u_i}{u_i^*} + \sum_{i\in\mathbf{N}} \frac{\bar w_i}{w_i^*} = n.$$

Now we apply the arithmetic-geometric mean inequality to derive

$$\left(\prod_{i\in\mathbf{B}}\frac{\bar u_i}{u_i^*}\prod_{i\in\mathbf{N}}\frac{\bar w_i}{w_i^*}\right)^{1/n}\le\frac{1}{n}\left(\sum_{i\in\mathbf{B}}\frac{\bar u_i}{u_i^*}+\sum_{i\in\mathbf{N}}\frac{\bar w_i}{w_i^*}\right)=1.$$

Hence,

$$\prod_{i\in\mathbf{B}}\bar u_i\prod_{i\in\mathbf{N}}\bar w_i\le\prod_{i\in\mathbf{B}}u_i^*\prod_{i\in\mathbf{N}}w_i^*,$$

proving that u^* is the analytic center of SP^*. The proof is complete. □

2.2.5. Identifying the Optimal Partition.

The Condition Number. To give bounds on the size of the variables along the central path, we need to find a quantity that in some sense characterizes the set of optimal solutions. For an optimal solution $u\in\mathrm{SP}^*$, we have

$$uw(u)=0 \quad\text{and}\quad u+w(u)\ge 0.$$

Our next question is about the size of the nonzero coordinates of optimal solutions. Following the definitions in Roos et al. [75] and Ye [96], we define a condition number of the problem (SP) that characterizes the magnitude of the nonzero variables on the optimal set SP^*.

Definition 3. Let us define

$$\sigma^u:=\min_{i\in\mathbf{B}}\max_{u\in\mathrm{SP}^*}\{u_i\},$$
$$\sigma^w:=\min_{i\in\mathbf{N}}\max_{u\in\mathrm{SP}^*}\{w(u)_i\}.$$

Then, the *condition number* of (SP) is defined as

$$\sigma=\min\{\sigma^u,\sigma^w\}=\min_i\max_{u\in\mathrm{SP}^*}\{u_i+w(u)_i\}.$$

To determine the condition number σ is, in general, more difficult than to solve the optimization problem itself. However, we can give an easily computable lower bound for σ. This bound depends only on the problem data.

Lemma 4 (Lower Bound for σ). *If M and q are integral*[6] *and all the columns of M are nonzero, then*

$$\sigma\ge\frac{1}{\pi(M)},$$

where $\pi(M)=\prod_{i=1}^n\|M_i\|$.

As we have seen, the condition that none of the columns of the matrix M is a zero vector is not restrictive. For the general problem (SP), a zero column M_i would imply that $w_i=q_i$ for all feasible solutions; thus, the pair (u_i,w_i) could be removed. More important is that for our embedding problem ($\overline{\mathrm{SP}}$) none of the columns of the coefficient matrix

$$\begin{pmatrix} M & r\\ -r^T & 0\end{pmatrix}$$

is zero. By definition we have $r=e-Me$ nonzero, because $e^Tr=e^Te-e^TMe=n$. Moreover, if $M_i=0$, then by using that matrix M is skew-symmetric, then we have $r_i=1$; thus, the ith column of the coefficient matrix is again nonzero.

[6] If the problem data are rational, then by multiplying by the least common multiple of the denominators, an equivalent LO problem with integer data is obtained.

The Size of the Variables Along the Central Path. Now, by using the condition number σ, we are able to derive lower and upper bounds for the variables along the central path. Let $(\mathbf{B},\mathbf{N})$ be the optimal partition of the problem (SP).

Lemma 5. *For each positive* μ, *one has*

$$u_i(\mu) \geq \frac{\sigma}{n} \quad i \in \mathbf{B}, \qquad u_i(\mu) \leq \frac{n\mu}{\sigma} \quad i \in \mathbf{N},$$
$$w_i(\mu) \leq \frac{n\mu}{\sigma} \quad i \in \mathbf{B}, \qquad w_i(\mu) \geq \frac{\sigma}{n} \quad i \in \mathbf{N}.$$

Identifying the Optimal Partition. For sufficiently small μ, the bounds presented in Lemma 5 enable us to identify the optimal partition $(\mathbf{B},\mathbf{N})$. We just have to calculate the μ value that ensures that the coordinates going to zero are certainly smaller than the coordinates that stay above the specified value and converge to a positive number.

Corollary 2. *If we have a central solution* $u(\mu) \in \mathrm{SP}$ *with*

$$\mu < \frac{\sigma^2}{n^2},$$

then the optimal partition $(\mathbf{B},\mathbf{N})$ *can be identified.*

The results of Lemma 5 and Corollary 2 can be generalized to the situation when a vector (u,w) is not on, but just in a certain neighborhood of, the central path. To keep our discussion short, we do not go into those details. The interested reader is referred to Roos et al. [75].

2.2.6. Rounding to an Exact Solution. Our next goal is to find a strictly complementary solution. This could be done by moving along the central path as $\mu \to 0$. Here, we show that we do not have to do that; we can stop at a sufficiently small $\mu > 0$ and round off the current "almost optimal" solution to a strictly complementary optimal one. We need some new notation. Let the optimal partition be denoted by $(\mathbf{B},\mathbf{N})$, and let $\omega := \|M\|_\infty = \max_{1\leq i\leq n} \sum_{j=1}^n |M_{ij}|$ and $\pi := \pi(M) = \prod_{i=1}^n \|M_i\|$.

Lemma 6. *Let* M *and* q *be integral and all the columns of* M *be nonzero. If* $(u,w) := (u(\mu), w(u(\mu)))$ *is a central solution with*

$$u^T s = n\mu < \frac{\sigma^2}{n^{3/2}(1+\omega)^2\pi}, \quad \textit{which certainly holds if } n\mu \leq \frac{1}{n^{3/2}(1+\omega)^2\pi^3},$$

then by a simple rounding procedure a strictly complementary optimal solution can be found in $\mathcal{O}(n^3)$ *arithmetic operations.*

Note that this rounding result can also be generalized to the situation when a vector (u,w) is not on, but just in a certain neighborhood of, the central path. For details the reader is again referred to Roos et al. [75]. This result makes clear that when one solves an LO problem by using an IPM, the iterative process can be stopped in a proper neighborhood of the central path at a sufficiently small value of μ. At that point a strictly complementary optimal solution can be identified easily.

2.2.7. Optimal Basis Identification. It is well known that simplex methods provide an optimal basis. An optimal basis is desired, e.g., when one solves integer LO problems using branch-and-cut methods. A reassuring fact is that, having a complementary optimal solution pair, an optimal basis can be identified in strongly polynomial time; thus, an interior optimal solution pair allows efficient identification of an optimal basis.

Theorem 8. *Let* (u,w) *be a pair of optimal solutions of problem* (SP). *Then, an optimal basis solution can be identified in strongly polynomial time, specifically by using at most* n *pivot steps.*

For a proof the reader may consult Megiddo [56], Roos et al. [75], and Ye [96]. Note that this result is valid only one way. One can find an optimal basis in strongly polynomial time if a complementary pair of optimal solutions is given, such as a "central" strictly complementary solution pair is provided by IPMs. However, identifying a strictly complementary optimal solution pair from an optimal basis solution cannot be done in strongly polynomial time. Either a series of decreasing size of LO problems needs to be solved in a procedure (not a polynomial time procedure) as proposed by Balinski and Tucker [11], or the problem needs to be solved again by a polynomial time IPM followed by the strongly polynomial time rounding procedure.

2.3. Summary of the Theoretical Results

Let us return to our general LO problem in canonical form

$$\min\{c^T x\colon\ Ax - z = b,\ x \geq 0,\ z \geq 0\}, \tag{P}$$

$$\max\{b^T y\colon\ A^T y + s = c,\ y \geq 0,\ s \geq 0\}, \tag{D}$$

where for convenience the slack variables are already included in the problem formulation. In what follows we recapitulate the results obtained so far.

- In §2.1 we see that to solve the LO problem it is sufficient to find a strictly complementary solution to the Goldman-Tucker model (2).
- This *homogeneous* system always admits the zero solution; however, to get a solution for our original problems we need a solution for which $\tau + \rho > 0$ holds. In more general terms, a strictly complementary solution pair is needed.
- If (x^*, z^*) is optimal for (P) and (y^*, s^*) for (D), then $(y^*, x^*, 1, z^*, s^*, 0)$ is a solution for the Goldman-Tucker model with the requested property $\tau + \rho > 0$. See Theorem 2.
- Any solution of the Goldman-Tucker model (y, x, τ, z, s, ρ) with $\tau > 0$ yields an optimal solution pair (scale the variables (x, z) and (y, s) by $1/\tau$) for LO. See Theorem 2.
- Any solution of the Goldman-Tucker model (x, z, y, s, τ, ρ) with $\rho > 0$ provides a certificate of primal or dual infeasibility (see Theorem 2).
- If $\tau = 0$ in every solution (y, x, τ, z, s, ρ), then (P) and (D) have no optimal solutions with zero duality gap.
- The Goldman-Tucker model can be transformed into a skew-symmetric self-dual problem (SP) satisfying the IPC. See §2.2.1.
- If problem (SP) satisfies the IPC, then
  - the central path exists (see Theorem 5);
  - the central path converges to a strictly complementary solution (see Theorem 6);
  - the central path converges to the analytic center of the optimal set (see Theorem 7);
  - if the problem data are integral and a solution on the central path with a sufficiently small μ is given, then the optimal partition (see Corollary 2) and an exact strictly complementary optimal solution (see Lemma 6) can be found.
- These results give a constructive proof of Theorem 3.
- This way, as we have seen in §2.1, the strong duality theorem of linear optimization (Theorem 4) is proved.
- From any complementary optimal pair of primal-dual solutions, an optimal basis can be identified in strongly polynomial time (see Theorem 8).

The above summary shows that the duality theory of LO can be built by using only elementary calculus and by the introduction of fundamental concepts of IPMs.

In the rest of this section a generic IP algorithm is presented and the complexity of some specific variants is discussed.

2.4. A General Scheme of IP Algorithms for Linear Optimization

In this section we briefly review the main elements of IPMs. For easy reference to the majority of the literature, the algorithmic concepts are presented for the standard form primal-dual pair of LO problems:

$$\min\{c^T x\colon Ax = b,\, x \geq 0\}, \tag{P}$$

$$\max\{b^T y\colon A^T y + s = c,\, y \geq 0\}, \tag{D}$$

where for ease of reference the slack variables are included in the problem formulation. For this form of LO problems, for any interior feasible $x, s > 0$ solution, the Newton system (3) is given as

$$\begin{aligned} A\Delta x &= 0,\\ A^T\Delta y + \Delta s &= 0,\\ s\Delta x + x\Delta s &= \mu e - xs. \end{aligned} \tag{4}$$

Assuming that rank $(A) = m$ and $x, s > 0$, the above Newton system (4) has a unique solution. Having determined this unique solution, one can make a—possibly damped—Newton step to update the current iterate:

$$\begin{aligned} x &:= x + \alpha\Delta x,\\ s &:= s + \alpha\Delta s. \end{aligned}$$

The various IPMs make full or damped Newton steps at each iteration, and the process is controlled by careful selection of the step length and the targeted central path parameter value μ or, in a broader sense, the right-hand side of the third Newton equation. A key tool in this process is the choice of a proximity measure that allows to quantify the distance from the central path and ensures that the iterates remain positive.

2.4.1. Proximity Measures. We have seen that the central path leads to a specific strictly complementary solution, to the analytic center of the set of optimal solutions. However, because of the nonlinearity of the equation system determining the central path, the iterates cannot stay on the central path, even if one uses the embedding model and the initial interior point was on the central path. For this reason we need some proximity measures that allow us to control the Newton process and keep the iterates in a proper neighborhood of the central path. These proximity measures depend on the current primal-dual iterate x and s, and the central path parameter value μ, which is either chosen independent of x and s or as $\mu = x^T s/n$. The goal is to quantify how far the iterate is from the point corresponding to μ on the central path. In general, the proximity measure is denoted by $\delta(x, s, \mu)$.

Observe that on the central path all the coordinates of the vector xs are equal to $\mu = x^T s/n$. This observation indicates that the proximity measure

$$\delta_c(xs) := \frac{\max(xs)}{\min(xs)},$$

where $\max(xs)$ and $\min(xs)$ denote the largest and smallest coordinates, respectively, of the vector xs, is an appropriate measure of centrality. One has $\delta_c(xs) \geq 1$, and the larger the value of $\delta_c(xs)$ is, the further the iterate is from the target point of the central path. If $\delta_c(xs) = 1$, then the current solution $(x, s) = (x(\mu), s(\mu))$ is on the central path. In the literature of IPMs, various centrality measures were developed (see den Hertog [26], Jansen [44], Roos et al. [75], Wright [94], Ye et al. [97], Ye [96]). Here, we present just another proximity measure, extensively used in Roos et al. [75]:

$$\delta_0(xs,\mu) := \frac{1}{2}\left\| \left(\frac{xs}{\mu}\right)^{1/2} - \left(\frac{\mu}{xs}\right)^{1/2} \right\|.$$

One has $\delta_0(xs,\mu) \geq 0$, and the larger the value of $\delta_0(xs)$ is, the further the iterate is from target point of the central path. If $\delta_0(xs,\mu) = 0$, then the current solution $(x,s) = (x(\mu), s(\mu))$ is on the central path. Both of the proximity measures allow us to design polynomial IPMs.

2.4.2. A Generic Interior Point Algorithm. Algorithm 1 gives a general framework for a large family of interior point methods.

Algorithm 1 (Generic Interior Point Newton Algorithm)

Input:
A proximity parameter γ; an accuracy parameter $\varepsilon > 0$;
a variable step-length damping factor α;
update parameter θ, $0 < \theta < 1$;
a feasible (x^0, s^0) and $\mu^0 > 0$ s.t. $\delta(x^0, s^0, \mu^0) \leq \gamma$.
begin:
$x := x^0$; $s := s^0$; $\mu := \mu^0$;
while $n\mu \geq \varepsilon$ **do**
begin:
$\mu := (1-\theta)\mu$;
while $\delta(x,s,\mu) \geq \gamma$ **do**
begin:
solve the Newton system (4);
choose the damping factor α;
$x := x + \alpha\Delta x$;
$s := s + \alpha\Delta s$;
end
end
end

To specify a concrete algorithm, the following choices need to be made:

- Choose the proximity measure $\delta(x,s,\mu)$,
- Choose the proximity parameter γ,
- Choose the update scheme for μ, and
- Specify the step length for the Newton step.

Proper choices allow us to prove polynomial iteration complexity of the resulting algorithm. Here, we present three sets of choices that give polynomial IPMs. For complexity proofs, see, e.g., Roos et al. [75] and the papers cited therein. Recall that by using the embedding model ($\overline{\text{SP}}$), one can use the all one vector with $\mu = 1$ as a perfectly centered initial interior solution.

The first IPM, which also enjoys the best possible iteration complexity bound, is a *primal-dual algorithm with full Newton steps*; see, e.g., Roos et al. [75]. Let us make the following choices:

- $\delta(x,s,\mu) := \delta_0(xs,\mu)$;
- $\mu^0 := 1$;
- $\theta := 1/(2\sqrt{n})$;
- $\gamma = 1/\sqrt{2}$; and
- $\alpha = 1$.

Theorem 9 (Theorem II.52 in Roos et al. [75]). *With the given parameter set, the full Newton step algorithm requires at most*

$$\left\lceil 2\sqrt{n}\log\frac{n}{\varepsilon}\right\rceil$$

iterations to produce a feasible solution (x,s) *such that* $\delta_0(xs,\mu)\le\gamma$ *and* $n\mu\le\varepsilon$.

The second IPM is a *primal-dual large-update algorithm*; see, e.g., Roos et al. [75]. The choices are as follows:

- $\delta(x,s,\mu):=\delta_0(xs,\mu)$;
- $\mu^0:=1$;
- $0<\theta<n/(n+\sqrt{n})$;
- $\gamma=\sqrt{R}/(2\sqrt{1+\sqrt{R}})$, where $R:=\theta\sqrt{n}/(1-\theta)$; and
- α is the result of a line search, when along the search direction the primal-dual logarithmic barrier function
$$\frac{x^T s}{\mu}-\sum_{i=1}^{n}\log\frac{x_i s_i}{\mu}$$
is minimized.

Theorem 10 (Theorem II.74 in Roos et al. [75]). *With the given parameter set, the primal-dual large-update algorithm requires at most*

$$\left\lceil \frac{1}{\theta}\left\lceil 2\left(1+\sqrt{\frac{\theta\sqrt{n}}{1-\theta}}\right)^4\right\rceil \log\frac{n}{\varepsilon}\right\rceil$$

iterations to produce a feasible solution (x,s) *such that* $\delta_0(xs,\mu)\le\tau$ *and* $n\mu\le\varepsilon$.

When we choose $0<\theta<1$ as a constant independent of n, e.g., $\theta=1/2$ or $\theta=0.9$, then the total complexity becomes $\mathcal{O}(n\log(n/\varepsilon))$, whereas the choice $\theta=\nu/\sqrt{n}$, with any fixed positive value ν, gives a complexity of $\mathcal{O}(\sqrt{n}\log(n/\varepsilon))$.

Other versions of this algorithm were studied in Peng et al. [66], where the analysis of large-update methods was based purely on the use of the proximity $\delta_0(xs,\mu)$.

The third IPM is the *Dikin step algorithm* studied, e.g., in Roos et al. [75]. This is one of the simplest IPMs, with an extremely elementary complexity analysis. The price for simplicity is that the polynomial complexity result is not the best possible. For this algorithm we make the following choices:

- $\delta(x,s,\mu):=\delta_c(xs)$, recall that this measure is always larger than or equal to 1;
- $\mu^0:=0$, which implies that μ stays equal to zero, thus θ is irrelevant;
- $\gamma=2$;
- $(\Delta x,\Delta s)$ is the solution of (4) when the right-hand side of the last equation is replaced by $-x^2s^2/\|xs\|$; and
- $\alpha=1/(2\sqrt{n})$.

Theorem 11 (Theorem I.27 in Roos et al. [75]). *With the given choices, the Dikin step algorithm requires not more than*

$$\left\lceil 2n\log\frac{n}{\varepsilon}\right\rceil$$

iterations to produce a feasible solution (x,s) *for* ($\overline{\mathrm{SP}}$) *such that* $\delta_c(xs)\le 2$ *and* $n\mu\le\varepsilon$.

2.5. Notes on the Log-Barrier Approach

To define the central path, so far we perturbed the optimality conditions for the primal-dual LO problems. An equivalent and, for people familiar with nonlinear optimization methods, more straightforward approach is to define the central path by the logarithmic barrier function. Let us consider the linear optimization problem in the standard form:

$$\begin{aligned} \min \quad & c^T x \\ \text{s.t.} \quad & Ax = b, \\ & x \geq 0. \end{aligned} \tag{P}$$

Assuming that problem (P) has a strictly positive solution, the standard logarithmic barrier approach puts the inequalities in the *log-barrier function*, and with the barrier parameter $\mu > 0$ adds the log-barrier term to the objective function:

$$\min \left\{ c^T x - \mu \sum_{i=1}^{n} \ln x_i \colon Ax = b \right\}. \tag{PBar}$$

This equality constrained optimization problem is defined on the intersection of the interior of the positive orthant and the affine space $\{x\colon Ax = b\}$. The idea behind the barrier method is to gradually reduce μ while approximately solving the log-barrier subproblems. When μ is sufficiently small, an almost-optimal solution of the original problem is obtained that can be rounded off to an exact optimal solution of the original problem.

The first-order optimality conditions for system (PBar) are

$$\begin{aligned} Ax &= b, \\ A^T y - c + \mu x^{-1} &= 0. \end{aligned}$$

Taking $s_i = \mu(1/x_i)$, $i = 1, \dots, n$, this equation yields $xs = \mu e$, the same central path equations that we obtained in Definition 1. An identical result can be derived starting from the standard dual LO problem.

2.5.1. Barrier Approaches. A natural extension of this idea is to replace the $-\ln t$ function with other barrier functions. Intriguing results were obtained by Nesterov and Nemirovski [64], who introduced the so-called *universal barrier function* that allows one to solve any smooth convex optimization problem in a polynomial number of iterations. Such an algorithm is not necessarily a polynomial time algorithm because the execution of a single iteration might not be possible in polynomial time. Copositive optimization, i.e., optimization of a linear objective function over the intersection of an affine subspace and the cone of copositive matrices (Bomze et al. [14]), is such a convex optimization problem, because it cannot be decided in polynomial time if a matrix is copositive or not; thus, a Newton step cannot be made in polynomial time. Nesterov and Nemirovski [64] also developed the powerful and comprehensive theory of self-concordant barrier functions that allows one to develop polynomial time algorithms for large classes of convex optimization problems.

The *volumetric barrier* method (Anstreicher [9, 10], Vaidya [86]) in a cutting plane (or column generation) framework allows us to solve problems in polynomial time when a polynomial time oracle is available. The significance of this result is that a large set of combinatorial optimization problems, previously solvable in polynomial time only by the ellipsoid method (Khachiyan [49]; see, e.g., Grötschel et al. [41]), can now be solved with better complexity, so IPMs completely supershade the ellipsoid method.

Self-regular barrier functions, by Peng et al. [66, 67], allow us to design IPMs that operate in a large neighborhood of the central path while having almost the same complexity as small-update IPMs. Another approach was suggested by Ai and Zhang [1] that has $O(\sqrt{n}\log(1/\varepsilon))$ iteration complexity.

3. IPMs, Klee-Minty Examples, and Complexity Bounds

Although path-following IPMs solve LO problems in polynomial time, their complexity depend on the number of inequalities in the problem and the condition number or input length (see §2.2.5; see, e.g., Roos et al. [75], Nematollahi [61]). In this section we review cases when the central path exhibits extreme behavior, and following from the extreme cases, we draw the limitations of IPMs.

3.1. Klee-Minty Examples for IPMs

Since the early 1970s, the Klee-Minty example (Klee and Minty [50], Roos [74]) has been giving evidence that simplex algorithms may require an exponential number of pivot steps to solve LO problems. Although IPMs are polynomial time algorithms, in this section we show that in any dimension, with an exponential number of constraints, the central path, and thus IPMs, may exhibit analogous behavior. In their paper, Deza et al. [31] showed in 2004 that the central path of a representation of the Klee-Minty cube with exponentially many redundant constraints follows the simplex path. More precisely, in their construction, exponentially many redundant constraints that are parallel to the facets passing through the optimal vertex are added to the Klee-Minty cube to force the central path to visit a predefined—but arbitrarily small—neighborhood of all the vertices of the Klee-Minty cube in the same order as simplex methods do. This fact implies that this central path is having $2^n - 2$ sharp turns. Uniform distances for the redundant constraints have been chosen in this first Klee-Minty construction, and consequently the number of the inequalities in dimension n becomes $N = O(n^2 2^{6n})$. In subsequent papers the authors developed other, leaner constructions, and significantly reduced the number of redundant inequalities. In Deza et al. [30], by a meticulous analysis, it is shown that $N = O(n2^{3n})$ redundant constraints suffice to reach the same result. By allowing the distances of the redundant constraints to the corresponding facets to decay geometrically, in Deza et al. [27], the number of the inequalities N reduced to $O(n^3 2^{2n})$. In the same paper, Deza et al. [27] show that after $O(\sqrt{N}n)$ iterations, a standard rounding procedure can be used to identify the optimal solution. These results tighten the iteration complexity lower and upper bounds.

In this chapter we discuss the tightest result to date. In Nematollahi and Terlaky [63], the authors simplify the construction given in Deza et al. [27]. They are placing the redundant constraints parallel to the coordinate hyperplanes at geometrically decaying distances and show that only $N = O(n2^{2n})$ redundant inequalities are needed to bend the central path along the edges of the n-dimensional Klee-Minty cube. This yields an $O(n^{3/2}2^n)$ iteration complexity upper bound, whereas the iteration complexity lower bound $\Omega(2^n)$ is given by the fact that the central path follows the simplex-edge-path arbitrarily close. This result almost completely closes the gap between the iteration complexity upper and lower bounds, because the number of iterations is bounded below by $\Omega(\sqrt{N/\ln N})$ and above by $O(\sqrt{N}\ln N)$.

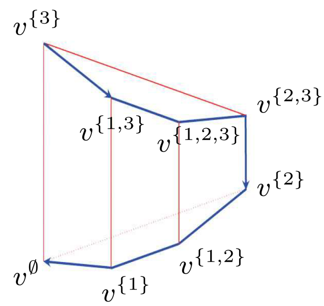
3.2. The Tight Klee-Minty Construction and Complexity Bounds

We consider the Klee-Minty problem (Klee and Minty [50]) in the following form. We use the convention that $x_0 = 0$ and τ is a small positive factor by which the unit cube $[0,1]^n$ is perturbed.

$$\begin{array}{rl} \min & x_n \\ \text{s.t.} & \tau x_{k-1} \le x_k \le 1 - \tau x_{k-1} \quad \text{for } k = 1, \ldots, n. \end{array}$$

This optimization problem has $N = 2n$ constraints, n variables, and the feasible region is a perturbed n-dimensional cube denoted by $\mathcal{C}$. Most variants of the simplex method take $2^n - 1$ iterations to solve this problem. Starting from the vertex $(0, \ldots, 0, 1)^T$, they may visit all the vertices because they are ordered by the decreasing value of the objective function, which is x_n the last coordinate, until the optimal point, which is the origin, is reached.

FIGURE 1. The vertices of the Klee-Minty 3-cube and the simplex path P_0.



To force the central path to behave analogously, we introduce redundant constraints induced by the hyperplanes $H_k := \{x\colon d_k + x_k = 0\}$, where d_k denotes the distance to the respective coordinate plane. Each such constraint is repeated h_k times, for $k = 1, \ldots, n$. Adding redundant constraints does not change the feasible region; however, the analytic center of the Klee-Minty cube and the central path is affected by the addition of the redundant constraints. Let us denote the repetition vector by $h = (h_1, \ldots, h_n)^T$ and the distance vector by $d = (d_1, \ldots, d_n)^T$. Finally, we define the redundant LO problem $\mathcal{C}_\tau^h$ as

$$\begin{aligned} \min\ & x_n \\ \text{s.t.}\ & \tau x_{k-1} \le x_k \le 1 - \tau x_{k-1} \quad \text{for } k = 1, \ldots, n, \\ & 0 \le d_k + x_k \text{ repeated } h_k \text{ times}, \quad \text{for } k = 1, \ldots, n. \end{aligned}$$

By analogy with the unit cube $[0,1]^n$, we denote the vertices of the Klee-Minty cube $\mathcal{C}$ by using subsets $\mathbf{S}$ of $\{1, \ldots, n\}$. For $\mathbf{S} \subseteq \{1, \ldots, n\}$, a vertex $v^{\mathbf{S}}$ of $\mathcal{C}$ is defined (see Figure 1) by

$$v_1^{\mathbf{S}} = \begin{cases} 1 & \text{if } 1 \in \mathbf{S}, \\ 0 & \text{otherwise;} \end{cases}$$

$$v_k^{\mathbf{S}} = \begin{cases} 1 - \varepsilon v_{k-1}^{\mathbf{S}} & \text{if } k \in \mathbf{S}, \\ \varepsilon v_{k-1}^{\mathbf{S}} & \text{otherwise;} \end{cases} \qquad k = 2, \ldots, n.$$

Using the convention $x_0 = 0$, the δ-neighborhood of a vertex $v^{\mathbf{S}}$ (see Figure 2) is defined by

$$\mathcal{N}_\delta(v^{\mathbf{S}}) = \left\{ x \in \mathcal{C}\colon \begin{cases} 1 - x_k - \varepsilon x_{k-1} \le \delta & \text{if } k \in \mathbf{S}, \\ x_k - \varepsilon x_{k-1} \le \delta & \text{otherwise;} \end{cases} \quad k = 1, \ldots, n \right\}.$$

We specify $\mathcal{C}_\tau^h$ by choosing the following parameters:

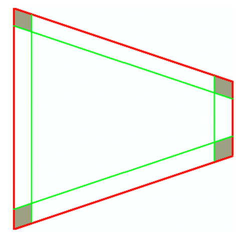
$$\tau = \frac{n}{2(n+1)}, \ \delta \le \frac{1}{4(n+1)},$$

$$d = \left(\frac{1}{\sqrt{\tau^{n-1}}}, \frac{1}{\sqrt{\tau^{n-2}}}, \ldots, \frac{1}{\sqrt{\tau}}, 0 \right),$$

$$h = \left(\left\lfloor \frac{4(1+\sqrt{\tau^{n-1}})}{\sqrt{\tau^{n-1}}\delta} \right\rfloor, \left\lfloor \frac{4(1+\sqrt{\tau^{n-2}})(2+\sqrt{\tau^{n-1}})}{\tau\sqrt{\tau^{n-2}}\delta} \right\rfloor, \ldots, \left\lfloor \frac{4(1+\sqrt{\tau})\prod_{i=2}^{n-1}(2+\sqrt{\tau^i})}{\tau^{n-2}\sqrt{\tau}\delta} \right\rfloor, \right.$$
$$\left. \left\lfloor \frac{4\prod_{i=1}^{n-1}(2+\sqrt{\tau^i})}{\tau^{n-1}\delta} \right\rfloor \right).$$

Although the geometrically decaying sequence of the components of d would give $d_n = 1$, we choose $d_n = 0$, which allows many of the redundant constraints to be active at optimality.

FIGURE 2. The δ-neighborhoods of the vertices of the Klee-Minty 2-cube.



Consequently, h_n does not obey either the sequence rule of the first $n-1$ components of h. The derivation of h is discussed in detail in Nematollahi and Terlaky [63].

To ensure that the δ-neighborhoods of all the 2^n vertices do not overlap, τ and δ have to satisfy $\tau + \delta < 1/2$. Observe that h depends linearly on $1/\delta$. The sharpest bounds are obtained by setting $\delta = 1/(4(n+1))$ and $\tau = n/(2(n+1))$; thus, the corresponding n-dimensional LO problem depends only on n, and thus it is denoted by $\mathcal{C}^n$.

By definition (see Definition 2), the analytic center χ^n of $\mathcal{C}^n$ is the unique maximizer of the product of the slack variables, or equivalently, the maximizer of the sum of the logarithms of the slack variables; i.e., it is the unique solution of the following optimization problem:

$$\max_x \sum_{k=1}^{n} (\ln s_k + \ln \bar{s}_k + h_k \ln \tilde{s}_k),$$

where $s_k = x_k - \tau x_{k-1}$, $\bar{s}_k = 1 - x_k - \tau x_{k-1}$, and $\tilde{s}_k = d_k + x_k$, for $k = 1, \ldots, n$.

The unique optimum of this strictly concave maximization problem is attained when the gradient of the objective function is zero; i.e.,

$$\begin{gathered}
\frac{1}{s_k} - \frac{\tau}{s_{k+1}} - \frac{1}{\bar{s}_k} - \frac{\tau}{\bar{s}_{k+1}} + \frac{h_k}{\tilde{s}_k} = 0 \quad \text{for } k = 1, \ldots, n-1, \\
\frac{1}{s_n} - \frac{1}{\bar{s}_n} + \frac{h_n}{\tilde{s}_n} = 0, \\
s_k > 0,\ \bar{s}_k > 0,\ \tilde{s}_k > 0 \quad \text{for } k = 1, \ldots, n.
\end{gathered}$$

An important observation is that, except the last equation, any point on the central path satisfies all the other above given equalities because the central path $\mathcal{P} = \{x(\mu)\colon \mu > 0\}$ is the set of strictly feasible solutions $x(\mu) > 0$ that are the unique maximizers of

$$\max_x \left\{ -x_n + \mu \sum_{k=1}^{n} (\ln s_k + \ln \bar{s}_k + h_k \ln \tilde{s}_k) \right\}.$$

After defining the central path of $\mathcal{C}^n$, we give a formal definition for the "*simplex path,*" i.e., the edge-path that is followed by simplex methods. We also define its δ-neighborhood in $\mathcal{C}^n$. Let us first define the following sets for $k = 2, \ldots, n$:

$$\begin{aligned}
T_\delta^k &= \{x \in \mathcal{C}\colon \bar{s}_k < \delta\}, \\
C_\delta^k &= \{x \in \mathcal{C}\colon \bar{s}_k \geq \delta, s_k \geq \delta\}, \\
B_\delta^k &= \{x \in \mathcal{C}\colon s_k < \delta\};
\end{aligned}$$

and for $k = 1, \ldots, n$,

$$\hat{C}_\delta^k = \{x \in \mathcal{C}\colon \bar{s}_k < \delta, s_{k-1} < \delta, \ldots, s_1 < \delta\}.$$

The sets T_δ^k, C_δ^k, and B_δ^k can be visualized as the top, central, and bottom parts of $\mathcal{C}$, respectively, and thus $\mathcal{C} = T_\delta^k \cup C_\delta^k \cup B_\delta^k$ for $k = 1, \ldots, n$. By defining the set $A_\delta^k = T_\delta^k \cup \hat{C}_\delta^{k-1} \cup B_\delta^k$ for $k = 2, \ldots, n$, a δ-neighborhood of the simplex path may be illustrated as on

FIGURE 3. The set P_δ for the Klee-Minty 3-cube.

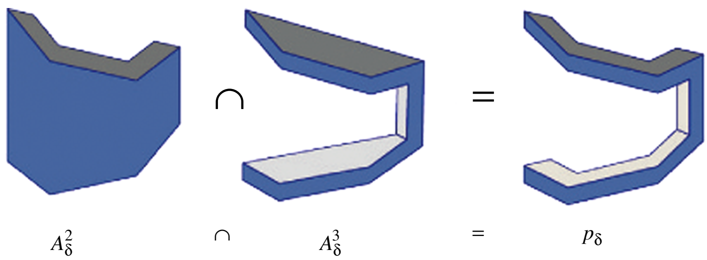


Figure 3, and it is given as $P_\delta = \bigcap_{k=2}^{n} A_\delta^k$. The simplex path itself is obtained in the limit as $\delta \to 0$, and it is given by $P_0 = \bigcap_{k=2}^{n} A_0^k$ (see Figure 1).

For the central path to follow the simplex path, it must originate in the δ-neighborhood of the initial vertex $v^{\{n\}}$. This is precisely the content of Proposition 1; i.e., it proves that the analytic center of $\mathcal{C}^n$ is in the δ-neighborhood $\hat{C}_\delta^n$ of the initial vertex $v^{\{n\}}$. The proof of the proposition can be found in Nematollahi and Terlaky [63].

Proposition 1. *The analytic center χ^n is in the δ-neighborhood of $v^{\{n\}}$; i.e., $\chi^n \in \hat{C}_\delta^n$.*

The next task is to establish that the central path stays in the δ-neighborhood of the simplex path. This is the content of Proposition 2. Consequently, the central path takes at least $2^n - 2$ sharp turns before reaching the optimal solution. In particular, it passes through the δ-neighborhood of all the 2^n vertices of the Klee-Minty n-cube. The proof of this statement can be found in Nematollahi and Terlaky [63] as well.

Proposition 2. *The central path $\mathcal{P}$ stays in the δ-neighborhood of the simplex path of $\mathcal{C}^n$; i.e., $\mathcal{P} \subset P_\delta$.*

As a consequence, as was also discussed in Deza et al. [27, 30], the number of iterations required by path-following interior point methods is at least the number of sharp turns of the central path. In this interpretation, Proposition 2 yields a theoretical lower bound for the iteration complexity of central-path-following IPMs when solving this n-dimensional linear optimization problem. This result is formalized in the following proposition.

Corollary 3. *For $\mathcal{C}^n$, the iteration complexity lower bound of path-following interior point methods is $\Omega(2^n)$.*

Because the theoretical iteration complexity upper bound (Roos et al. [75]) for central-path-following interior point methods is $O(\sqrt{N}L)$, where N and L respectively denote the number of inequalities and the bit length of the input data, we have the following result for $\mathcal{C}^n$:

Theorem 12. *For $\mathcal{C}^n$, the iteration complexity upper bound for central-path-following interior point methods is $O(2^n\sqrt{n}L)$, i.e., $O(2^{3n}n^{5/2})$.*

Proof. Because $\mathcal{C}^n$ is given by $N = 2n + \sum_{k=1}^{n} h_k$ inequalities, and because $e^{x/2} \geq 1 + x/2$, we have

$$h_k = \frac{\prod_{i=n-k}^{n-1}(2+\sqrt{\tau^i})}{\tau^{k-1}\sqrt{\tau^{n-k}}\delta} \leq \frac{2^k e^{1/2\sum_{i=n-k}^{n-1}\sqrt{\tau^i}}}{\tau^{k-1}\sqrt{\tau^{n-k}}\delta}.$$

Because $e^{(1/2)\sum_{i=n-k}^{n-1}\sqrt{\tau^i}} \leq \sqrt{\tau^{n-k}}/(1-\sqrt{\tau})$ and $\tau < 1/2$, we get $h_k \leq 2^k/(\tau^{k-1}(1-\sqrt{\tau})\delta) \leq 2^{k+1}/(\tau^k\delta)$. Therefore, using the set values of $\tau = n/(2(n+1))$ and $\delta = 1/(4(n+1))$, we have

$$N \leq 2n + \sum_{k=1}^{n} 2^{2k+3}\left(\frac{n}{n+1}\right)^k (n+1) \leq 2n + (n+1)\sum_{k=1}^{n} 2^{2k+3} = 2n + \frac{32}{3}(n+1)(2^{2n}-1).$$

Thus, we conclude that $N = O(n2^{2n})$ and $L \leq N\lceil \ln d_1 \rceil = O(n^2 2^{2n})$. $\square$

Because along the simplex path all the vertices are ordered by their decreasing last components, for the last two vertices $v^{\{1\}}$ and $v^{\varnothing}$ we have $v_n^{\{1\}} = \tau^{n-1}$ and $v_n^{\varnothing} = 0$, respectively. As known, the $N\mu^* < \epsilon$ (see, e.g., Theorem 9 and Roos et al. [75]) ϵ-precision stopping criterion can be replaced by $N\mu^* < (1/2)\tau^{n-1}$. Then, the standard rounding procedure (see, e.g., Roos et al. [75] and Lemma 6) can be used to compute this problem's unique, exact optimal solution. Additionally, one can check (see Nematollahi and Terlaky [63]) that the starting point can be chosen on the central path corresponding to $\mu^0 = \tau^{n-1}\delta = n^{n-1}/(2^{n+1}(n+1)^n)$, which is in the interior of the δ-neighborhood of the first vertex $v^{\{n\}}$. This observation yields an input-length-independent iteration complexity bound $O(\sqrt{N}\ln(N\mu^0/(N\mu^*))) = O(\sqrt{N}n)$, and thus the results of Theorem 12 are strengthened to an input-length-independent version as follows.

Theorem 13. *For $\mathcal{C}^n$, the iteration complexity upper bound of path-following interior point methods is $O(n^{3/2}2^n)$.*

Observe that, unlike the examples of Deza et al. [31, 27], in the construction presented here a set of the redundant constraints touches the feasible region; therefore, it is less obvious that preprocessors detect redundancy of these constraints.

Remark 1. For $\mathcal{C}^n$, by Corollary 3 and Theorem 13, the iteration complexity lower and upper bounds for central-path-following interior point methods are $O(2^n)$ and $O(n^{3/2}2^n)$ or, expressed in terms of the number of inequalities, $O(\sqrt{N/\ln N})$ and $O(\sqrt{N}\ln N)$, respectively. The gap between the lower and upper bounds is $O(\ln^2 N)$, and thus the gap between the upper and lower bounds is essentially closed.

Remark 2. The following discussions review some of the results related to extreme behavior of interior trajectories.

(i) Megiddo and Shub [57] considered the nonredundant Klee-Minty cube. They proved that from certain starting points affine scaling trajectories, just as the central path in our redundant Klee-Minty construction, may trace the simplex path. They also noted that their approach does not extend to projective scaling.

(ii) Todd and Ye [83] gave an $\Omega(\sqrt[3]{N})$ iteration complexity lower bound for the number of iterations between two updates of the central path parameter μ.

(iii) Vavasis and Ye [89] provided an $O(N^2)$ upper bound for the number of approximately straight segments of the central path.

(iv) The knapsack problem with proper objective function yields an n-dimensional example with $n+1$ constraints and n sharp turns.

(v) In Nematollahi and Terlaky [62] another intriguing Klee-Minty construction is presented where all the redundant constraints are touching the feasible set. The price to pay for such a construction is a significantly higher number of redundant constraints. In this example the number of constraints is $N = O(2^{n^2})$, which is exponentially larger than the one in the previously presented and cited constructions.

The Klee-Minty construction presented in this section provides an example where the central path follows the simplex path. To force such extreme behavior, an exponential number of redundant constraints is needed. One may wonder what worst-case examples for the central path are possible when no redundant constraints are in the problem formulation. Although in the Klee-Minty examples the number of turns of the central path is exponential in terms of the problem dimension, the number of turns is less than the square root of the number of constraints. A natural question is whether the sharp turns of the central path can be equal (or at least almost equal) to the number of constraints. Such questions and related conjectures are discussed in the next section.

4. Curvature and Conjectures

In this section we review the results and conjectures of Deza et al. [29, 30]. In Deza et al. [29], a nonredundant construction with N constraints and $N-4$ sharp turns is given. Whereas in the examples of the previous section the number of constraints grows exponentially, in this nonredundant example the input length and the condition number of the problem grows rapidly as the number of inequalities grow.

The Hirsch conjecture (Dantzig [20]) is arguably the most intriguing open problem in optimization theory. By now it stands unsolved for over 60 years. The conjecture was presented by W. M. Hirsch to Dantzig in a letter in 1957. The Hirsch conjecture asserts a linear upper bound for the diameter of polytopes. Although the conjecture is still unsolved, numerous partial results are published (Klee and Walkup [51], Kalai and Kleitman [47], Holt and Klee [43], Fritzsche and Holt [34]). Analogous conjectures and results for the central path curvature (Malajovich et al. [53]) are presented in Deza et al. [30] and are summarized in this section as well.

4.1. Continuous Analogue of the Hirsch Conjecture

Let in dimension n a full dimensional convex polyhedron P be defined by N inequalities. For any pair of vertices $v = \{v_1, v_2\}$ of P, let $\delta^v(P)$ denote the length of the shortest edge path that connects v_1 and v_2. *The diameter of polyhedron* P is defined as the maximum of $\delta^v(P)$ over all possible pairs of vertices, i.e., the smallest number such that any two vertices of polyhedron P can be connected by an edge path formed by at most $\delta^v(P)$ edges. The Hirsch conjecture, posed to Dantzig in 1957 and published in Dantzig [20], states that the diameter of a polyhedron defined by N inequalities in dimension n is not greater than $N-n$.

Conjecture 1 (The Hirsch Conjecture for Polytopes). *The diameter of a polytope defined by N inequalities in dimension n is at most $N-n$.*

The conjecture is proved to be false for unbounded polyhedra and stays unsolved for polytopes, i.e., for bounded polyhedra. Thus, from now on we consider only polytopes and denote those by P. To date, even no polynomial bound is known for the diameter of a polytope. The best bound is $2N^{\log(n)+1}$; see Kalai and Kleitman [47].

4.2. Linear Bound in Terms of the Number of Inequalities

In what follows, as introduced in Deza et al. [29], we consider *the largest possible total curvature of the central path* as the continuous analogue of the diameter. We first recall the definitions of the central path (see Definition 1 and the equivalent definition in §2.5 by problem (PBar) by using the logarithmic barrier function) and of the total curvature (Malajovich et al. [53], Deza et al. [29]). For a polytope $P = \{x\colon Ax \geq b\}$ with $A \in \mathbb{R}^{N\times n}$, the central path of the LO problem $\min\{c^T x\colon x \in P\}$ is the set of minimizers of $\min\{c^T x + \mu f(x)\colon x \in P\}$ for $\mu \in (0, \infty)$, where $f(x)$ is the standard logarithmic barrier function of P given by $f(x) = -\sum_{i=1}^{N} \ln(A_i x - b_i)$. The logarithmic barrier function is a strictly convex function, and thus, as we proved also in §2.2, the central path is unique. It is also known that the central path is an analytic curve (Sonnevend et al. [78]) and thus is fully determined by any of its arbitrary small segments, or by all of its derivatives at any single point. Conversely, one may expect that the analyticity of the central path would allow us to reconstruct the associated polytope P up to its orientation given by the objective vector c and possibly some equivalence relationship between the polytopes.

The *total curvature*, intuitively, measures how far a certain curve differs from a straight line. Let $\psi\colon [\alpha, \beta] \to \mathbb{R}^n$ be a two-times continuously differentiable curve with nonzero derivatives on the interval $(\alpha - \varepsilon, \beta + \varepsilon)$ for some $\varepsilon > 0$. Let denote its arc length by $l(t) = \int_\alpha^t \|\dot\psi(\tau)\|\, d\tau$ and its parameterization by the arc length by $\psi_{\text{arc}} = \psi \circ l^{-1}\colon [0, l(\beta)] \to \mathbb{R}^d$. The curvature of ψ at the point t is denoted by $\kappa(t) = \ddot\psi_{\text{arc}}(t)$. The total curvature is defined as $\int_0^{l(\beta)} \|\kappa(t)\|\, dt$. The assumption $\dot\psi \neq 0$ insures that any given segment of the curve is traversed only once and allows us to define the curvature at any point on the curve.

Let $\lambda^c(P)$ denote the total curvature of the central path of the LO problem $\min\{c^T x \colon x \in P\}$. We aim to draw an analogy between the total curvature $\lambda^c(P)$ and the diameter $\delta^v(P)$. However, $\lambda^c(P)$ still depends on the objective vector c, whereas the diameter $\delta^v(P)$ is independent of it. By taking the supremum of $\lambda^c(P)$ over all possible objective coefficient vectors c, we obtain the quantity $\lambda(P)$, referred to as the curvature of the polytope. The quantity $\lambda(P)$ is solely a characteristic of polytope P or, more precisely, to its current representation. The following continuous analogue of the Hirsch conjecture was proposed in Deza et al. [29].

Conjecture 2 (Continuous Analogue of the Hirsch Conjecture). *The order of the curvature of a polytope defined by N inequalities in dimension n is N.*

Although the Hirsch conjecture is open, several intriguing results related to it were proved. Holt and Klee [43] proved that, for $N > n \geq 13$, the Hirsch conjecture is tight. Fritzsche and Holt [34] strengthened this result to $N > n \geq 8$. The continuous analogue would be that Conjecture 2 is tight as well. In Deza et al. [28], the authors showed that a redundant Klee-Minty n-cube $\mathcal{C}$ satisfies $\lambda(\mathcal{C}) \geq (3/2)^n$ and in Deza et al. [29] provided a family of n-dimensional polytopes $\mathcal{P}$ defined by $N > 2n$ nonredundant inequalities satisfying $\liminf_{N\to\infty} \lambda(\mathcal{P})/N \geq \pi$ for a fixed n. In other words, the continuous analogue of the result of Holt and Klee [43] holds; i.e.,

Theorem 14 (Continuous Analogue of the Result of Holt and Klee (Deza et al. [29]). The conjectured order N of the curvature of a polytope is tight.

This result was proved by presenting a series of polytopes that asymptotically provide the tight bound. In dimension $n = 2$, let the objective coefficient vector be $c = (0, 1)^T$, and the feasible set of the corresponding series of LO problems is given for $N \geq 4$ by the polytope $P^*_{N,2}$ that is defined by the following N inequalities:

$$\begin{aligned} x_2 &\leq 1, \\ x_1 &\leq \frac{x_2}{10} + \frac{1}{2}, \\ -x_1 &\leq \frac{x_2}{3} + \frac{1}{3}, \\ (-1)^i x_1 &\leq \frac{10^{i-2}}{11} x_2 + \frac{5}{11} - \frac{10^{-4}}{m}\frac{i}{N} \qquad \text{for } i = 4, \ldots, N. \end{aligned} \tag{5}$$

The resulting LO problem and its central path for $P^*_{6,2}$ are illustrated on Figure 4. Furthermore, the central path of $P^*_{34,2}$ on a logarithmic scale is presented in Figure 5.

FIGURE 4. The polytope $P^*_{6,2}$ and its central path.

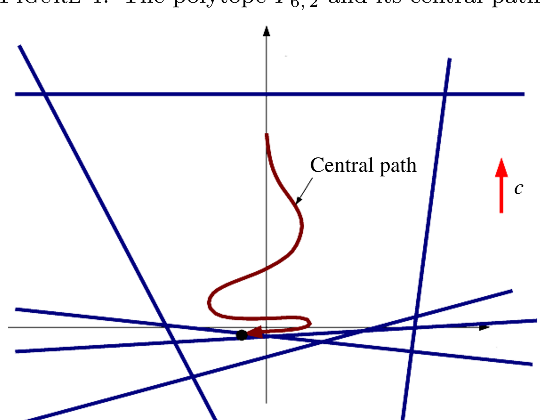
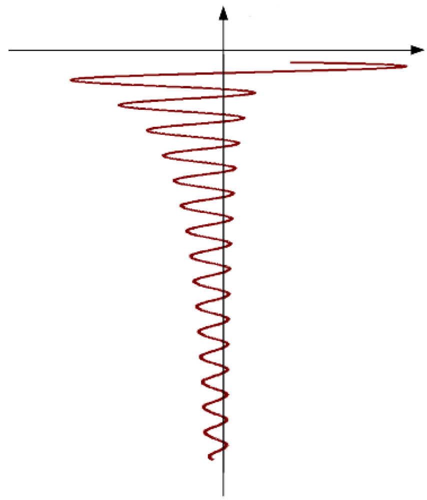


FIGURE 5. The central path for $P^*_{34,2}$ on the logarithmic scale in x_2.



This example leads to the following explicit form of Theorem 14, which provides the desired asymptotically tight bound. For details of the proof, the reader is referred to Deza et al. [29].

Theorem 15. *The total curvature of the central path of* $\min\{x_2\colon (x_1, x_2) \in P^*_{N,2}\}$ *satisfies*

$$\liminf_{N\to\infty} \frac{\lambda^{(0,1)}(P^*_{N,2})}{N} \geq \pi.$$

One can easily see that the polytope $P^*_{N,2}$, and the result of Theorem 15, can straightforwardly be generalized to any dimensions $n \geq 3$ by adding the box constraints $0 \leq x_i \leq 1$ for $3 \leq i \leq n$.

4.3. Conjectures and Results

A motivation to study the diameter and the curvature of polytopes can come from the consideration that the diameter may be thought of as a worst-case lower bound on the number of pivots for a simplex method, whereas the curvature of the central path influences the performance of central-path-following interior point methods.

After seeing some examples where the central path has exponentially many sharp turns in any given dimension, and also seeing that in a nonredundant example the central path can make as many 180-degree turns asymptotically as the number of inequalities, one may think about upper bounds for the curvature of a polytope and also the average curvature in some appropriate probabilistic models.

An exponential upper bound on the curvature of a polytope is provided by the result of Dedieu et al. [53]. Let us consider a simple arrangement (Grünbaum [42], Ziegler [98]) formed by N hyperplanes in dimension[7] n. In a simple arrangement, let k_c denote the number of bounded cells $P_i, i = 1, \ldots, k_c$, i.e., the closures of the bounded connected components of the complement of the hyperplanes. It is known that $k_c = \binom{N-1}{n}$. Malajovich et al. [53] proved that for any fixed objective vector c, one has

$$\sum_{i=1}^{i=I} \lambda^c(P_i)/k_c \leq 2\pi n.$$

This result implies that $\lambda^c(P) \leq 2\pi n\binom{N-1}{n} < 2\pi n N^{n-1} < 2\pi N^n$ for any c, i.e., $\lambda(P) \leq 2\pi N^n$.

[7] Recall that an arrangement of hyperplanes is called simple if $N \geq n+1$ and any n hyperplanes intersect at a unique distinct point.

Theorem 16 (Exponential Upper Bound on the Curvature). *The curvature of a polytope defined by N inequalities in dimension n is at most $2\pi N^n$.*

The special case of $N = 2n$ of the Hirsch conjecture is known as the d-step[8] conjecture, and its continuous analogue is

Conjecture 3 (Continuous Analogue of the d-Step Conjecture). *The order of the curvature of a polytope in dimension n defined by $N = 2n$ inequalities is n.*

Klee and Walkup [51] proved that the special case $N = 2n$, for all n, is equivalent to the Hirsch conjecture. The continuous analogue of this result, i.e, the equivalence of Conjecture 2 and Conjecture 3, was proved in Deza et al. [30].

4.4. The Continuous Analogue of the Result of Klee and Walkup [51]

Following the analogy with the diameter, let $\Lambda(n, N)$ be the largest total curvature $\lambda(P)$ of the central path over all polytopes P defined by N inequalities in dimension n. Klee and Walkup [51] showed that for the Hirsch conjecture the special case where the number of inequalities is twice the dimension is equivalent to the general case. We prove that the same holds for the continuous variant of the Hirsch conjecture.

Theorem 17 (Continuous Analogue of the Result of Klee and Walkup). *The continuous Hirsch conjecture is equivalent to the continuous d-step conjecture; that is, if $\Lambda(n, 2n) = \mathcal{O}(n)$ for all n, then $\Lambda(n, N) = \mathcal{O}(N)$.*

For a proof of Theorem 17, the reader is referred to Deza et al. [30]. Here we just sketch the main ideas and illustrate how the proof goes. Suppose $\Lambda(n, 2n) \leq 2kn$, $n \geq 2$ for some constant $k > 0$. Consider $\min\{c^T x\colon x \in P\}$, where $P = \{x\colon Ax \geq b\}$ with $A \in \mathbb{R}^{N\times n}$ and $\mathcal{P}$ and χ respectively denote the associated central path and the analytic center, i.e., the minimizer of the standard logarithmic barrier function $f(x)$ over P. There are two cases, $n < N < 2n$ and $N > 2n$, that are considered separately. Let $\mathbf{0}$ denote the vector of all zeros and $\operatorname{int}(P)$ denote the interior of P. We may assume P is full dimensional, for if not, we may reduce the problem dimension to satisfy the assumption. Note A is full rank because P is bounded.

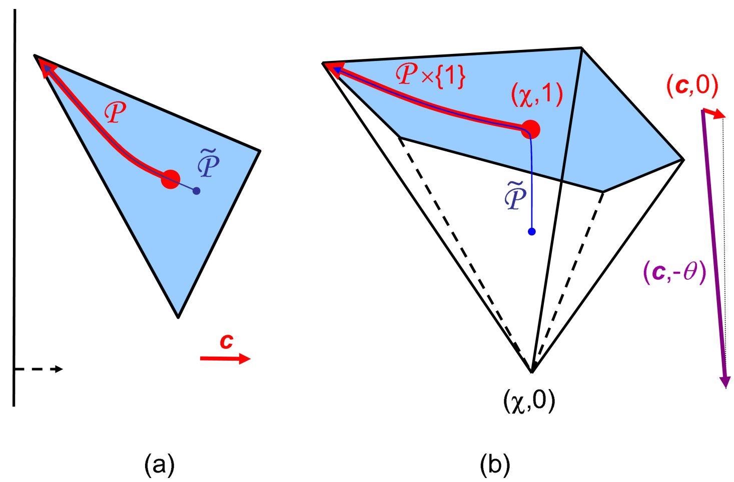
Case $n < N < 2n$. Without loss of generality, assume $c = (1; 0; \ldots; 0) \in \mathbb{R}^n$, and denote x_1^* the optimal value of $\min\{c^T x\colon x \in P\}$. Consider $\min\{c^T x\colon x \in \tilde{P}\}$, where $\tilde{P} = \{x\colon \tilde{A}x \geq \tilde{b}\}$, with $\tilde{A} \in \mathbb{R}^{2n,n}$ and $\tilde{b} \in \mathbb{R}^{2n}$ given by

$$\tilde{A}_{i,j} = \begin{cases} A_{i,j} & \text{for } i = 1, \ldots, N \text{ and } j = 1, \ldots, n, \\ 1 & \text{for } i = N+1, \ldots, 2n \text{ and } j = 1, \\ 0 & \text{for } i = N+1, \ldots, 2n \text{ and } j = 2, \ldots, n; \end{cases} \qquad \tilde{b}_i = \begin{cases} b_i & \text{for } i = 1, \ldots, N, \\ x_1^* - 1 & \text{for } i = N+1, \ldots, 2n, \end{cases}$$

and $\tilde{\mathcal{P}}$ denoting the associated central path (see Figure 6(a)). Recall that the central path may be parameterized as the collection of the analytic centers of the level sets between

[8] In polyhedral theory, it is customary to denote the dimension of the Euclidean space by d.

FIGURE 6. An illustration for the proof of Theorem 17 for $n=2$: (a) $2n > N = 3$; (b) $2n < N = 5$.



the optimal solution and the analytic center χ; i.e., $\mathcal{P}$ is the set of minimizers of $f(x)$ over $P \cap \{x\colon c^T x = \omega\}$, where $\omega \in (x_1^*; \chi_1)$. Therefore, we have $\mathcal{P} \subseteq \tilde{\mathcal{P}}$ and consequently $\lambda^c(P) \le \lambda^c(\tilde{P})$. As $\tilde{P}$ is defined by $2n$ inequalities in dimension n, we have $\lambda^c(\tilde{P}) \le 2kn$, and thus, for $n < N < 2n$, $\lambda^c(P) \le 2kn$; i.e., $\Lambda(n,N) = \mathcal{O}(N)$.

Case $N > 2n$. Without loss of generality, assume $\chi = \mathbf{0}$. Consider $\min\{(c;-\theta)^T(x;x_{n+1})\colon (x;x_{n+1}) \in \tilde{P}\}$, where $\tilde{P} = \{(x,x_{n+1}) \in \mathbb{R}^n \times \mathbb{R}\colon Ax - bx_{n+1} \ge 0,\ x_{n+1} \le 1\}$ with the associated central path $\tilde{\mathcal{P}}$. In particular, if the definition of P is nonredundant, $\tilde{P}$ is the $(n+1)$-dimensional flipped pyramid with base $P \times \{1\}$ and the apex $(\chi;0) = \mathbf{0}$. We show that for a large enough value of $\theta \gg \|c\|$, the central path $\mathcal{P}$ of the original optimization problem may be well approximated by a segment of the central path $\tilde{\mathcal{P}}$, so that the total curvature of $\mathcal{P}$ is bounded from above by the total curvature of $\tilde{\mathcal{P}}$. Intuitively, by choosing θ large enough we should be able to force $\tilde{\mathcal{P}}$ to first follow almost a straight line from the analytic center of $\tilde{P}$ to the face containing $P \times \{1\}$; and once $\tilde{\mathcal{P}}$ is forced almost onto this face, the path looks just like the central path $\mathcal{P}$ for $\min\{c^T x\colon x \in P\}$ in one less dimension (see Figure 6(b)). Consequently, we argue that the total curvature of $\mathcal{P}$ may not be less then that of $\tilde{\mathcal{P}}$.

Roughly speaking, we argue that two similar curves might not differ much in their total curvatures either. Finally, by carefully passing to the limits and using the pointwise convergence of two paths, we may argue that as $\theta \to \infty$, the total curvature of $\mathcal{P}$ may be arbitrarily well approximated by the total curvature of the segment of $\tilde{\mathcal{P}}$; i.e., for a large enough θ, we have $\lambda^c(P) \le \lambda^{(c;-\theta)}(\tilde{P})$, noting that $\tilde{\mathcal{P}}$ is also bound to make an additional sharp turn before it starts to converge to $\mathcal{P}$.

Now, inductively at each step, one may increase the dimension and the number of inequalities by 1 by carefully using the limit argument, and by repeatedly applying the same construction, the result $\Lambda(n,N) = \mathcal{O}(n)$ follows.

4.5. Summary

The above discussions demonstrate that the curvature of a polytope has analogous properties as the diameter of a polytope. Analogous conjectures and analogous partial results can be proved for both objects.

However, for the curvature of polytopes an average case result was proved in Malajovich et al. [53]. They proved that for any fixed objective coefficient vector c, the average curvature $\sum_{i=1}^{i=k_c} \lambda^c(P_i)/k_c$ of the bounded cells of a simple arrangement[9] is bounded by $2\pi n$. Observe

[9] We may refer to this quantity as the average curvature of the simple arrangement.

that this result states that the average curvature of an arrangement is independent of the number of hyperplanes defining the arrangement, and the bound is linear in the dimension. There is no analogous result known for the average diameter of an arrangement, so we may present the following conjecture (Deza et al. [30]):

Conjecture 4 (Average Diameter Conjecture). *The average diameter* $(\sum_{i=1}^{i=k_c} \delta(P_i))/k_c$ *of an arrangement is* $\Omega(n)$.

Observe that if the Hirsch conjecture (Conjecture 1) holds, then the average diameter conjecture holds as well; however, the reverse implication does not hold.

5. Extensions

IPMs spread much beyond LO. Some of the previously mentioned books (den Hertog [26], Jansen [44], Nesterov and Nemirovski [64], Terlaky [81], Ye [96]) discuss extensions of IPMs for classes of nonlinear problems. In the second decade, the majority of research was devoted to IPMs for nonlinear optimization problems, in particular, for conic linear optimization (CLO). The best-known classes of CLO problems are second-order and semidefinite optimization that both have numerous interesting applications, not only in such traditional areas as combinatorial optimization (Alizadeh [2]) and control (Ben-Tal and Nemirovski [12], Pólik and Terlaky [68]), but also in various areas of engineering, including structural (de Klerk et al. [22], Ben-Tal and Nemirovski [12]) and electrical engineering (Vandenberghe and Boyd [87], Boyd and Vandenberghe [16]). For surveys on algorithmic and complexity issues, the reader may consult de Klerk [21], de Klerk et al. [23, 24, 25], Nesterov and Nemirovski [64], Nesterov and Todd [65], Potra and Sheng [70], Sturm [79], Wolkowicz et al. [92], Ye [96], and Pólik and Terlaky [69].

IPMs not only allowed the development of a rich theory, but they have also been implemented with great success for solving LO, CLO, and general NLO problems. It is now a common sense that for large-scale, sparse, structured LO problems, IPMs are the method of choice, and all major commercial optimization software systems contain implementations of IPMs. The reader can find thorough discussions of implementation strategies of IPMs for LO in Gondzio and Terlaky [38], Andersen et al. [7], McShane et al. [54], Mehrotra [58], and Xu et al. [95]. The books by Roos et al. [75], Wright [94], and Ye [96] also devote chapters to that subject. For implementations of IPMs for NLO problems, the reader may consult Andersen and Ye [5], Wächter [90], Wächter and Biegler [91], Byrd et al. [17], and Vanderbei [88]. For discussions of the implementation strategies and documentation of available software for CLO, Andersen et al. [6, 8], Sturm [80], Toh et al. [84], and Borchers and Young [15] provide a rich source of information.

6. Summary

The quarter century of interior point methods, in Wright's [93] words, the "interior point revolution," with lasting consequences, has brought the theory and practice of optimization to a new level. Optimization problems, and in particular LO problems, that were beyond imagination 25 years ago are solved routinely today. The improvement of the efficiency of optimization software is measured in the order of 10^6 or more (Bixby [13]). As Bixby concluded, the advances in computer hardware and the advances of optimization theory, primarily because of IPMs, contributed equally to the dramatic advances of optimization software. Core theory, such as duality theory and sensitivity analysis (Roos et al. [75], Koltai and Terlaky [52], Ghaffari Hadigheh et al. [35]), was rejuvenated. Optimization computation and optimization software development is a rich area of research. IPMs not only spread to all areas of optimization but also allow us to solve new problem classes, such as second-order conic and semidefinite optimization, and smooth nonlinear optimization problems. Novel paradigms, such as the Ben-Tal–Nemirovski robust optimization methodology (Ben-Tal and Nemirovski [12]), opened never-seen opportunities to solve large and, for practice, important problem classes, including problems in truss-topology design, in signal processing, in VLSI design, and in robust and intensity-modulated radiation therapy treatment (Chu et al. [18], Craig et al. [19]). The reader may consult the books by Ben-Tal and Nemirovski [12] and Boyd and Vandenberghe [16] for treasures of novel modeling methodologies and a broad range of applications in various engineering areas.

Twenty-five years after the publication of Karmarkar's path-breaking paper [48], the theory of IPMs has matured so that we also understand the limitations of IPMs. Nevertheless, the theory and practice of optimization has observed transformative changes and has become an essential enabling technology.

References

[1] W. Ai and S. Zhang. An $O(\sqrt{n}L)$ Iteration primal-dual path-following method, based on wide neighborhoods and large updates, for Monotone LCP. *SIAM Journal on Optimization* 16:400–417, 2005.

[2] F. Alizadeh. Combinatorial optimization with interior point methods and semidefinite matrices. Ph.D. thesis, University of Minnesota, Minneapolis, 1991.

[3] F. Alizadeh. Interior point methods in semidefinite programming with applications to combinatorial optimization. *SIAM Journal on Optimization* 5(1):13–51, 1995.

[4] F. Alizadeh and D. Goldfarb. Second-order cone programming. *Mathematical Programming, Series B* 95:3–51, 2002.

[5] E. D. Andersen and Y. Ye. A computational study of the homogeneous algorithm for large-scale convex optimization. *Computational Optimization and Applications* 10(3):243–280, 1998.

[6] E. D. Andersen, C. Roos, and T. Terlaky. On implementing a primal-dual interior-point method for conic quadratic optimization. *Mathematical Programming* 95(2):249–277, 2003.

[7] E. D. Andersen, J. Gondzio, Cs. Mészáros, and X. Xu. Implementation of interior point methods for large scale linear programming. T. Terlaky, ed. *Interior Point Methods of Mathematical Programming*. Kluwer APC, Dordrecht, The Netherlands, 189–252, 1996.

[8] E. D. Andersen, B. Jensen, R. Sandvik, and U. Worsøe. The improvements in MOSEK version 5. Technical Report 1-2007, MOSEK ApS, Copenhagen, 2007.

[9] K. Anstreicher. Large step volumetric potential reduction algorithms for linear programming. *Annals of Operations Research* 62:521–538, 1996.

[10] K. Anstreicher. Volumetric path following algorithms for linear programming. *Mathematical Programming* 76:245–263, 1997.

[11] M. L. Balinski and A. W. Tucker. Duality theory of linear programs: A constructive approach with applications. *SIAM Review* 11:499–581, 1969.

[12] A. Ben-Tal and A. Nemirovski. *Lectures on Modern Convex Optimization: Analysis, Algorithms, and Engineering Applications*. MPS-SIAM Series on Optimization, SIAM, Philadelphia, 2001.

[13] R. E. Bixby. Solving real-world linear programs: A decade and more of progress. *Operations Research* 50(1):3–15, 2002.

[14] I. M. Bomze, M. Duerr, E. de Klerk, C. Roos, A. J. Quist, and T. Terlaky. On copositive programming and standard quadratic optimization problems. *Journal of Global Optimization* 18(2):301–320, 2000.

[15] B. Borchers and J. G. Young. Implementation of a primal-dual method for SDP on a shared memory parallel architecture. *Computational Optimization and Applications* 37(3):355–369, 2007.

[16] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, Cambridge, UK, 2004.

[17] R. Byrd, J. Nocedal, and R. Waltz. KNITRO: An integrated package for nonlinear optimization. G. Di Pillo and M. Roma, eds. *Large-Scale Nonlinear Optimization*, *Nonconvex Optimization and Its Applications*, Vol. 83. Springer, New York, 35–59, 2006.

[18] M. Chu, Y. Zinchenko, S. G. Henderson, and M. B. Sharpe. Robust optimization for intensity modulated radiation therapy treatment planning under uncertainty. *Physics in Medicine and Biology* 53:3231–3250, 2008.

[19] T. Craig, M. B. Sharpe, T. Terlaky, and Y. Zinchenko. Controlling the dose distribution with gEUD-type constraints within the convex IMRTP framework. *Physics in Medicine and Biology* 53:3231–3250, 2008.

[20] G. B. Dantzig. *Linear Programming and Extensions*. Princeton University Press, Princeton, NJ, 1963.

[21] E. de Klerk. *Aspects of Semidefinite Programming: Interior Point Algorithms and Selected Applications*. Kluwer Academic Publishers, Dordrecht, The Netherlands, 2002.

[22] E. de Klerk, C. Roos, and T. Terlaky. Semidefinite problems in truss topology optimization. Technical Report 95-128, Faculty of Technical Mathematics and Informatics, Technische Universiteit Delft, Delft, The Netherlands, 1995.

[23] E. de Klerk, C. Roos, and T. Terlaky. Initialization in semidefinite programming via a self-dual, skew-symmetric embedding. *Operations Research Letters* 20:213–221, 1997.

[24] E. de Klerk, C. Roos, and T. Terlaky. Infeasible-start semidefinite programming algorithms via self-dual embeddings. P. M. Pardalos and H. Wolkowicz, eds. *Topics in Semidefinite and Interior Point Methods*, *Fields Institute Communications*, Vol. 82. AMS, Providence, RI, 215–236, 1998.

[25] E. de Klerk, C. Roos, and T. Terlaky. On primal-dual path-following algorithms for semidefinite programming. F. Gianessi, S. Komlósi, and T. Rapcsák, eds. *New Trends in Mathematical Programming*. Kluwer APC, Dordrecht, The Netherlands, 137–157, 1998.

[26] D. den Hertog. *Interior Point Approach to Linear, Quadratic and Convex Programming*, *Mathematics and Its Applications*, Vol. 277. Kluwer APC, Dordrecht, The Netherlands, 1994.

[27] A. Deza, E. Nematollahi, and T. Terlaky. How good are interior point methods? Klee-Minty cubes tighten iteration-complexity bounds. *Mathematical Programming* 113:1–14, 2008.

[28] A. Deza, T. Terlaky, and Y. Zinchenko. Central path curvature and iteration-complexity for redundant Klee-Minty cubes. D. Gao and H. Sherali, eds. *Advances in Mechanics and Mathematics and Global Optimization*, Vol. 17. Springer Science, New York, 195–233, 2009.

[29] A. Deza, T. Terlaky, and Y. Zinchenko. Polytopes and arrangements: Diameter and curvature. *Operations Research Letters* 36:215–222, 2008.

[30] A. Deza, T. Terlaky, and Y. Zinchenko. The continuous d-step conjecture for polytopes. *Discrete and Computational Geometry* 41(2):318–327, 2009.

[31] A. Deza, E. Nematollahi, R. Peyghami, and T. Terlaky. The central path visits all the vertices of the Klee-Minty cube. *Optimization Methods and Software* 21:851–865, 2006.

[32] A. V. Fiacco and G. P. McCormick. *Nonlinear Programming: Sequential Unconstrained Minimization Techniques*. John Wiley & Sons, New York, 1968.

[33] R. Frisch. The logarithmic potential method for solving linear programming problems. Memorandum, University Institute of Economics, Oslo, Norway, 1955.

[34] K. Fritzsche and F. Holt. More polytopes meeting the conjectured Hirsch bound. *Discrete Mathematics* 205:77–84, 1999.

[35] A. R. Ghaffari Hadigheh, O. Romanko, and T. Terlaky. Sensitivity analysis in convex quadratic optimization: Simultaneous perturbation of the objective and right-hand-side vectors. *Algorithmic Operations Research* 2(2):94–111, 2007.

[36] M. X. Goemans and D. P. Williamson. Improved approximation algorithms for maximum cut and satisfiability problems using semidefinite programming. *Journal of the ACM* 42:1115–1145, 1994.

[37] A. J. Goldman and A. W. Tucker. Theory of linear programming. H. W. Kuhn and A. W. Tucker, eds. *Linear Inequalities and Related Systems*, *Annals of Mathematical Studies*, Vol. 38. Princeton University Press, Princeton, NJ, 53–97, 1956.

[38] J. Gondzio and T. Terlaky. A computational view of interior point methods for linear programming. J. E. Beasley, ed. *Advances in Linear and Integer Programming*. Oxford University Press, Oxford, UK, 103–185, 1996.

[39] C. C. Gonzaga. An algorithm for solving linear programming problems in $O(n^3L)$ operations. N. Megiddo, ed. *Progress in Mathematical Programming: Interior Point and Related Methods*. Springer-Verlag, New York, 1–28, 1989.

[40] H. J. Greenberg. The use of the optimal partition in a linear programming solution for postoptimal analysis. *Operations Research Letters* 15:179–186, 1994.

[41] M. Grötschel, L. Lovász, and A. Schrijver. *Geometric Algorithms and Combinatorial Optimization*. Springer-Verlag, Berlin, 1988.

[42] B. Grünbaum. *Convex Polytopes*. Springer-Verlag, New York, 2003.

[43] F. Holt and V. Klee. Many polytopes meeting the conjectured Hirsch bound. *Discrete and Computational Geometry* 20:1–17, 1998.

[44] B. Jansen. *Interior Point Techniques in Optimization. Complexity, Sensitivity and Algorithms*, *Applied Optimization*, Vol. 6. Kluwer APC, Dordrecht, The Netherlands, 1996.

[45] B. Jansen, C. Roos, and T. Terlaky. The theory of linear programming: Skew symmetric self-dual problems and the central path. *Optimization* 29:225–233, 1994.

[46] B. Jansen, J. J. de Jong, C. Roos, and T. Terlaky. Sensitivity analysis in linear programming: Just be careful! *European Journal of Operations Research* 101(1):15–28, 1997.

[47] G. Kalai and D. Kleitman. A quasi-polynomial bound for the diameter of graphs of polyhedra. *Bulletin of the American Mathematical Society* 26:315–316, 1992.

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

[49] L. Khachiyan. A polynomial algorithm in linear programming. *Soviet Mathematics Doklady* 20(1):191–194, 1979.

[50] V. Klee and G. J. Minty. How good is the simplex algorithm. O. Shisha, ed. *Inequalities III.* Academic Press, New York, 159–175, 1972.

[51] V. Klee and D. Walkup. The d-step conjecture for polyhedra of dimension $d < 6$. *Acta Mathematica* 133:53–78, 1967.

[52] T. Koltai and T. Terlaky. The difference between managerial and mathematical interpretation of sensitivity analysis results in linear programming. *International Journal of Production Economics* 65:257–274, 200.

[53] G. Malajovich, J.-P. Dedieu, and M. Shub. On the curvature of the central path of linear programming theory. *Foundations of Computational Mathematics* 5:145–171, 2005.

[54] K. A. McShane, C. L. Monma, and D. F. Shanno. An implementation of a primal-dual interior point method for linear programming. *ORSA Journal on Computing* 1:70–83, 1989.

[55] N. Megiddo. Pathways to the optimal set in linear programming. N. Megiddo, ed. *Progress in Mathematical Programming: Interior Point and Related Methods.* Springer-Verlag, New York, 131–158, 1989.

[56] N. Megiddo. On finding primal and dual optimal bases. *ORSA Journal on Computing* 3:63–65, 1991.

[57] N. Megiddo and M. Shub. Boundary behavior of interior point algorithms in linear programming. *Mathematics of Operations Research* 14(1):97–146, 1989.

[58] S. Mehrotra. On the implementation of a (primal-dual) interior point method. *SIAM Journal on Optimization* 2(4):575–601, 1992.

[59] S. Mehrotra and Y. Ye. On finding the optimal facet of linear programs. *Mathematical Programming* 62:497–515, 1993.

[60] R. D. C. Monteiro and S. Mehrotra. A general parametric analysis approach and its implications to sensitivity analysis in interior point methods. *Mathematical Programming* 72:65–82, 1996.

[61] E. Nematollahi. Worst-case examples of interior-point trajectories. Ph.D. thesis, McMaster University, Hamilton, ON, Canada, 2008.

[62] E. Nematollahi and T. Terlaky. A redundant Klee-Minty construction with all the redundant constraints touching the feasible region. *Operations Research Letters* 36(4):414–418, 2008.

[63] E. Nematollahi and T. Terlaky. A simpler and tighter redundant Klee-Minty construction. *Optimization Letters* 2(3):403–414, 2008.

[64] Y. E. Nesterov and A. S. Nemirovski. *Interior-Point Polynomial Algorithms in Convex Programming*, *SIAM Studies in Applied Mathematics*, Vol. 13. SIAM Publications, Philadelphia, 1994.

[65] Y. E. Nesterov and M. J. Todd. Self-scaled barriers and interior-point methods for convex programming. *Mathematics of Operations Research* 22(1):1–42, 1997.

[66] J. Peng, C. Roos, and T. Terlaky. New complexity analysis of the primal-dual Newton method for linear optimization. *Annals of Operations Research* 99:23–39, 2000.

[67] J. Peng, C. Roos, and T. Terlaky. *Self-Regularity: A New Paradigm for Primal-Dual Interior-Point Algorithms.* Princeton University Press, Princeton, NJ, 2002.

[68] I. Pólik and T. Terlaky. A survey of the S-lemma. *SIAM Review* 49(3):371–418, 2007.

[69] I. Pólik and T. Terlaky. Interior point methods for nonlinear optimization. G. DiPillo and F. Schoen, eds. *Nonlinear Optimization. C.I.M.E. Lecture Series.* Springer Science, Heidelberg, Germany, 209–271, 2009.

[70] F. Potra and R. Sheng. On homogeneous interior-point algorithms for semidefinite programming. *Optimization Methods and Software* 9(3):161–184, 1998.

[71] J. Renegar. A polynomial-time algorithm, based on Newton's method, for linear programming. *Mathematical Programming* 40:59–93, 1988.

[72] J. Renegar. *A Mathematical View of Interior-Point Methods in Convex Optimization.* SIAM, Philadelphia, 2001.

[73] C. Roos. An $O(n^3L)$ approximate center method for linear programming. S. Dolecki, ed. *Optimization: Proceedings of the 5th French-German Conference in Castel-Novel (Varetz), France, October 1988*, *Lecture Notes in Mathematics*, Vol. 1405. Springer-Verlag, Berlin, 147–158, 1989.

[74] C. Roos. An exponential example for Terlaky's pivoting rule for the criss-cross simplex method. *Mathematical Programming* 46:79–84, 1990.

[75] C. Roos, T. Terlaky, and J.-Ph. Vial. *Interior Point Methods for Linear Optimization*, 2nd ed. Springer, New York, 2006.

[76] R. Saigal. *Linear Programming: A Modern Integrated Analysis.* Kluwer APC, Dordrecht, The Netherlands, 1995.

[77] G. Sonnevend. An "analytic center" for polyhedrons and new classes of global algorithms for linear (smooth, convex) programming. A. Prékopa, J. Szelezsán, and B. Strazicky, eds. *System Modeling and Optimization: Proceedings of the 12th IFIP-Conference Held in Budapest, Hungary, September 1985*, *Lecture Notes in Control and Information Sciences*, Vol. 84. Springer Verlag, Berlin, 866–876, 1986.

[78] G. Sonnevend, J. Stoer, and G. Zhao. On the complexity of following the central path of linear programs by linear extrapolation. *Mathematical Programming* 52(1-3):527–553, 1991.

[79] J. F. Sturm. Primal-dual interior point approach to semidefinite programming. J. B. G. Frenk, C. Roos, T. Terlaky, and S. Zhang, eds. *High Performance Optimization.* Kluwer APC, Dordrecht, The Netherlands, 1999.

[80] J. F. Sturm. Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones. *Optimization Methods and Software* 11–12:625–653, 1999.

[81] T. Terlaky, ed. *Interior Point Methods of Mathematical Programming*, *Applied Optimization*, Vol. 5. Kluwer APC, Dordrecht, The Netherlands, 1996.

[82] T. Terlaky. An easy way to teach interior-point methods. *European Journal of Operational Research* 130(1):1–19, 2001.

[83] M. J. Todd and Y. Ye. A lower bound on the number of iterations of long-step and polynomial interior-point linear programming algorithms. *Annals of Operations Research* 62:233–252, 1996.

[84] K. C. Toh, R. H. Tütüncü, and M. J. Todd. On the implementation of SDPT3 (version 3.1)—A Matlab software package for semidefinite-quadratic-linear programming. *Proceedings of the IEEE Conference on Computer-Aided Control System Design*, IEEE, Piscataway, NJ, 290–296, 2004.

[85] A. W. Tucker. Dual systems of homogeneous linear relations. H. W. Kuhn and A. W. Tucker, eds. *Linear Inequalities and Related Systems*, *Annals of Mathematical Studies*, Vol. 38. Princeton University Press, Princeton, NJ, 3–18, 1956.

[86] P. M. Vaidya. A new algorithm for minimizing convex functions over convex sets. *Mathematical Programming* 73(3):291–341, 1996.

[87] L. Vandenberghe and S. Boyd. Semidefinite programming. *SIAM Review* 38:49–95, 1996.

[88] R. J. Vanderbei. *LOQO User's Guide–Version 4.05.* Department of Operations Research and Financial Engineering, School of Engineering and Applied Science, Princeton University, Princeton, NJ, 2006.

[89] S. Vavasis and Y. Ye. A primal-dual interior-point method whose running time depends only on the constraint matrix. *Mathematical Programming* 74:79–120, 1996.

[90] A. Wächter. An interior point algorithm for large-scale nonlinear optimization with applications in process engineering. Ph.D. thesis, Carnegie Mellon University, Pittsburgh, 2002.

[91] A. Wächter and L. T. Biegler. On the implementation of a primal-dual interior point filter line search algorithm for large-scale nonlinear programming. *Mathematical Programming* 106(1):25–57, 2006.

[92] H. Wolkowicz, R. Saigal, and L. Vandenberghe, eds. *Handbook of Semidefinite Programming: Theory, Algorithms, and Applications.* Kluwer APC, Boston, 2000.

[93] M. H. Wright. The interior-point revolution in optimization: History, recent developments, and lasting consequences. *Bulletin of the American Mathematical Society* 42(1):39–56, 2004.

[94] S. J. Wright. *Primal-Dual Interior-Point Methods*. SIAM, Philadelphia, 1997.

[95] X. Xu, P. Hung, and Y. Ye. A simplified homogeneous and self-dual linear programming algorithm and its implementation. *Annals of Operations Research* 62:151–172, 1996.

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

[97] Y. Ye, M. J. Todd, and S. Mizuno. An $O(\sqrt{n}L)$-iteration homogeneous and self-dual linear programming algorithm. *Mathematics of Operations Research* 19:53–67, 1994.

[98] G. Ziegler. *Lectures on Polytopes*. Springer-Verlag, Heidelberg, Germany, 1995.