# Interior-Point Polynomial Algorithms in Convex Programming

Yurii Nesterov
Arkadii Nemirovskii

siam Studies in Applied Mathematics

# Interior-Point Polynomial Algorithms in Convex Programming

# SIAM Studies in Applied and Numerical Mathematics

This series of monographs focuses on mathematics and its applications to problems of current concern to industry, government, and society. These monographs will be of interest to applied mathematicians, numerical analysts, statisticians, engineers, and scientists who have an active need to learn useful methodology.

*Series List*

Vol. 1 *Lie-Bäcklund Transformations in Applications*
Robert L. Anderson and Nail H. Ibragimov

Vol. 2 *Methods and Applications of Interval Analysis*
Ramon E. Moore

Vol. 3 *Ill-Posed Problems for Integrodifferential Equations in Mechanics and Electromagnetic Theory*
Frederick Bloom

Vol. 4 *Solitons and the Inverse Scattering Transform*
Mark J. Ablowitz and Harvey Segur

Vol. 5 *Fourier Analysis of Numerical Approximations of Hyperbolic Equations*
Robert Vichnevetsky and John B. Bowles

Vol. 6 *Numerical Solution of Elliptic Problems*
Garrett Birkhoff and Robert E. Lynch

Vol. 7 *Analytical and Numerical Methods for Volterra Equations*
Peter Linz

Vol. 8 *Contact Problems in Elasticity: A Study of Variational Inequalities and Finite Element Methods*
N. Kikuchi and J. T. Oden

Vol. 9 *Augmented Lagrangian and Operator-Splitting Methods in Nonlinear Mechanics*
Roland Glowinski and P. Le Tallec

Vol. 10 *Boundary Stabilization of Thin Plate Splines*
John E. Lagnese

Vol. 11 *Electro-Diffusion of Ions*
Isaak Rubinstein

Vol. 12 *Mathematical Problems in Linear Viscoelasticity*
Mauro Fabrizio and Angelo Morro

Vol. 13 *Interior-Point Polynomial Algorithms in Convex Programming*
Yurii Nesterov and Arkadii Nemirovskii

Vol. 14 *The Boundary Function Method for Singular Perturbation Problems*
Adelaida B. Vasil'eva, Valentin F. Butuzov, and Leonid V. Kalachev

Vol. 15 *Linear Matrix Inequalities in System and Control Theory*
Stephen Boyd, Laurent El Ghaoui, Eric Feron, and Venkataramanan Balakrishnan

Vol. 16 *Indefinite-Quadratic Estimation and Control: A Unified Approach to $H^2$ and $H^\infty$ Theories*
Babak Hassibi, Ali H. Sayed, and Thomas Kailath

*Yurii Nesterov and Arkadii Nemirovskii*

# Interior-Point Polynomial Algorithms in Convex Programming

*Society for Industrial and Applied Mathematics*

siam *Philadelphia*

**Library of Congress Cataloging-in-Publication Data**

Nesterov, I͡U. E.
Interior-point polynomial algorithms in convex programming
/Yurii Nesterov, Arkadii Nemirovskii.
p. cm. — (SIAM studies in applied mathematics ; vol.13)
Includes bibliographical references and index.
ISBN 0-89871-319-6
ISBN 0-89871-515-6 (pbk.)
1. Programming (Mathematics) 2. Convex programming
I. Nemirovski, Arkadii Semenovich. II.Title. III. Series: SIAM
studies in applied mathematics ; 13
QA402.5.N4615 1993
519.7'6--dc20 93-38912

# Contents

# Foreword

In this book, Nesterov and Nemirovskii describe the first *unified theory* of polynomial-time interior-point methods. Their approach provides a simple and elegant framework in which all known polynomial-time interior-point methods can be explained and analyzed. Perhaps more important for applications, their approach yields polynomial-time interior-point methods for a very wide variety of problems beyond the traditional linear and quadratic programs.

The book contains new and important results in the general *theory* of convex programming, e.g., their "conic" problem formulation in which duality theory is completely symmetric. For each algorithm described, the authors carefully derive precise bounds on the computational effort required to solve a given family of problems to a given precision. In several cases they obtain better problem complexity estimates than were previously known.

The detailed proofs and lack of "numerical examples" might suggest that the book is of limited value to the reader interested in the practical aspects of convex optimization, but nothing could be further from the truth. An entire chapter is devoted to potential reduction methods precisely because of their great efficiency in practice (indeed, some of these algorithms are worse than path-following methods from the complexity theorist's point of view). Although it is not reported in this book, several of the new algorithms described (e.g., the projective method) have been implemented, tested on "real world" problems, and found to be extremely efficient in practice.

Nesterov and Nemirovskii's work has profound implications for the applications of convex programming. In many fields of engineering we find convex problems that are not linear or quadratic programs, but are of the form readily handled by their methods. For example, convex problems involving matrix inequalities arise in control system engineering. Before Nesterov and Nemirovskii's work, we could observe that such problems can be solved in polynomial time (by, e.g., the ellipsoid method) and therefore are, at least in a *theoretical* sense, tractable. The methods described in this book make these problems tractable in *practice.*

Karmakar's contribution was to demonstrate the first algorithm that solves linear programs in polynomial time and with practical efficiency. Similarly, it is one of Nesterov and Nemirovskii's contributions to describe algorithms that solve, in polynomial time and with practical efficiency, an extremely wide class of convex problems beyond linear and quadratic programs.

Stephen Boyd *Stanford, California*

# Acknowledgments

During our five-year work on this book, we had (and intensively used) the opportunity to discuss the subject and to collaborate with many of our colleagues. We are greatly indebted to Farid Alizadeh, Alfred Auslender, Aharon Ben-Tal, Stephen Boyd, Ronald Freund, Eugene Golsthein, Osman Guler, Dick den Hertog, Florian Jarre, Leonid Khachiyan, Ken Kortanek, Claude Lemarechal, Olvi Mangasarian, Boris Polyak, Roman Polyak, Cornelis Roos, Naum Shor, Linus Shrage, Michael Todd, Pravin Vaidya, Jean-Philippe Vial, and Yinyu Ye for their advice and remarks; in particular, the courses given by Michael Todd and Ronald Freund on the basis of initial version of the book helped us to hopefully improve the order of exposition. We are especially thankful to Stephen Boyd, Claude Lemarechal, and Olvi Mangasarian for encouraging us to write this monograph.

It was our pleasure to collaborate with SIAM in processing the manuscript. We highly appreciate valuable comments of the anonymous referees, which helped to improve the initial text. We are greatly impressed by the professionalism of Acquisitions Editor Susan Ciambrano, and also by her care and patience.

Having finished this book in Paris, we express our gratitude to Claude Lemarechal and Jean-Philippe Vial for their hospitality and for the fine working facilities we were given.

# Preface

The purpose of this book is to present the general theory of interior-point polynomial-time methods for convex programming. Since the publication of Karmarkar's famous paper in 1984, the area has been intensively developed by many researchers, who have focused on linear and quadratic programming. This monograph has given us the opportunity to present in one volume all of the major theoretical contributions to the theory of complexity for interior-point methods in optimization. Our aim is to demonstrate that all known polynomial-time interior-point methods can be explained on the basis of general theory, which allows these methods to extend into a wide variety of nonlinear convex problems. We also have presented for the first time a definition and analysis of the self-concordant barrier function for a compact convex body.

The abilities of the theory are demonstrated by developing new polynomial-time interior-point methods for many important classes of problems: quadratically constrained quadratic programming, geometrical programming, approximation in $L_p$ norms, finding extremal ellipsoids, and solving problems in structural design. Problems of special interest covered by the approach are those with positive semidefinite matrices as variables. These problems include numerous applications in modern control theory, combinatorial optimization, graph theory, and computer science.

This book has been written for those interested in optimization in general, including theory, algorithms, and applications. Mathematicians working in numerical analysis and control theory will be interested, as will computer scientists who are developing theory for computation of solutions of problems by digital computers. We hope that mechanical and electrical engineers who solve convex optimization problems will find this a useful reference.

Explicit algorithms for the aforementioned problems, along with detailed theoretical complexity analysis, form the main contents of this book. We hope that the theory presented herein will lead to additional significant applications.

*This page intentionally left blank*

# Chapter 1

# Introduction

## 1.1 Subject

The introduction of polynomial-time interior-point methods is one of the most remarkable events in the development of mathematical programming in the 1980s. The first method of this family was suggested for linear programming in the landmark paper of Karmarkar (see [Ka 84]). An excellent complexity result of this paper, as well as the claim that the performance of the new method on real-world problems is significantly better than the one of the simplex method, made this work a sensation and subsequently inspired very intensive and fruitful studies.

Until now, the activity in the field of interior-point methods focuses mainly on linear programming. At the same time, we find that the nature of the methods, is in fact, independent of the specific properties of LP problems, so that these methods can be extended onto more general convex programs. The aim of this book is twofold:

- To present a general approach to the design of polynomial-time interior-point methods for nonlinear convex problems, and
- To illustrate the abilities of the approach by a number of important examples (quadratically constrained quadratic programming, geometrical programming, approximation in $L_p$ norm, minimization of eigenvalues, among others).

## 1.2 Essence of the approach

After the seminal paper of Renegar (see [Re 86]), it became absolutely clear that the new polynomial-time algorithms belong to the traditional class of *interior penalty methods* studied in the classical monograph of Fiacco and McCormick (see [FMcC 68]). To solve a convex problem

$$(f) \qquad \text{minimize } f_0(x) \text{ s.t. } f_i(x) \le 0,\ i = 1, ..., m(f), \qquad x \in \mathbf{R}^{n(f)}$$

by an interior penalty method, it is first necessary to form a barrier function for the feasible domain

$$G_f = \{x \mid f_i(x) \le 0,\ i = 1, ..., m(f)\}$$

of the problem, i.e., smooth and strongly convex on the interior of the domain function $F$ tending to infinity along each sequence of interior points converging to a boundary point of $G_f$. Given such a barrier, one approximates the constrained problem $(f)$ by the family of unconstrained problems, e.g., by the *barrier-generated family*

$$(f_t) \qquad \text{minimize } f_t(x) = tf_0(x) + F(x),$$

where $t > 0$ is the penalty parameter. Under extremely mild restrictions, the solutions $x(t)$ to $(f_t)$ tend to the optimal set of $(f)$ as $t$ tends to $\infty$. The classical scheme suggests following the trajectory $x(t)$ along certain sequence $t_i \to \infty$ of values of the penalty. By applying to $(f_t)$ a method for unconstrained minimization, one forms "tight" approximations to $x(t_i)$, and these approximations are regarded as approximate solutions to $(f)$. This scheme leads to *barrier methods.*

Another "unconstrained approximation" of the constrained problem $(f)$ is given by the family

$$(f_t^c) \qquad \text{minimize } f_t^c(x) = \phi(t - f_0(x)) + F(x),$$

where $t > f^*$ ($f^*$ is the optimal value in $(f)$) and $\phi$ is a barrier for the nonnegative half-axis. As $t \to f^* + 0$, the solutions $x^c(t)$ to the problems $(f_t^c)$ tend to the optimal set of $(f)$, and one can follow the path $x^c(t)$ along a sequence $t_i \to f^* + 0$ by applying to $(f_t^c)$ a method for unconstrained minimization. The latter scheme originating from Huard (see, e.g., [BH 66]) leads to what is called *methods of centers.*

Note that the above schemes possess two main "degrees of freedom": First, it is possible to use various barriers; second, one can implement any method for unconstrained minimization. Regarding the first issue, the classical recommendation, at least in the case of smooth convex constraints, is to use barriers that are compositions of constraints,

$$F(x) = \sum_{i=1}^{m(f)} \psi(-f_i(x)),$$

where $\psi(s)$ is a barrier for the nonnegative half-axis, e.g.,

$$\psi(s) = s^{-\kappa}, \quad \kappa > 0; \qquad \psi(s) = -\kappa \ln s, \quad \kappa > 0; \qquad \psi(s) = \mathrm{e}^{1/s}, \quad \text{etc.}$$

Regarding choice of the method for unconstrained minimization, there were almost no firm theoretical priorities; the computational experience was in favor of the Newton method, but this recommendation had no theoretical background.

Such a background was first given by Renegar in [Re 86]. Renegar demonstrated that in the case of a linear programming problem $(f)$ (all $f_i$, $i =$

$1,\ldots,m(f)$, are linear), the method of centers associated with the *standard logarithmic barrier*

$$F(x) = -\sum_{i=1}^{m} \ln(-f_i(x))$$

for the feasible polytope $G_f$ of the problem and with

$$\phi(s) = -\omega \ln(s), \qquad \omega > 0$$

allows us to decrease the residual $t_i - f^*$ at a linear rate at the cost of a *single* step of the Newton method as applied to $(f^c_{t_i})$. Under appropriate choice of the weight $\omega$ at the term $\ln(t - f_0(x))$ (namely, $\omega = O(m(f))$), one can force the residual $t_i - f^*$ to decrease as $\exp\{-O(1)i/m^{1/2}(f)\}$. Thus, to improve the accuracy of the current approximate solution by an absolute constant factor, it suffices to perform $O(m^{1/2}(f))$ Newton steps, which requires a polynomial in *the size* $(n(f), m(f))$ of the problem number of arithmetic operations; in other words, the method proves to be *polynomial*. Similar results for the barrier method associated with the same logarithmic barrier for a linear programming problem were established by Gonzaga [Go 87].

We see that the central role in the modern interior-point methods for linear programming is played by the standard logarithmic barrier for the feasible polytope of the problem. To extend the methods onto nonlinear problems, one should understand the properties of the barrier responsible for polynomiality of the associated interior-point methods. Our general approach originates in [Ns 88b], [Ns 88c], [Ns 89]. It is as follows: Among all various properties of the logarithmic barrier, only two are responsible for all nice features of the associated with $F$ interior-point methods. These two properties are (i) the Lipschitz continuity of the Hessian $F''$ of the barrier with respect to the local Euclidean metric defined by the Hessian itself as

$$|D^3F(x)[h,h,h]| \leq \text{const}_1\{D^2F(x)[h,h]\}^{3/2}$$

for all $x$ from the interior of $G$ and all $h \in \mathbf{R}^n$; and (ii) the Lipschitz continuity of the barrier itself with respect to the same local Euclidean structure

$$|DF(x)[h]| \leq \text{const}_2\{D^2F(x)[h,h]\}^{1/2}$$

for the same as above $x$ and $h$.

Now (i) and (ii) do not explicitly involve the polyhedral structure of the feasible domain $G$ of the problem; given an arbitrary closed convex domain $G$, we can consider a interior penalty function for $G$ with these properties (such a function will be called a *self-concordant barrier* for $G$). The essence of the theory is that, given a self-concordant barrier $F$ for a closed convex domain $G$, we can associate with this barrier interior point methods for minimizing *linear* objectives over $G$ in the same way as is done in the case of the standard logarithmic barrier for a polytope. Moreover, all polynomial-time interior-point methods known for LP admit the above extension. To improve the accuracy of a given approximate solution by an absolute constant factor, the resulting methods require the amount of steps that depends only on the *parameter of the barrier*, i.e., on certain combination of the above $\mathrm{const}_1$ and $\mathrm{const}_2$, while each of the steps is basically a step of the Newton minimization method as applied to $F$.

Note that the problem of minimizing a linear objective over a closed convex domain is universal for convex programming: Each convex program can be reformulated in this form. It follows that the possibility to solve convex programs with the aid of interior-point methods is limited only by our ability to point out self-concordant barriers for the resulting feasible domains. The result is that such a barrier always exists (with the parameter being absolute constant times the dimension of the domain); unfortunately, to obtain nice complexity results, we need a barrier with moderate arithmetic cost of computing the gradient and the Hessian, which is not always the case. Nevertheless, in many cases we can to point out "computable" self-concordant barriers, so that we can develop efficient methods for a wide variety of nonlinear convex problems of an appropriate analytical structure.

Thus, we see that there exist not only heuristic, but also theoretical reasons for implementing the Newton minimization method in the classical schemes of the barrier method and the method of centers. Moreover, we understand how to use the freedom in choice of the barrier: It should be self-concordant, and we are interested in this intrinsic property, in contrast to the traditional recommendations where we are offered a number of possibilities for constructing the barrier but have no priorities for choosing one of them.

## 1.3 Motivation

In our opinion, the main advantage of interior-point machinery is that, in many important cases, it allows us to utilize the knowledge of the analytical structure of the problem under consideration to develop an efficient algorithm. Consider a family $\mathcal{A}$ of solvable optimization problems of the type $(f)$ with convex finite (say, on the whole $\mathbf{R}^{n(f)}$) objective and constraints.

Assume that we have fixed analytical structure of the functionals involved into our problems, so that each problem instance $(f)$ belonging to $\mathcal{A}$ can be identified by a finite-dimensional real vector $\mathsf{D}(f)$ ("the set of coefficients of the instance"). Typical examples here are the classes of linear programming problems, linearly/quadratically constrained convex quadratic problems, and so forth. Assume that, when solving $(f)$, the set of data $\mathsf{D}(f)$ form the input to the algorithm, and we desire to solve $(f)$ to a prescribed accuracy $\varepsilon$, i.e., to find an approximate solution $x_\varepsilon$ satisfying the relations

$$f_0(x_\varepsilon) \le f^* + \varepsilon, \quad f_i(x_\varepsilon) \le \varepsilon, \quad i = 1, \dots, m(f),$$

where $f^*$ is the optimal value in $(f)$.

An algorithm that transforms the input $(\mathsf{D}(f), \varepsilon)$ into an $\varepsilon$-solution to $(f)$ in a finite number of operations of precise real arithmetic will be called *polynomial*, if the total amount of these operations for all $(f) \in \mathcal{A}$ and all $\varepsilon > 0$ is bounded from above by $p(m(f), n(f), \dim\{\mathsf{D}(f)\}) \ln(V(f)/\varepsilon)$, where $p$ is a polynomial. Here $V(f)$ is certain *scale parameter*, which can depend on the magnitudes of coefficients involved into $(f)$ (a reasonable choice of the parameter is specific for the family under consideration). The ratio $\varepsilon/V(f)$ can be regarded as the relative accuracy, so that $\ln(V(f)/\varepsilon)$ is something like the amount of accuracy digits in an $\varepsilon$-solution. Thus, a polynomial-time algorithm is a procedure in which the arithmetic cost "per accuracy digit" does not exceed a polynomial of the *problem size* $(m(f), n(f), \dim\{\mathsf{D}(f)\})$. Polynomiality usually is treated as theoretical equivalent to the unformal notion "an effective computational procedure," and the efficiency of a polynomial-time algorithm, from the theoretical viewpoint, is defined by the corresponding "cost per digit" $p(m(f), n(f), \dim\{\mathsf{D}(f)\})$.

The concept of a polynomial-time algorithm was introduced by Edmonds [Ed 65] and Cobham [Co 65] (see also Aho et al. [AHU 76], Garey and Johnson [GJ 79], and Karp [Kr 72], [Kr 75]). This initial concept was oriented onto discrete problems; in the case of continuous problems with real data, it seems to be more convenient to deal with the above (relaxed) version of this concept.

Note that polynomial-time algorithms do exist in a sense, for "all" convex problems. Indeed, there are procedures (e.g., the ellipsoid method; see [NY 79]) that solve all convex problems $(f)$ to relative (in a reasonable scale) accuracy $\varepsilon$ at the cost of $O(p(n,m)\ln(n/\varepsilon))$ arithmetic operations and $O(q(n,m)\ln(n/\varepsilon))$ computations of the values and subgradients of the objective and the constraints, where $p$ and $q$ are polynomials (for the ellipsoid method, $p(n,m) = n^3(m+n)$, $q(n,m) = n^2$). Now, if our class of problems $\mathcal{A}$ is such that, given the data $\mathsf{D}(f)$, we can compute the above values and subgradients at a given point $x$ in polynomial in $m(f)$, $n(f)$, $\dim\{\mathsf{D}(f)\}$ number of arithmetic operations, then the above procedure proves to be polynomial on $\mathcal{A}$.

A conceptual drawback of the latter scheme is that, although from the very beginning we possess complete information about the problem instance, we make only "local" conclusions from this "global" information; in fact, in this scheme, we ignore our knowledge of the analytical structure of the problem under consideration (more accurately, this information is used only when computing the values and the subgradients of $f_i$). At the same time, the interior-point machinery is now the only known way to utilize the knowledge of analytical structure to improve—sometimes significantly—the theoretical efficiency of polynomial-time algorithms. Indeed, as already mentioned, the efficiency of a polynomial-time interior-point method is defined first by the parameter of the underlying barrier and second by the arithmetic cost at which one can form and solve the corresponding Newton systems; both these quantities depend more on the analytical structure of the objective and constraints than on the dimensions $m(f)$ and $n(f)$ of the problem.

## 1.4 Overview of the contents

Chapter 2 forms the technical basis of the book. Here we introduce and study our main notions of self-concordant functions and barriers.

Chapter 3 is devoted to the path-following interior-point methods. In their basic form, these methods allow us to minimize a linear objective $f$ over a bounded closed convex domain $G$, provided that we are given a self-concordant barrier for the domain and a starting point belonging to the interior of the domain. In a path-following method, the barrier and the objective generate certain penalty-type family of functions and, consequently, the trajectory of minimizers of these functions; this trajectory converges to the optimal set of the problem. The idea of the method is to follow this path of minimizers: Given a strictly feasible approximate solution close, in a sense, to the point of the path corresponding to a current value of the penalty parameter, we vary the parameter in the desired direction and then compute the Newton iterate of the current approximate solution to restore the initial closeness between the updated approximate solution and the new point of the path. Of course, this scheme is quite traditional, and, generally speaking, it does not result in polynomial-time procedure. The latter feature is provided by self-concordance of the functions comprising the family.

We demonstrate that path-following methods known for LP (i.e., for the case when $G$ is a polytope) can be easily explained and extended onto the case of general convex domains $G$. We prove that the efficiency ("cost per digit") of these methods is $O(\vartheta^{1/2})$, where $\vartheta$ is the parameter of the barrier (for the standard logarithmic barrier for an $m$-facet polytope one has $\vartheta = m$).

In Chapter 4 we extend onto the general convex case the potential reduction interior-point methods for LP problems; we mean the method of Karmarkar [Ka 84], the projective method [Nm 87], the primal-dual method of Todd and Ye [TY 87], and Ye [Ye 88a], [Ye 89]. The efficiency of the resulting method is $O(\vartheta)$ (for the generalized method of Karmarkar and the projective method) or $O(\vartheta^{1/2})$ (the generalized primal-dual method), where $\vartheta$ denotes the parameter of the underlying self-concordant barrier. Thus, the potential reduction methods, theoretically, have no advantages as compared to the path-following algorithms. From the computational viewpoint, however, these methods are much more attractive. The reason is that, for a potential reduction method, one can point out an explicit Lyapunov's function, and the accuracy of a feasible approximate solution can be expressed in terms of the potential (the less the potential, the better the approximate solution). At each strictly feasible solution, the theory prescribes a direction and a stepsize, which allows us to obtain a new strictly feasible solution with the value of the potential being "considerably" less than that at the previous approximate solution. To ensure the theoretical efficiency estimate, it suffices to perform this theoretical step, but we are not forbidden to achieve a deeper decreasing of the potential, say, with the aid of one-dimensional minimization of the potential in the direction prescribed by the theory. In real-world problems, these "large steps" significantly accelerate the method. In contrast to this, in a path-following method, we should maintain closeness to the corresponding trajectory, which, at least theoretically, is an obstacle for "large steps."

To extend potential reduction interior-point methods onto the general convex case, we use a special reformulation of a convex programming problem, the so-called conic setting of it (where we should minimize a linear functional over the intersection of an affine subspace and a closed convex cone). An important role in the extension is played by duality, which for conic problems attains very symmetric form and looks quite similar to the usual LP duality. Another advantage of the "conic format" of convex programs, which is especially important to the design of polynomial-time methods, is that this format allows us to exploit the widest group of transformations preserving convexity of feasible domains; we mean the projective transformations (to subject a conic problem to such a transformation is basically the same as to intersect the cone with another affine subspace).

As already mentioned, to solve a convex problem by an interior-point method, we should first reduce the problem to one of minimizing a linear objective over convex domain (which is quite straightforward) and, second, point out a "computable" self-concordant barrier for this domain (which is the crucial point for the approach). As shown in Chapter 2, every $n$-dimensional closed convex domain admits a self-concordant barrier with the parameter of order of $n$; unfortunately, the corresponding "universal barrier" is given by a multivariate integral and therefore cannot be treated as "computable." Nevertheless, the result is that there exists a kind of calculus of "computable" self-concordant barriers, which forms the subject of Chapter 5. We first point out "simple" self-concordant barriers for a number of standard domains arising in convex programming (epigraphs of standard functions on the axis, level sets of convex quadratic forms, the epigraph of the Euclidean norm, the cone of positive semidefinite matrices, and so forth). Second, we demonstrate that all standard (preserving convexity) operations with convex domains (taking images/inverse images under affine mappings and projective transformations, intersection, taking direct products, and so forth) admit simple rules for combining self-concordant barriers for the operands into a self-concordant barrier for the resulting domain. This calculus involves "rational linear algebra" tools only and, as applied to our "raw materials"—concrete self-concordant barriers for "standard" convex sets—allows us to form "computable" self-concordant barriers for a wide variety of convex domains arising in convex programming.

In Chapter 6 we illustrate the abilities of the developed technique. Namely, we present polynomial-time interior-point algorithms for a number of classes of nonlinear convex programs, including quadratically constrained quadratic programming, geometrical programming (in exponential form), approximation in $L_p$-norm, and minimization of the operator norm of a matrix linearly depending on the control vector. An especially interesting application is semidefinite programming, i.e., minimization of a linear functional of a symmetric matrix subjected to positive semidefiniteness restriction and a number of linear equality constraints. Note that, first, semidefinite programming is a nice field for interior-point methods (all path-following and potential reduction methods can be easily implemented for this class); second, semidefinite programming covers many important problems arising in various areas, from control theory to combinatorial optimization (e.g., the problem of minimizing the largest eigenvalue or the sum of $k$ largest eigenvalues of a symmetric matrix). We conclude Chapter 6 with developing polynomial-time interior-point algorithms for two geometrical problems concerning extremal ellipsoids (the problems are to inscribe the maximum volume ellipsoid into a given polytope and to cover a given finite set in $\mathbf{R}^n$ by the ellipsoid of minimum volume). The first of these is especially interesting for nonsmooth convex optimization (it arises as an auxiliary problem in the *inscribed ellipsoid method* (see Khaciyan et al. [KhTE 88])).

Chapter 7 is devoted to variational inequalities with monotone operators. Here we extend the notion of self-concordance onto monotone operators and develop a polynomial-time path-following method for inequalities involving operators "compatible," in a sense, with a self-concordant barrier for the feasible domain of the inequality. Although the compatibility condition is a rather severe restriction, it is automatically satisfied for linear monotone operators, as well as for some interesting nonlinear operators (e.g., the operator arising, under some natural assumptions, in the pure exchange model of Arrow–Debreu).

In Chapter 8 we consider possibilities for acceleration of the path-following algorithms as applied to linearly constrained convex quadratic (in particular, LP) problems. Until now, the only known acceleration strategies were more or less straightforward modifications of the Karmarkar speed-up (see [Ka 84]) based on recursive updatings of approximate inverses to the matrices arising at the sequential Newton-type steps of the procedure under consideration. We describe four more strategies: Three are based on following the path with the aid of (prescaled) multistep methods for smooth optimization; in the fourth strategy, to find an approximate solution of a Newton system, we use the prescaled conjugate gradient method. All our strategies lead to the same worst-case complexity estimates as the known ones, but the new strategies seem to be more flexible and therefore can be expected to be more efficient in practice.

We conclude the exposition with Bibliography Comments. It seems to be impossible to give a detailed survey of the activity in the very intensively developing area of polynomial-time interior-point methods. Therefore we have restricted ourselves only with the papers closely related to the monograph. We realize that the "level of completeness" of our comments is far from being perfect and apologize in advance for possible lacunae.

The methods presented in the book are new, and we believe that they are promising for practical computations. The very preliminary experience we now possess supports this hope, but it in no sense is sufficient to make definite conclusions. Therefore our decision was to completely omit any numerical results.

Of course, we realize that it is computational experience, not theoretical results alone, that proves practical potential of an algorithm, and we hope that, by the methods presented in this book, this experience can soon be gained.

## 1.5 How to read this book

Basically, there are two ways of reading this book, depending on whether the reader is interested in the interior-point theory itself or its applications to concrete optimization problems. The theoretical aspect is detailed in Chapters 2–5, while applications of variational inequalities, in addition to the theory, can be found in Chapter 7 (it is possible to exclude Chapters 4 and 5 here). For specific explications of the theory of linear and linearly constrained quadratic programming, refer to Chapter 8.

Chapters 2 and 3 deal in general theory rather than in concrete applications; the reader interested in applications is expected to be familiar with the main concepts and results found there (with the exception of §2.5 and possible §2.4), but not necessarily with the proofs. Note that, for some applications (e.g., geometrical programming, approximation in $L_p$-norm, and finding extremal ellipsoids), only path-following methods are developed, and, consequently, those interested in these applications may move from chapter 3 directly to Chapter 6. Quadratically constrained quadratic problems, especially semidefinite programming with applications to control theory, can be found in §2.4 and Chapter 4 (at the level of concepts and schemes). If one wishes to deal with concrete applications that are not explicitly presented in this book and would like to attempt to develop new interior-point methods, refer to Chapter 5 for the techniques of constructing self-concordant barriers.

*This page intentionally left blank*

# Chapter 2

# Self-concordant functions and Newton method

All interior-point algorithms heavily exploit the classical Newton minimization method as applied to "unconstrained smooth approximations" of the problem under consideration. To be successful, this scheme requires approximations that can be effectively minimized by the Newton method. The approach presented in this book deals with a special class of functions used in these approximations, namely, *self-concordant* functions and barriers.

In this chapter, which forms the technical basis for the book, we introduce and study the notions of a self-concordant function (§2.1) and a self-concordant barrier (§2.3). We demonstrate (§2.2) that self-concordant functions form a nice field for the Newton method, which allows us to develop general theory of polynomial-time path-following algorithms (Chapter 3); as demonstrated by the general existence theorem (§2.5), the related methods *in principle* can be used to solve an *arbitrary* convex problem. Section 2.4 is devoted to the Legendre transformation of self-concordant functions and barriers. The results of this section form the background for potential reduction methods (Chapter 4).

## 2.1 Self-concordant functions

To motivate the notion of self-concordance, let us start with analyzing the traditional situation in which, given some strongly convex function $F : \mathbf{R}^n \to \mathbf{R}$, one desires to minimize it by the Newton method. In its classical form, the method generates the iterates

$$x_{i+1} = x_i - (F''(x_i))^{-1} F'(x_i).$$

It is well-known that, under mild regularity assumptions, this scheme locally converges quadratically. The typical result here is as follows: If $F$ is strongly convex with some constant $m$, i.e.,

$$h^T F''(x) h \geq m \parallel h \parallel_2^2, \qquad x, h \in \mathbf{R}^n$$

and $F''$ is Lipschitz continuous with constant $L$, i.e.,

$$\parallel (F''(x) - F''(y))h \parallel_2 \leq L \parallel x - y \parallel_2 \parallel h \parallel_2, \qquad x, y, h \in \mathbf{R}^n,$$

then

$$\| F'(x_{i+1}) \|_2 \leq \frac{L}{2m^2} \| F'(x_i) \|_2^2 .$$

In particular, in the region defined by the inequality

$$\frac{L}{2m^2} \| F'(x_i) \|_2 < 1,$$

the method converges quadratically.

Note that the above sufficient condition for quadratic convergence of the Newton method involves an Euclidean structure on $\mathbf{R}^n$; one should choose such a structure to define the Hessian matrix and consequently the constants $m$ and $L$. Thus, the description of the region of quadratic convergence depends on an ad hoc choice of Euclidean structure, which, generally speaking, has nothing in common with $F$. The resulting uncertainty contradicts the affine-invariant nature of the Newton method.

Now note that the second-order differential of $F$ induces an infinitesimal Euclidean metric on $\mathbf{R}^n$, intrinsically connected with $F$. The result is that the Lipschitz continuity of $F''$ with respect to this metric implies very interesting results on the behaviour of the Newton method as applied to $F$. This property is called *self-concordance.* The precise definition is as follows.

**Definition 2.1.1** *Let $E$ be a finite-dimensional real vector space, $Q$ be an open nonempty convex subset of $E$, $F : Q \to \mathbf{R}$ be a function, $a > 0$. $F$ is called* self-concordant *on $Q$ with the parameter value $a$ ($a$-self-concordant; notation: $F \in S_a(Q, E)$), if $F \in C^3$ is a convex function on $Q$, and, for all $x \in Q$ and all $h \in E$, the following inequality holds:*

$$| D^3F(x)[h, h, h] | \leq 2a^{-1/2}(D^2F(x)[h, h])^{3/2} \tag{2.1.1}$$

*($D^kF(x)[h_1, ..., h_k]$ henceforth denotes the value of the $k$th differential of $F$ taken at $x$ along the collection of directions $h_1, ..., h_k$).*

*An $a$-self-concordant on $Q$ function $F$ is called* strongly $a$-self-concordant *(notation: $F \in S_a^+(Q, E)$) if the sets $\{x \in Q \mid F(x) \leq t\}$ are closed in $E$ for each $t \in \mathbf{R}$.*

**Remark 2.1.1** *Note that a self-concordant on $Q$ function $F$ is strongly self-concordant on $Q$ if and only if $Q$ is the "natural domain" of $F$, i.e., if and only if $F(x_i)$ tends to infinity along every sequence $\{x_i \in Q\}$ converging to a boundary point of $Q$.*

Self-concordance is an affine-invariant property; see the following proposition.

**Proposition 2.1.1** (i) Stability with respect to affine substitutions. *Let $F \in S_a(Q, E)$ ($F \in S_a^+(Q, E)$), let $x = \mathcal{A}(y) = Ay + b$ be an affine transformation from a finite-dimensional real vector space $E^+$ into $E$ such that $Q^+ \equiv \{y \mid$*

*$\mathcal{A}(y) \in Q\} \neq \emptyset$, and let $F^+(y) \equiv F(\mathcal{A}(y)) : Q^+ \to \mathbf{R}$. Then $F^+ \in S_a(Q^+, E^+)$ (respectively, $F^+ \in S_a^+(Q^+, E^+)$).*

(ii) Stability under summation. *Let $F_i \in S_{a_i}(Q_i, E)$, $p_i > 0$, $i = 1, 2$, $Q \equiv Q_1 \bigcap Q_2 \neq \emptyset$, $F(x) = p_1 F_1(x) + p_2 F_2(x) : Q \to \mathbf{R}$ and let $a = \min\{p_1 a_1, p_2 a_2\}$. Then $F \in S_a(Q, E)$. If under the above assumptions either $F_i \in S_{a_i}^+(Q_i, E)$, $i = 1, 2$, or $F_1 \in S_{a_1}^+(Q_1, E)$ and $Q_1 \subseteq Q_2$, then $F \in S_a^+(Q, E)$.*

(iii) Stability with respect to direct products. *Let $F_1(x) \in S_a(Q_1, E_1)$, $F_2(y) \in S_a(Q_2, E_2)$. Then $F_1(x) + F_2(y) \in S_a(Q_1 \times Q_2, E_1 \times E_2)$. If $F_1(x) \in S_a^+(Q_1, E_1)$, $F_2(y) \in S_a^+(Q_2, E_2)$, then $F_1(x) + F_2(y) \in S_a^+(Q_1 \times Q_2, E_1 \times E_2)$.*

The proof is quite straightforward.

Let $F$ be self-concordant on $Q$ with the parameter value $a$ and let $x \in Q$. The second-order differential $D^2F(x)$ defines the Euclidean seminorm on $E$ as follows:

$$\| h \|_{x,F} \equiv \left(\frac{1}{a} D^2F(x)[h,h]\right)^{1/2}.$$

An $r$-neighbourhood of $x$ in this seminorm, i.e., the set

$$W_r(x) = \{y \mid \| x - y \|_{x,F} \leq r\}$$

will be called *Dikin's ellipsoid* of $F$ centered at $x$ of the radius $r$.

The following statement is a finite-difference version of the differential inequality (2.1.1); this statement is used in what follows as the main technical tool.

**Theorem 2.1.1** *Let $F$ be $a$-self-concordant on $Q$ and let $x \in Q$. Then*

(i) *For each $y \in Q$ such that $r \equiv \| x - y \|_{x,F} < 1$ we have, for all $h \in E$,*

$$(2.1.2) \quad (1 - r)^2 D^2F(x)[h,h] \leq D^2F(y)[h,h] \leq \frac{1}{(1-r)^2} D^2F(x)[h,h];$$

(ii) *If $F$ is strongly self-concordant, then every Dikin's ellipsoid of $F$ centered at $x$ of the radius $r < 1$ is contained in $Q$, namely,*

$$r < 1 \quad \Rightarrow \quad W_r(x) \subset Q.$$

**Proof.** (i) Denote

$$e = y - x, \quad x(s) = x + se, \quad \delta = \{s \geq 0 \mid \| x(s) - x \|_{x,F} < 1\}.$$

Let $s \in \delta$ be such that $x(s) \in Q$ and let $h \in E$. Let $\Delta = [0, s]$ and let the functions $\psi(\rho)$ and $\phi(\rho)$ be defined for $\rho \in \Delta$ as

$$\psi(\rho) = D^2F(x(\rho))[e,e],$$

$$\phi(\rho) = D^2F(x(\rho))[h,h].$$

From (2.1.1) it follows (see §8.4.3) that, for each triple of vectors $h_i \in E$, $i = 1, 2, 3$, we have

$$| D^3F(u)[h_1, h_2, h_3] | \leq 2a^{-1/2} \prod_{i=1}^{3} \{D^2F(u)[h_i, h_i]\}^{1/2}, \qquad u \in Q.$$

The latter relation implies that

$$(2.1.3) \qquad | \psi'(\rho) | \leq 2a^{-1/2}(\psi(\rho))^{3/2}, \quad | \phi'(\rho) | \leq 2a^{-1/2}(\psi(\rho))^{1/2}\phi(\rho).$$

From the first relation in (2.1.3), we have either $\psi(\rho) \equiv 0$, $\rho \in \Delta$, and therefore, by virtue of the second relation in (2.1.3), $\phi(s) = \phi(0)$, or $\psi(\rho) > 0$ on $\Delta$.[1]

In the second case, we have $| (\psi^{-1/2}(\rho))'_\rho | \leq a^{-1/2}$, $\rho \in \Delta$, which implies that

$$(2.1.4) \qquad \psi^{-1/2}(\rho) \geq \psi^{-1/2}(0) - \rho a^{-1/2}.$$

In the latter case, by virtue of $\psi(0) = a \parallel y - x \parallel_{x,F}^2 = ar^2$, we obtain

$$\psi^{1/2}(\rho) \leq \frac{a^{1/2}r}{1 - \rho r},$$

since $\psi^{-1/2}(0) > \rho a^{-1/2}$ for $\rho \in \Delta \subseteq \delta$. Consequently, the second relation in (2.1.3) can be written as

$$| \phi'(\rho) | \leq \frac{2r\phi(\rho)}{1 - \rho r}, \qquad \rho \in \Delta.$$

Thus, either $\phi \equiv 0$ on $\Delta$, or $\phi$ is positive on $\Delta$, and in the latter case we have

$$| \ln \frac{\phi(s)}{\phi(0)} | \leq 2 \ln \frac{1}{1 - sr},$$

which, by the definition of $\phi$, leads to (2.1.2). Evidently, (2.1.2) also holds in the cases where $\psi(\rho) \equiv 0$, $\rho \in \Delta$, and $\phi(\rho) = \phi(0)$, $\rho \in \Delta$. Thereby (i) is proved.

(ii) We should prove that, if $y \in E$ is such that $r \equiv \parallel y - x \parallel_{x,F} < 1$, then $y \in Q$. Let $e$, $x(s)$, $\delta$ be defined in the same manner as above and let

$$\sigma = \sup\{s \in \delta \mid x(s) \in Q\}.$$

It suffices to prove that $\sigma = 1/r$ $(1/0 = \infty)$. Assume that the latter relation does not hold, so that $\sigma r < 1$. Applying (i) to the points $y = x(s)$, $0 \leq s < \sigma$ (these points do belong to $Q$), we conclude that the function $g(s) = F(x(s))$

[1] We have used the well-known consequence of the uniqueness theorem for linear differential equations: If an absolute continuous real-valued function $f$ defined on a segment $\Delta \subset \mathbf{R}$ satisfies the inequality $| f'(t) | \leq g(t) | f(t) |$, where $g$ is summable, then either $f = 0$ identically on $\Delta$ or $f$ does not take zero value on this segment.

has a bounded second derivative for $0 \le s < \sigma$ and hence itself is bounded for these $s$. Since $F \in S_a^+(Q, E)$, this leads to

$$x(\sigma) = \lim_{s \to \sigma - 0} x(s) \in Q.$$

Since $Q$ is open, we have $x(s) \in Q$ for some $s > \sigma$. The latter inclusion in the case of $\sigma r < 1$ contradicts the definition of $\sigma$, and thereby (ii) is proved. □

**Corollary 2.1.1** *Let $F$ be self-concordant on $Q$. Then the subspace*

$$\{h \in E \mid D^2F(x)[h,h] = 0\}$$

*does not depend on $x \in Q$.*

**Proof.** Let us fix $h \in E$. The set $X_h = \{x \in Q \mid D^2F(x)[h,h] = 0\}$ is closed in $Q$ by virtue of the continuity of $D^2F(x)[h,h]$ in $x$ and is open in $Q$ by virtue of (2.1.2). Hence this set is either empty, or coincides with $Q$. □

The subspace mentioned in Corollary 2.1.1 will be called the *recessive subspace* of $F$ and will be denoted $E_F$. We call $F$ *nondegenerate*, if $E_F = \{0\}$.

## 2.2 Damped Newton method

In this section, we describe the behaviour of the Newton method as applied to a self-concordant function.

### 2.2.1 Newton decrement

Let us start with introducing a convenient accuracy measure—the *Newton decrement* $\lambda(F, x)$. Let $F$ be $a$-self-concordant on $Q$. We define $\lambda(F, x)$, $x \in Q$ as

$$\lambda(F,x) = \inf\{\lambda \mid \ \mid DF(x)[h] \mid \le \lambda a^{1/2}(D^2F(x)[h,h])^{1/2} \quad \forall h \in E\} \tag{2.2.1}$$

(if the set on the right is empty, then $\lambda(F, x) = \infty$ by definition). The quantity $\lambda(F, x)$ can be interpreted as follows. Consider the quadratic approximation

$$\Phi_x(y) = F(x) + DF(x)[y-x] + \tfrac{1}{2}D^2F(x)[y-x, y-x] \tag{2.2.2}$$

of the function $F$ at the point $x$. Then we clearly have

$$\frac{a}{2}\lambda^2(F,x) = \Phi_x(x) - \inf\{\Phi_x(y) \mid y \in E\} \tag{2.2.3}$$

or, which is the same,

$$a\lambda^2(F,x) = 2\sup\{DF(x)[h] - \tfrac{1}{2}D^2F(x)[h,h] \mid h \in E\}. \tag{2.2.4}$$

In the case of nondegenerate $F$ (i.e., $E_F = \{0\}$), the Newton decrement can be expressed in terms of the gradient vector, and the Hessian of $F$ taken with respect to (any) Euclidean structure of $E$ as

$$\lambda^2(F,x) = \frac{1}{a}(F'(x))^T[F''(x)]^{-1}F'(x).$$

It is worth noting that $\lambda(F,x)$ is, up to a constant factor, the norm of the linear form $DF(x)$ in the metric induced by $D^2F(x)$.

**Proposition 2.2.1** *For any $F \in S_a(Q,E)$, either $\lambda(F,x) = \infty$ for all $x \in Q$, or $\lambda(F,x)$ is a finite continuous function on $Q$. This latter case occurs if and only if $DF(x)[h] = 0$ for all $x \in Q$ and all $h \in E_F$.*

**Proof.** Let $x \in Q$. It is clear that

$$\{\lambda(F,x) < \infty\} \Leftrightarrow \{DF(x)[h] = 0 \quad \forall h \in E_F\}.$$

Assume that $\lambda(F,x) < \infty$ for chosen $x$ and let $h \in E_F$. For $\psi(y) = DF(y)[h]$, we have $D\psi(y)[e] = D^2F(y)[h,e] = 0$, so that $\psi$ is constant on $Q$; since $\psi(x) = 0$, we have $\psi \equiv 0$. Thus, if $\lambda(F,x) < \infty$, then $DF(y)[h] = 0$ for all $h \in E_F$ and all $y \in E$, so that

$$a^{1/2}\lambda(F,y) = \max\{|\ DF(y)[h]\ |\ (D^2F(y)[h,h])^{-1/2} \mid h \in E^F,\ h \neq 0\},$$

where $E^F$ is a subspace complement to $E_F$ in $E$. The quadratic form $D^2F(y)$ $[h,h]$ is positive definite on the subspace $E^F$; therefore the statement follows from the $C^3$-smoothness of $F$. □

### 2.2.2 Preliminary results

Now we are ready to analyze the behaviour of the Newton method as applied to a self-concordant function.

Let $F$ be $a$-self-concordant on $Q$, let $x \in Q$, and let $\lambda(F,x) < \infty$. Then the form $\Phi_x(y)$ is below bounded in $y \in E$ and therefore attains its minimum over $y$. Let $x^*(F,x)$ be some minimizer of this form and let $e(F,x) = x^*(F,x) - x$. We clearly have

$$DF(x)[h] = -D^2F(x)[e(F,x),h] \quad \forall h \in E, \tag{2.2.5}$$

$$D^2F(x)[e(F,x),e(F,x)] = a\lambda^2(F,x). \tag{2.2.6}$$

The point $x^*(F,x)$ is called the *Newton iterate* of $x$. In the nondegenerate case, as it should be,

$$e(F,x) = -[F''(x)]^{-1}F'(x), \quad x^*(F,x) = x - [F''(x)]^{-1}F'(x). \tag{2.2.7}$$

The following result describes the behaviour of the quantities $\lambda(F,\cdot)$ and $F(\cdot)$ on the segment $[x, x^*(F,x)]$.

**Theorem 2.2.1** *Let $F \in S_a(Q, E)$, $x \in Q$, and $\lambda(F, x) < \infty$. Let*

$$\sigma \in (0, \min\{1, \lambda^{-1}(F, x)\}]$$

*be such that the points $x(s) = x + se(F, x)$ belong to $Q$ for all $s \in \Delta \equiv [0, \sigma)$.*
*Then for all $s \in \Delta$ we have*

$$\lambda(F, x(s)) \leq \lambda \frac{1 - s - s\lambda + 2s^2\lambda}{(1 - s\lambda)^2}, \tag{2.2.8}$$

$$F(x) - F(x(s)) \geq a\lambda^2 \left\{ s\frac{1 + \lambda}{\lambda} + \frac{1}{\lambda^2} \ln(1 - s\lambda) \right\}, \tag{2.2.9}$$

*where $\lambda = \lambda(F, x)$.*

**Proof.** Let $e = e(F, x)$. Note that by (2.2.6) $\lambda = \| e \|_{x,F}$. Since $\sigma \leq 1/\lambda$, we have by virtue of Theorem 2.1.1(i), for all $s \in \Delta$ and $h \in E$,

$$(1 - s\lambda)^2 D^2F(x)[h, h] \leq D^2F(x(s))[h, h] \leq \frac{1}{(1 - s\lambda)^2} D^2F(x)[h, h], \tag{2.2.10}$$

whence for $s \in \Delta$

$$| D^2F(x)[h, h] - D^2F(x(s))[h, h] | \leq \left\{ \frac{1}{(1 - s\lambda)^2} - 1 \right\} D^2F(x)[h, h]. \tag{2.2.11}$$

It follows that

$$\begin{aligned} &\left| \frac{d}{ds} DF(x(s))[h] - D^2F(x)[e, h] \right| \\ &\leq \left\{ \frac{1}{(1 - s\lambda)^2} - 1 \right\} (D^2F(x)[h, h])^{1/2} (D^2F(x)[e, e])^{1/2} \\ &\leq \left\{ \frac{1}{(1 - s\lambda)^2} - 1 \right\} (D^2F(x)[h, h])^{1/2} a^{1/2} \lambda, \end{aligned} \tag{2.2.12}$$

or, in view of (2.2.5),

$$| DF(x(s))[h] - (1 - s)DF(x)[h] | \leq a^{1/2} \frac{(s\lambda)^2}{1 - s\lambda} (D^2F(x)[h, h])^{1/2}. \tag{2.2.13}$$

Now, by definition of $\lambda(F,x)$ combined with (2.2.13) and (2.2.10) we have

$$
\begin{aligned}
&2\sup\left\{DF(x(s))[h] - \frac{1}{2}D^2F(x(s))[h,h] \,\Big|\, h \in E\right\} \\
&\le 2\sup\left\{(1-s)DF(x)[h] + \frac{a^{1/2}(s\lambda)^2}{1-s\lambda}(D^2F(x)[h,h])^{1/2}\right. \\
&\qquad \left. -\frac{1}{2}(1-s\lambda)^2 D^2F(x)[h,h] \mid h \in E\right\} \\
&\le 2\sup\left\{a^{1/2}\lambda(1-s)(D^2F(x)[h,h])^{1/2}\right. \\
&\qquad \left. +\frac{a^{1/2}(s\lambda)^2}{1-s\lambda}(D^2F(x)[h,h])^{1/2} - \frac{1}{2}(1-s\lambda)^2 D^2F(x)[h,h] \,\Big|\, h \in E\right\} \\
&\le a\lambda^2\left(\frac{1-s-s\lambda+2s^2\lambda}{(1-s\lambda)^2}\right)^2,
\end{aligned}
$$

which, combined with (2.2.1), implies (2.2.8) (note that by definition $\sigma \le 1$ and therefore $(1-s) \ge 0$).

Let $f(s) = F(x(s)) - F(x)$. Relation (2.2.13) with $h = e$ and relation (2.2.5) lead to

$$
\begin{aligned}
f'(s) = DF(x(s))[e] &\le (1-s)DF(x)[e] + \frac{a^{1/2}(s\lambda)^2}{1-s\lambda}(D^2F(x)[e,e])^{1/2} \\
&= -(1-s)D^2F(x)[e,e] + \frac{a^{1/2}(s\lambda)^2}{1-s\lambda}(D^2F(x)[e,e])^{1/2} \\
&= -(1-s)a\lambda^2 + \frac{a\lambda(s\lambda)^2}{1-s\lambda}.
\end{aligned}
$$

Hence

$$
f(s) \le f(0) - a\lambda^2 \int_0^s \left\{1 - \rho - \frac{\rho^2\lambda}{1-\rho\lambda}\right\} d\rho,
$$

which implies (2.2.9). □

The following modification of the above theorem sometimes is more convenient.

**Theorem 2.2.2** *Let $F \in S_a(Q,E)$, $x \in Q$, and $\lambda(F,x) < 1$. Let one of the sets*

$$
X_x = \{y \in Q \mid \lambda(F,y) \le \lambda(F,x)\}, \qquad Y_x = \{y \in Q \mid F(y) \le F(x)\}
$$

*be closed in $E$. Then*

(i) *$F$ attains its minimum over $Q$ : $X_* \equiv \operatorname{Argmin}_Q F \ne \emptyset$;*

(ii) *If $\lambda(F,x) \le \lambda_* \equiv 2 - 3^{1/2} = 0.2679...$, then $x^*(F,x) \in Q$ and*

$$
\lambda(F,x^*(F,x)) \le \frac{\lambda^2(F,x)}{(1-\lambda(F,x))^2} \le \frac{1}{2}\lambda(F,x); \tag{2.2.14}
$$

(iii) *For each* $y \in Q$ *such that* $\lambda(F,y) < \frac{1}{3}$ *and for each* $x_* \in X_*$, *we have*

$$\frac{1}{a}\left\{F(y) - \min_Q F\right\} \leq \frac{1}{2}\omega^2(\lambda(F,y))\frac{1+\omega(\lambda(F,y))}{1-\omega(\lambda(F,y))}, \tag{2.2.15}$$

$$D^2F(x_*)[y-x_*, y-x_*] \leq a\left(\frac{\omega^2(\lambda(F,y))}{1-\omega(\lambda(F,y))}\right)^2, \tag{2.2.16}$$

$$D^2F(y)[x_*-y, x_*-y] \leq a\omega^2(\lambda(F,y)), \tag{2.2.17}$$

*where* $\omega(\lambda) = 1-(1-3\lambda)^{1/3}$;

(iv) *For each* $y \in Q$ *such that* $\delta^2(F,y) \equiv 2a^{-1}\{F(y) - \min_Q F\} < 4/9$, *we have*

$$\lambda(F,y) \leq \frac{24\delta(F,y)}{(3+(9+12\delta(F,y))^{1/2})((9-12\delta(F,y))^{1/2}-1)^2}. \tag{2.2.18}$$

**Proof.** $1^0$. Let $\sigma(\lambda) = \min\{1, (1-\lambda)/\lambda(3-\lambda)\}$, $\Delta_\lambda = [0,\sigma(\lambda)]$, and

$$\phi_\lambda(s) = \frac{1-s-s\lambda+2s^2\lambda}{(1-s\lambda)^2}, \quad \psi_\lambda(s) = \frac{s(1+\lambda)}{\lambda} + \frac{1}{\lambda^2}\ln(1-s\lambda), \quad s \in \Delta_\lambda,$$

for $0 \leq \lambda < 1$. It is easy to show that in view of choice of $\sigma(\lambda)$ the function $\phi_\lambda(s)$ decreases on $\Delta_\lambda$, and the function $\psi_\lambda(s)$ is nonnegative on $\Delta_\lambda$. Let $X = X_x \bigcap Y_x$ and $u \in X$. Then $\lambda \equiv \lambda(F,u) \leq \lambda(F,x) < 1$. Let

$$s' = \sup\{s \in \Delta_\lambda \mid u + se(F,u) \in Q\}.$$

By virtue of (2.2.8), (2.2.9), for $0 \leq s < s'$, we have

$$\lambda(F, u+se(F,u)) \leq \lambda\phi_\lambda(s) \leq \lambda,$$

$$F(u+se(F,u)) \leq F(u) - a\lambda^2\psi_\lambda(s) \leq F(u). \tag{2.2.19}$$

The sets $X_x$ and $Y_x$ are closed in $Q$ (since $\lambda(F,\cdot)$ and $F(\cdot)$ are continuous on $Q$), and one of the sets is closed in $E$; hence $X$ is closed in $E$. By (2.2.19) $u + se(F,u) \in X$, $0 \leq s < s'$, and by virtue of closedness of $X$ in $E$ we have $u + s'e(F,u) \in X$. For continuity reasons, it follows that (2.2.19) holds for $s = s'$. Thus, $u + s'e(F,u) \in X \subset Q$. Since $Q$ is open, the inclusion $u + s'e(F,u) \in Q$ is possible only if $s' = \sigma(\lambda)$ (see the definition of $s'$). Thus, we obtain

$$u \in X \Rightarrow u^*(u) \equiv u + \sigma(\lambda(F,u))e(F,u) \in X$$

and

$$\lambda(F,u^*(u)) \leq \lambda(F,u)\phi_{\lambda(F,u)}(\sigma(\lambda(F,u))), \tag{2.2.20}$$

$$F(u^*(u)) \leq F(u) - a\lambda^2(F,u)\psi_{\lambda(F,u)}(\sigma(\lambda(F,u))). \tag{2.2.21}$$

$2^0$. Assume that $u \in X$ is such that $\lambda(F,u) \leq 2-3^{1/2}$. Then $\sigma(\lambda(F,u)) = 1$, whence $u^*(u) = x^*(F,u) \in Q$, and, by (2.2.20),

$$\lambda(F,x^*(F,u)) \leq \phi_{\lambda(F,u)}(1)\lambda(F,u) = \frac{\lambda^2(F,u)}{(1-\lambda(F,u))^2} \leq \frac{2-3^{1/2}}{(3^{1/2}-1)^2}\lambda(F,u),$$

which leads to (2.2.14).

$3^0$. Consider the following process:

$$x_0 = x; \quad x_{i+1} = u^*(x_i) \equiv x_i + \sigma(\lambda(F, x_i))e(F, x_i), \quad i \geq 0 \tag{2.2.22}$$

(the process is well defined by the arguments of $1^0$, namely, $x_i \in X \subset Q$ for all $i \geq 0$). This is the Newton method as applied to $F$ starting at $x$ with special choice of step length. We shall prove that, for all sufficiently large $i$, we have $\sigma(\lambda(F, x_i)) = 1$, which means that for those $i$ (2.2.22) is the standard Newton process.

Let $\lambda_i = \lambda(F, x_i)$. Then, by virtue of (2.2.20), (2.2.21),

$$\lambda_{i+1} \leq \lambda_i \phi_{\lambda_i}(\sigma(\lambda_i)), \quad F(x_{i+1}) < F(x_i), \tag{2.2.23}$$

which implies that $\lambda_i \to 0$, $i \to \infty$ (recall that $\phi_\lambda(s)$, as a function of $s$, decreases on the segment $\Delta_\lambda$ and equals 1 at $s = 0$). In particular, for all sufficiently large $i$, we have $\lambda_i < \lambda_*$, or $\sigma(\lambda_i) = 1$, as claimed. Let $i_* = \min\{i \mid \lambda_i < \lambda_*\}$. Then for $i \geq i_*$ we have, by virtue of (2.2.14),

$$\lambda_{i+1} \leq \frac{\lambda_i^2}{(1-\lambda_i)^2} < \frac{1}{2}\lambda_i. \tag{2.2.24}$$

Note that the behaviour of $\lambda_i$ depends solely on $\lambda_0 \equiv \lambda(F, x)$ (this quantity must be less than 1), and $\lambda_i$ quadratically converges to 0 by virtue of (2.2.24).

$4^0$. Let us now prove (i). We can assume that $\lambda_i > 0$, $i > 0$, since otherwise (i) is evident. Let $E^F$ be a subspace complement to $E_F$ in $E$. Let

$$V_i = \{y \in x_i + E^F \mid D^2F(x_i)[y - x_i, y - x_i] \leq 100a\lambda_i^2\}.$$

Then $V_i$ is a compact set, because $D^2F(x)[\cdot,\cdot]$ is positive definite on $E^F$. Let

$$\omega_{i,\varepsilon}(s) = a\left\{\frac{1}{2}s^2 + \varepsilon\left(\lambda_i s + \int_0^s \frac{\rho^2}{1-\rho}d\rho\right)\right\}$$

for $\varepsilon = \pm 1$. Assume that an integer $i \geq 2$ is such that, for $s_i = 10\lambda_i$, we have

$$s_i < 1;$$

$$\omega_{i,-1}(s_i) > 0; \quad \omega_{i,1}(s) \leq F(x_0) - F(x_1), \quad 0 \leq s < s_i;$$

$$\frac{s + \lambda_i - \lambda_i s}{(1-s)^{-2}} < \lambda_0, \qquad 0 \leq s \leq s_i \tag{2.2.25}$$

(since $\lambda_i > 0$ and $\lambda_i \to 0$, (2.2.25) holds for all sufficiently large $i$). Let us verify that, for the above value of $i$, the following inclusion holds:

$$V_i \subset X. \tag{2.2.26}$$

Indeed, let $e \in E^F$ satisfy $D^2F(x_i)[e, e] = a$ and let

$$\sigma_e = \sup\{s \in [0, 1] \mid x(s, e) \equiv x_i + se \in Q\}.$$

By virtue of Theorem 2.1.1, for all $s \in [0, \sigma_e)$ and $h \in E$ we have

$$| D^2F(x(s,e))[h,h] - D^2F(x_i)[h,h] | \leq \left( \frac{1}{(1-s)^2} - 1 \right) D^2F(x_i)[h,h],$$

$$D^2F(x(s,e))[h,h] \geq (1-s)^2 D^2F(x_i)[h,h],$$

which leads to

$$\left| \frac{d}{ds}(DF(x(s,e))[h]) - D^2F(x_i)[e,h] \right| \leq a^{1/2} \left( \frac{1}{(1-s)^2} - 1 \right) (D^2F(x_i)[h,h])^{1/2}.$$

Therefore

$$| DF(x(s,e))[h] - sD^2F(x_i)[e,h] - DF(x_i)[h] | \leq \frac{a^{1/2}s^2}{1-s}(D^2F(x_i)[h,h])^{1/2},$$

and, by Cauchy's inequality,

$$DF(x(s,e))[h] \leq a^{1/2}s(D^2F(x_i)[h,h])^{1/2} + DF(x_i)[h] + \frac{a^{1/2}s^2}{1-s}(D^2F(x_i)[h,h])^{1/2}.$$

Hence

$$\begin{aligned} & 2\sup\left\{ DF(x(s,e))[h] - \frac{1}{2}D^2F(x(s,e))[h,h] \mid h \in E \right\} \\ & \leq 2\sup\left\{ DF(x_i)[h] + a^{1/2}\left\{ s + \frac{s^2}{1-s} \right\} \{D^2F(x_i)[h,h]\right\}^{1/2} \\ & \qquad -\frac{1}{2}(1-s)^2 D^2F(x_i)[h,h] \mid h \in E \Big\} \\ & \leq 2\sup\left\{ a^{1/2}\left( \lambda_i + \frac{s}{1-s} \right) (D^2F(x_i)[h,h])^{1/2} \right. \\ & \qquad \left. -\frac{1}{2}(1-s)^2 D^2F(x_i)[h,h] \mid h \in E \right\} \\ & \leq \frac{a(s + \lambda_i - \lambda_i s)^2}{(1-s)^4}, \end{aligned}$$

which, in view of (2.2.1) and (2.2.25), leads to

$$\lambda(x(s,e)) < \lambda_0, \qquad 0 \leq s < \sigma_e. \tag{2.2.27}$$

Let $f(s) = F(x(s,e)) - F(x_i)$. Then, for $0 \leq s < \sigma_e$, we have

$$f''(s) = D^2F(x(s,e))[e,e],$$

so that, in view of Theorem 2.1.1(i),

$$| f''(s) - D^2F(x_i)[e,e] | \leq a \left( \frac{1}{(1-s)^2} - 1 \right).$$

Since $f'(0) = DF(x_i)[e]$, this leads to

$$sDF(x_i)[e]+a\left(\frac{s^2}{2}-\int_0^s \frac{\rho^2}{1-\rho}d\rho\right) \le f(s) \le sDF(x_i)[e]+a\left(\frac{s^2}{2}+\int_0^s \frac{\rho^2}{1-\rho}d\rho\right).$$

By virtue of $| DF(x_i)[e] | \le \lambda_i a^{1/2}(D^2F(x_i)[e,e])^{1/2} = a\lambda_i$, we now have

$$f(s) \le \omega_{i,1}(s), \quad f(s) \ge \omega_{i,-1}(s), \quad 0 \le s < \sigma_e^*, \tag{2.2.28}$$

where $\sigma_e^* = \min\{s_i, \sigma_e\}$. By virtue of (2.2.25) and (2.2.23), relations (2.2.27) and (2.2.28) imply that $x(s,e) \in X$ for $0 \le s < \sigma_e^*$. Since $X$ is closed in $E$, we have $x(\sigma_e^*, e) \in X$. Since $Q$ is open, it follows from the definition of $\sigma_e^*$ that the latter inclusion holds only if $\sigma_e > s_i$. This implies (2.2.26) because $e$ is an arbitrary vector from $E^F$ satisfying $D^2F(x_i)[e,e] = a$.

Note that the points belonging to the (relative) boundary $\partial V_i$ of $V_i$ can be represented as $x_i + s_i e$, $e \in E^F$, $D^2F(x_i)[e,e] = a$. Thus, taking into account (2.2.28) and (2.2.25), we obtain $F(u) > F(x_i)$, $u \in \partial V_i$. Hence there exists a point $x_* \in V_i$ such that $DF(x_*)[h] = 0$ for all $h \in E^F$. By virtue of Proposition 2.2.1, under the assumptions of the theorem, we have $DF(u)[h] = 0$ for all $u \in Q$ and $h \in E_F$. Hence $DF(x_*) = 0$, and (i) is thereby proved.

$5^0$. Let us prove (iii). Let

$$e = y - x_*, \quad \lambda = \lambda(F,y), \quad \omega = (D^2F(y)[e,e]/a)^{1/2},$$

$$x(s) = x_*+se, \quad y(s) = y-se = x(1-s), \quad f(s) = F(y(s))-F(x_*), \quad 0 \le s \le 1.$$

We have

$$DF(y(s))[e] = -f'(s) \ge 0, \qquad 0 \le s \le 1. \tag{2.2.29}$$

Let $\sigma = \min\{1, \omega^{-1}\}$. Since $(d/ds)\{-DF(y(s))[e]\} = D^2F(y(s))[e,e]$, by virtue of Theorem 2.2.1 we obtain

$$\frac{d}{ds}\{-DF(y(s))[e]\} \ge (1-s\omega)^2 D^2F(y)[e,e] = a\omega^2(1-s\omega)^2, \qquad 0 \le s \le \sigma,$$

and therefore

$$DF(y(s))[e] \le DF(y)[e]-a\omega^2\int_0^s(1-\rho\omega)^2 d\rho \le a\omega\left\{\lambda - s\omega\left(1 - 1s\omega + \frac{1}{3}s^2\omega^2\right)\right\},$$

$0 \le s \le \sigma$. The latter inequality combined with (2.2.29) implies that

$$3\lambda \ge s\omega(3 - 3s\omega + s^2\omega^2), \qquad 0 \le s \le \sigma. \tag{2.2.30}$$

If $\omega \ge 1$, then $\sigma = \omega^{-1}$ and (2.2.30) holds for $s = \omega^{-1}$. It follows that $\lambda \ge \frac{1}{3}$, which contradicts the assumptions of (iii). Hence $\omega < 1$, so that $\sigma = 1$, and

(2.2.30) implies $\omega(3-3\omega+\omega^2) \le 3\lambda$. In the latter inequality, the left-hand side is monotone in $\omega > 0$, so that

$$\omega \le \omega(\lambda) \equiv 1-(1-3\lambda)^{1/3} \tag{2.2.31}$$

($\omega(\lambda)$ is the unique root of the equation $\omega(3-3\omega+\omega^2) = 3\lambda$). By definition of $\omega$, (2.2.31) is equivalent to (2.2.17). Now let

$$g(s) = F(x(s)) - F(x(0))(= F(x(s)) - F(x_*)).$$

Then $g(0) = g'(0) = 0$, and, taking into account Theorem 2.2.1, we obtain

$$g''(s) = D^2F(x(s))[e,e] \le \frac{1}{(1-(1-s)\omega)^2} D^2F(y)[e,e] \quad \text{for } 0 \le s \le 1.$$

Therefore

$$\begin{aligned} g(1) &\le a\omega^2 \int_0^1 \int_0^s \frac{1}{(1-(1-\rho)\omega)^2} d\rho ds = a\left\{\frac{\omega}{1-\omega} + \ln(1-\omega)\right\} \\ &= a\left\{(\omega+\omega^2+\omega^3+\cdots) - \left(\omega + \frac{\omega^2}{2} + \frac{\omega^3}{3} + \cdots\right)\right\} \\ &\le a\left\{\frac{\omega^2}{2} + \frac{\omega^3}{1-\omega}\right\} = \frac{a\omega^2(1+\omega)}{2(1-\omega)}. \end{aligned}$$

Combining this relation with (2.2.31), we obtain (2.2.15). Furthermore, by virtue of (2.2.31) and Theorem 2.1.1, we have

$$D^2F(x_*)[e,e] \le \frac{1}{(1-\omega)^2} D^2F(y)[e,e],$$

which, combined with (2.2.31), implies (2.2.16). Thereby (iii) is proved.

$6^0$. Let us prove (iv). Let $x_* \in \operatorname{Argmin} F$ and let $y \in Q$ be such that

$$\delta^2 \equiv \tfrac{2}{a}(F(y) - F(x_*)) < \tfrac{4}{9}.$$

Let

$$e = y - x_*, \quad \omega = (D^2F(x_*)[e,e]/a)^{1/2}, \quad \sigma = \min\{1, \omega^{-1}\},$$
$$x(s) = x_* + se, \ 0 \le s \le 1, \quad f(s) = F(x(s)) - F(x_*).$$

By virtue of Theorem 2.1.1, we have $f''(s) \ge a\omega^2(1-\omega s)^2$, $0 \le s \le \sigma$. Since $f'(0) = 0$, it follows that

$$f(s) \ge a\omega^2 \int_0^s \int_0^t (1-z\omega)^2 dz dt = \frac{a\omega^2 s^2}{12}(6 - 4\omega s + \omega^2 s^2), \qquad 0 \le s \le \sigma.$$

If $\omega \geq 1$, then $\sigma = \omega^{-1}$, and we obtain $\delta^2 = 2a^{-1}f(1) \geq \frac{1}{2} > \frac{4}{9}$, which is a contradiction. Therefore $\omega < 1$, whence $\sigma = 1$, and the above inequality implies that

$$\left\{\tfrac{1}{6}\omega^2(6 - 4\omega + \omega^2) \leq \delta^2\right\} \ \& \ \{\omega < 1\}. \tag{2.2.32}$$

Hence

$$\left\{\tfrac{1}{9}\omega^2(3 - \omega)^2 \leq \delta^2\right\} \ \& \ \{\omega < 1\}. \tag{2.2.33}$$

For $0 \leq s \leq 1$ and $h \in E$, we have, by virtue of Theorem 2.2.1,

$$\begin{aligned}\left|\frac{d}{ds}DF(x(s))[h]\right| = \mid D^2F(x(s))[h,e] \mid &\leq (D^2F(x(s))[h,h])^{1/2}(D^2F(x(s))[e,e])^{1/2}\\ &\leq \frac{a^{1/2}\omega}{(1-s\omega)^2}(D^2F(x_*)[h,h])^{1/2},\end{aligned}$$

whence, in view of $DF(x_*)[h] = 0$, we obtain

$$\mid DF(y)[h] \mid \leq \frac{a^{1/2}\omega}{1-\omega}(D^2F(x_*)[h,h])^{1/2}.$$

By virtue of Theorem 2.1.1, we also have $D^2F(y)[h,h] \geq (1-\omega)^2D^2F(x_*)[h,h]$. The above inequalities imply that $\lambda(F,y) \leq \omega(1-\omega)^{-2}$, which, combined with (2.2.33), implies (2.2.18). □

### 2.2.3 Newton method: Summary

The following theorem provides us with complete description of the behaviour of the Newton method.

**Theorem 2.2.3** *Let $F \in S_a(Q,E), x \in Q$ and let the set $X = \{y \in Q \mid F(y) \leq F(x)\}$ be closed in $E$. Then*

(i) *$F$ is below bounded on $Q$ if and only if it attains its minimum over $Q$. If $\lambda(F,x) < 1$, then $F$ attains its minimum over $Q$;*

(ii) *Let $\lambda(F,x) < \infty$, $\lambda_* = 2 - 3^{1/2} = 0.2679...$ and $\lambda' \in [\lambda_*, 1)$.*

*Consider the* Newton iteration *starting at $x$,*

$$x_0 = x; \quad x_{i+1} = x_i + \sigma'(\lambda(F,x_i))e(F,x_i), \qquad i \geq 0, \tag{2.2.34}$$

*where*

$$e(F,x_i) \in \text{Argmin}\,\{DF(x_i)[h] + \tfrac{1}{2}D^2F(x_i)[h,h] \mid h \in E\}, \tag{2.2.35}$$

$$\sigma'(\lambda) = \begin{cases} \frac{1}{1+\lambda}, & \lambda > \lambda'; \\ \frac{1-\lambda}{\lambda(3-\lambda)}, & \lambda' \geq \lambda \geq \lambda_*; \\ 1, & \lambda < \lambda_*. \end{cases} \tag{2.2.36}$$

*The iterations are well defined (i.e., for all $i$ we have $x_i \in X, \lambda_i \equiv \lambda(F, x_i) < \infty$ and $e(F, x_i)$ is well defined), and the following relations hold:*

(2.2.37)
$$\{\lambda_i > \lambda'\} \Rightarrow$$
$$\Rightarrow \{F(x_{i+1}) \le F(x_i) - a(\lambda_i - \ln(1+\lambda_i)) \le F(x_i) - a(\lambda' - \ln(1+\lambda'))\};$$

(2.2.38)
$$\{\lambda' \ge \lambda_i \ge \lambda_*\} \Rightarrow$$
$$\Rightarrow \left\{\lambda_{i+1} \le \frac{6\lambda_i - \lambda_i^2 - 1}{4} < \lambda_i\right\} \& \left\{1 - \lambda_{i+1} \ge \frac{5-\lambda_i}{4}(1-\lambda_i) \ge \frac{5-\lambda'}{4}(1-\lambda_i)\right\};$$

$$\text{(2.2.39)} \qquad \{\lambda_i < \lambda_*\} \Rightarrow \left\{\lambda_{i+1} \le \left(\frac{\lambda_i}{1-\lambda_i}\right)^2 < \frac{\lambda_i}{2}\right\}.$$

*Moreover,*

$$\text{(2.2.40)} \qquad \lambda_i < 1/3 \Rightarrow F(x_i) - \min_Q F \le \frac{a\omega^2(\lambda_i)(1+\omega(\lambda_i))}{2(1-\omega(\lambda_i))}.$$

**Proof.** $1^0$. Let $\lambda(F, x) < \infty$. Let $J$ be the set of all integers $j \ge 0$ that satisfy the following conditions:

($1_j$) Process (2.2.34) is well defined for $0 \le i \le j$; i.e., for all these $i$, we have $x_i \in X$, $\lambda_i < \infty$, and $e(F, x_i)$ are well defined;
($2_j$) For $0 \le i < j$, the implications (2.2.38)–(2.2.39) are true.

Let us prove that $J = \{j \ge 0\}$. First, let us show that $0 \in J$. Indeed, ($2_0$) is evident, and to prove ($1_0$) we only must establish that $e(F, x)$ is well defined. This, however, follows immediately from Proposition 2.2.1 and the above assumption that $\lambda(F, u)$ is finite for $u = x$.

It remains to prove the implication $j \in J \Rightarrow j+1 \in J$. Let $j \in J$. Then $x_j \in X$, and $e_j \equiv e(F, x_j)$ is well defined.

Assume that $\lambda_j > \lambda'$. Let

$$\sigma = \sup\{s \in [0, \sigma'(\lambda_j)] \mid x(s) = x_j + se_j \in Q\}.$$

Then $\sigma \le \min\{1, \lambda_j^{-1}\}$. Using Theorem 2.2.1 (see (2.2.9)) and the inequality

$$\sigma \le (1+\lambda_j)^{-1} = \sigma'(\lambda_j),$$

we have in the case under consideration

$$\text{(2.2.41)} \quad F(x(s)) - F(x_j) \le -a\lambda_j^2\left(s\frac{1+\lambda_j}{\lambda_j} + \frac{1}{\lambda_j^2}\ln(1 - s\lambda_j)\right) \le 0$$

for $0 \le s < \sigma$. Thus, $x(s) \in X$, $0 \le s < \sigma$, and since $X$ is closed we have $x(\sigma) \in X$. By the definition of $\sigma$, the latter inclusion necessitates $\sigma = \sigma'(\lambda_j)$, which in turn implies that $x_{j+1} \in X$. Hence, in view of ($1_j$) and the above remarks on $\lambda(F,u)$ and $e(F,u)$, we are led to $(1_{j+1})$. Since $x_{j+1} \in X$, relation (2.2.41), by the continuity arguments, holds for $s = \sigma = \sigma'(\lambda_j)$. In view of $(2_j)$, this implies that $(2_{j+1})$. Thus, in the case under consideration, we have $j+1 \in J$.

Now let $\lambda_j < \lambda'$. Following the argument of item $1^0$ of the proof of Theorem 2.2.2 with $u = x = x_j$ (the assumptions of the theorem are satisfied since $X_{x_j}$ is closed in $E$ together with $X$ in view of $x_j \in X$) and making use of the fact that $\sigma'(\lambda) = \sigma(\lambda)$ for $\lambda < \lambda'$, we obtain $x_{j+1} \in X$, which combined with $(1_j)$ leads to $(1_{j+1})$. Relations (2.2.36), when applied to the above $u$, imply (2.2.38)–(2.2.39) for $i = j$. These latter relations, combined with $(2_j)$, lead to $(2_{j+1})$. Thus, $j+1 \in J$.

$2^0$. Now we are able to prove (ii). All statements in (ii), excluding (2.2.40), follow immediately from the system of implications $\{(1_j), (2_j) \mid j \geq 0\}$. Let us verify (2.2.40). Assume that $\lambda_i < 1/3$. The set

$$\{y \in Q \mid F(y) \leq F(x_i)\}$$

is closed in $E$ together with $X$, since $x_i \in X$. Therefore (2.2.40) follows from Theorem 2.2.2.(iii), where we set $x = x_i$, $y = x_i$.

$3^0$. It remains to prove (i). Under the assumptions of the theorem, the set $\{y \in Q \mid F(y) \leq F(x')\}$ is closed in $E$ for each $x' \in X$, so that, by Theorem 2.2.2.(i), the implication

$$(\exists x' \in X : \lambda(F,x') < 1) \Rightarrow F \textit{ attains its minimum over } Q \tag{2.2.42}$$

holds. Therefore, to prove (i), it suffices to show that, if $F$ is bounded from below, then $\lambda(F,x) < \infty$, and the premise in (2.2.42) is true.

The first statement follows immediately from the fact that, in the case of $\lambda(F,x) = \infty$, there exists $h \in E$ such that $D^2F(u)[h,h] = 0$ for all $u \in Q$, while $DF(x)[h] < 0$. This means that, on the intersection of $Q$ and the ray $\{x + th \mid t \geq 0\}$, the function $F$ linearly decreases. Since $X$ is closed and $Q$ is open, the above ray is contained in $Q$, and $F$ is not below bounded on $Q$. This contradicts the assumption.

Thus, in the case of a below bounded function $F$, we have $\lambda(F,x) < \infty$. Consider process (2.2.34). By virtue of (ii) and in view of (2.2.38), the first stage of this process does terminate, i.e., $x_j \in X$ and $\lambda(F,x_j) \leq \lambda' < 1$ for some $j$. Thus, the premise in (2.2.42) holds. □

When proving (2.2.38), we never used the assumption that $\lambda' > \lambda_*$ and, in fact, have established the following.

**Proposition 2.2.2** *Let $F \in S_a^+(Q,E)$ be such that $\lambda(F,x) < \infty$ for some (and then, in view of Proposition* 2.2.1, *for all) $x \in Q$. Let $u \in Q$. Then the point*

$$u^+ = u + \frac{e(F,u)}{1 + \lambda(F,u)}$$

*belongs to $Q$, and*

$$F(u^+) \le F(u) - a\{\lambda(F,u) - \ln(1+\lambda(F,u))\}.$$

**Comments.** Let $F$ be $a$-self-concordant and below bounded on $Q$ and let $x \in Q$. Assume that the set $\{y \in Q \mid F(y) \le F(x)\}$ is closed in $E$. By Theorem 2.2.3, $F$ attains its minimum over $Q$ and $\lambda_0 \equiv \lambda(F,x) < \infty$, while the above-described Newton iterations converge (in the objective) to the minimizer of $F$ over $Q$.

Moreover, $\lambda_i \to 0$, $i \to \infty$. Theorem 2.2.3 shows that the Newton process can be divided into three sequential stages, each corresponding to one of the following three conditions on the iteration number $i$:

$$i < i_*(1) \equiv \min\{i \mid \lambda_i \le \lambda'\};$$

$$i_*(1) \le i < i_*(2) \equiv \min\{i \mid \lambda_i < \lambda_*\};$$

$$i \ge i_*(2).$$

At the first stage, $F$ decreases at each iteration by a quantity that is not smaller than $a\{\lambda' - \ln(1+\lambda')\} \equiv a\lambda^*$; the number of iterations at this stage, $i_*(1)$, is not greater than

$$t(1) \equiv \rfloor (a\lambda^*)^{-1}(F(x) - \min_Q F) \lfloor.$$

At the second stage, the quantities $\lambda_i$ decrease, and the quantities $(1-\lambda_i)$ increase as a geometric progression with the ratio $\kappa = (5-\lambda')/4$; the number of iterations at this stage, $i_*(2) - i_*(1)$, is not greater than

$$t(2) \equiv 1 + \left\rfloor \ln^{-1}(\kappa) \ln\left(\frac{1-\lambda_*}{1-\lambda'}\right) \right\lfloor.$$

At the third stage, the quantities $\lambda_i$ decrease quadratically. It is important that the behaviour of $\lambda_i$ at the second stage depends on $\lambda'$ only, while at the third stage it does not depend on any parameter of the objects involved.

The inequality $\lambda(F,x) < 1$, which under the assumptions of Theorem 2.2.3 ensures the below boundedness of $F$, cannot be weakened. This is demonstrated by the example

$$F(x) = \ln\frac{1}{x} \in S_1^+((0,\infty), \mathbf{R}),$$

where $\lambda(F,x) \equiv 1$.

### 2.2.4 Behaviour of a strongly self-concordant function on its Lebesgue set

The results on self-concordant functions established so far demonstrate that such a function in its Dikin ellipsoid of a reasonable (less than 1) radius is fairly well approximated by its second-order Taylor expansion, which is the reason for nice behaviour of the Newton minimization method as applied to the function. Now, what happens outside the ellipsoid? In a "large" neighbourhood of a point, the behaviour of a self-concordant function $F$ may be bad—it can go to infinity at a finite distance. Nevertheless, it turns out—and this is important for our further goals—that this is, in a sense, the only bad thing that can happen. Namely, if $F$ is below bounded, then its behaviour on any Lebesgue set $\{x \mid F(x) < b\}$ resembles its behaviour on the ellipsoid. We are about to represent the related properties; to simplify notation, we restrict ourselves to strongly 1-self-concordant functions, as it always will be the case in our coming applications.

Let $Q$ be an open convex subset in $E$ and $F$ be a strongly 1-self-concordant *below bounded* function on $Q$. Since $F$ is strongly self-concordant, its Lebesgue sets

$$\mathcal{F}_b = \left\{x \in Q \mid F(x) \le b + \inf_Q F\right\}$$

are closed in $E$ for any $b > 0$; this observation and the below boundedness of $F$ imply that $F$ attains its minimum on $Q$ (Theorem 2.2.3(i)). Let $x(F)$ be a minimizer of $F$ over $Q$. The results on the behaviour of $F$ on $\mathcal{F}_b$ are as follows.

(i) *Boundedness of the Lebesgue set in the "central metric."* For any $x \in \mathcal{F}_b$, we have

$$\| x - x(F) \|_{x(F),F} \le \chi^{-1}(b) \le 3b + \tfrac{1}{4}, \tag{2.2.43}$$

where $\chi^{-1}$ is the function inverse to

$$\chi(t) = \int_0^t \int_0^s (1-r)_+^2 dr ds = \begin{cases} t^2/2 - t^3/3 + t^4/12, & 0 \le t \le 1, \\ \frac{1}{4} + \frac{1}{3}(t-1), & t > 1. \end{cases}$$

This is an immediate corollary of the following lower bound, which is useful in its own right.

**Proposition 2.2.3** *Let $f \in S_1(Q, E)$ and $x \in Q$. Then, for every $y \in Q$, we have*

$$f(y) \ge f(x) + Df(x)[y-x] + \chi(\| y - x \|_{x,f}). \tag{2.2.44}$$

**Proof.** Taking restriction of $f$ on the line passing through $x$ and $y$, we immediately reduce the situation to the following one: $Q$ is an interval on the axis, $x = 0$, $y > 0$, and $f$ is 1-self-concordant on $Q$. Assume first that $f''(0) = 0$; then (2.2.44) is nothing but convexity of $f$. Now assume that $f''(0) > 0$; then, without loss of generality, we may restrict ourselves to the case of $f''(0) = 1$ (this requires nothing but scaling of the argument), so that $\| z \|_{x,f}$ is simply $| z |$ and, in particular, $\| y \|_{x,f} = | y | = y$. Now, in view of Theorem 2.2.1(i), from $f''(0) = 1$ it follows that $f''(t) \ge (1-t)^2$ for all $t \in Q$, $0 \le t \le 1$, while, from convexity of $f$, it follows that $f''(t) \geq 0,\ t \in Q,\ t \geq 1$. Thus,

$$f(y) - f(0) - f'(0)y \geq \int_0^y \int_0^s (1-t)_+^2 dtds = \chi(y),$$

as claimed. $\square$

Note that the final estimate $3b + \frac{1}{4}$ in (2.2.43) in the case of small $b$ (less than $\chi(1) = \frac{1}{4}$) can be improved (in this range of values of $b$ the quantity $\chi^{-1}$ is given by another expression), but this is not the issue in which we are interested now.

(ii) *Boundedness of the Newton decrement.* For any $x \in \mathcal{F}_b$, we have

$$\lambda(F, x) \leq \psi^{-1}(b), \tag{2.2.45}$$

where $\psi^{-1}$ is the function inverse to

$$\psi(\lambda) = \lambda - \ln(1 + \lambda).$$

Indeed, since $F \in S_1^+(Q, E)$ is below bounded on $Q$, its Newton decrement is finite (Proposition 2.2.1; the case of $\lambda(F, \cdot) \equiv \infty$ is excluded, since $\lambda(F, x(F)) = 0$). Now, since $\lambda(F, x) < \infty$, Proposition 2.2.2 tells us that an appropriate step from $x$ keeps the point in $Q$ and reduces $F$ by at least $\psi(\lambda(F, x))$, so that $\psi(\lambda(F, x)) \leq b$, as claimed.

(iii) *Compatibility of Hessians.* For any $x \in \mathcal{F}_b$, we have

$$\begin{aligned} \alpha(b) D^2 F(x(F))[h, h] &\leq D^2 F(x)[h, h] \\ &\leq A(b) D^2 F(x(F))[h, h], \qquad h \in E, \end{aligned} \tag{2.2.46}$$

where

$$\alpha(b) = \left(2\chi^{-1}(b + \tfrac{1}{2} + \tfrac{1}{2}\psi^{-1}(b))\right)^{-2},$$

$$A(b) = 2(1 + \psi^{-1}(b))^{2 + 2b/\psi(0.06)}.$$

Indeed, let $W_{1/2} = \{y \in E \mid D^2 F(x)[y - x, y - x] \leq \frac{1}{4}\}$ be the Dikin ellipsoid of $F$ centered at $x$ of radius $\frac{1}{2}$. Since $F$ is strongly self-concordant on $Q$, this ellipsoid is contained in $Q$, and in this ellipsoid we have $D^2 F(y)[h, h] \leq 4 D^2 F(x)[h, h]$ (both claims follow from Theorem 2.2.1). We immediately conclude that

$$F(y) - F(x) - DF(x)[y - x] \leq \tfrac{1}{2}, \qquad y \in W_{1/2},$$

whence $F(y) \leq \min_Q F + b + \frac{1}{2} + \psi^{-1}(b)/2$, $y \in W_{1/2}$ (we have considered that $\lambda(F, x) \leq \psi^{-1}(b)$ and that $| DF(x)[y - x] | \leq \lambda(F, x) \| y - x \|_{x,F}$ by definition of $\lambda(F, x)$). Thus,

$$W_{1/2} \subset \mathcal{F}_{b + 1/2 + \psi^{-1}(b)/2},$$

whence, in view of (i), $\|\cdot\|_{x(F),F}$-diameter of $W_{1/2}$ does not exceed $\alpha^{-1/2}(b)$, and we come to the left inequality in (2.2.46).

To prove the right inequality, let us act as follows. Consider the Newton-type process

$$x_0 = x; \quad x_{i+1} = x_i + (1+\lambda(F,x_i))^{-1}e(F,x_i), \quad i = 1,2,\ldots.$$

In view of Proposition 2.2.2, this process is well defined, keeps the iterates in $\mathcal{F}_b$, and decreases $F$ at the $i$th step by at least $\psi(\lambda_i)$, $\lambda_i = \lambda(F,x_i)$. Let $i^*$ be the first $i$ such that $\lambda_i < 0.06$. Then

$$i^* \leq 1 + \frac{b}{\psi(0.06)}.$$

We have $\| x_i - x_{i+1} \|_{x_i,F} = \lambda_i/(1+\lambda_i)$ (since $\| e(F,u) \|_{u,F}$ clearly equals $\lambda(F,u)$). In view of Theorem 2.2.1,

$$\begin{aligned} D^2F(x_i)[h,h] &\leq (1- \| x_{i+1} - x_i \|_{x_i,F})^{-2} D^2F(x_{i+1})[h,h] \\ &= (1+\lambda_i)^2 DF(x_{i+1})[h,h], \qquad h \in E, \end{aligned}$$

and, in view of (2.2.46), we have $\lambda_i \leq \psi^{-1}(b)$ (recall that $x_i \in \mathcal{F}_b$), whence

$$D^2F(x_i)[h,h] \leq (1+\psi^{-1}(b))^2 D^2F(x_{i+1})[h,h], \qquad h \in E.$$

It follows that

$$D^2F(x)[h,h] \equiv D^2F(x_0)[h,h] \leq (1+\psi^{-1}(b))^{2i} D^2F(x_i)[h,h], \qquad h \in E,$$

whence

$$D^2F(x)[h,h] \leq (1+\psi^{-1}(b))^{2+2b/\psi(0.06)} D^2F(x_{i^*})[h,h], \qquad h \in E.$$

On the other hand, relation $\lambda(F,x_{i^*}) \leq 0.06$ in view of the first inequality in (2.2.17) ensures that

$$\| x_{i^*} - x(F) \|^2_{x(F),F} \leq \frac{\omega^2(0.06)}{(1-\omega(0.06))^2} \leq \frac{1}{81},$$

whence, again in view of Theorem 2.2.1,

$$D^2F(x_{i^*})[h,h] \leq 2D^2F(x(F))[h,h], \qquad h \in E.$$

Thus,

$$D^2F(x)[h,h] \leq 2(1+\psi^{-1}(b))^{2+2b/\psi(0.06)} D^2F(x(F))[h,h], \qquad h \in E,$$

as claimed in the right inequality in (2.2.46). $\square$

**Remark 2.2.1** *Unpleasant exponential type of the second inequality in* (2.2.46) *is inavoidable. Indeed, consider a below bounded* 1*-strongly self-concordant function*

$$F(t) = -\ln t + t, \quad t > 0.$$

*Here* $x(F) = 1$ *and* $\exp(-b) \in \mathcal{F}_b$ *for* $b > 0$, *so that the best upper bound for the ratio* $f''(t)/f''(x(F)) = t^{-2}$ *on* $\mathcal{F}_b$ *is at least* $\exp(2b)$ *and is therefore actually exponential in* $b$.

The last property we present is a "quantitive" version of the sufficient condition for below boundedness (see Theorem 2.2.3(i)).

(iv) Let $x \in Q$ be such that $\lambda(F, x) < 1$. Then

$$x \in \mathcal{F}_b \quad \text{with } b = 0.215 + 16\left(1 + \log_2 \frac{1}{1 - \lambda(F, x)}\right). \tag{2.2.47}$$

To establish (2.2.47), consider the Newton-type process (2.2.34)–(2.2.36) starting at $x$ and set $\lambda' = \lambda(F, x)$

$$x_0 = x; \quad x_{i+1} = x_i + \sigma_i e_i, \quad i = 1, 2, ...,$$

where

$$\sigma_i = \begin{cases} \frac{1-\lambda_i}{\lambda_i(3-\lambda_i)}, & \lambda' \geq \lambda_i \geq \lambda_* = 2 - 3^{1/2}, \\ 1, & \lambda_i \leq \lambda_*, \end{cases}$$

$$\lambda_i = \lambda(F, x_i), \quad e_i = e(F, x_i).$$

As we know from (2.2.39), (2.2.39), the process is well defined, keeps the iterates in Q, and decreases $\lambda_i$'s. Consider the stage of the process comprised of iterations with $\lambda_i > \lambda_*$, let the corresponding iteration numbers be $0, ..., i^*$. Let us bound from above the quantity by which the objective is reduced at the stage in question. Since $F$ is convex, we have

$$F(x_{i+1}) \geq F(x_i) + \sigma_i DF(x_i)[e(F, x_i)] = F(x_i) - \sigma_i \lambda_i^2 = F(x_i) - \frac{\lambda_i(1-\lambda_i)}{3-\lambda_i}$$
$$\geq F(x_i) - (1 - \lambda_i) \equiv F(x_i) - \rho_i,$$

whence

$$F(x) - F(x_{i^*+1}) \leq \sum_{i=0}^{i^*} \rho_i. \tag{2.2.48}$$

On the other hand, in view of (2.2.39), we have

$$\rho_{i+1} \geq \rho_i + \tfrac{1}{4}\rho_i^2, \qquad i = 0, ..., i^* - 1. \tag{2.2.49}$$

Let $I_1$ be the group comprised of all $i \leq i^*$ with $\rho_i \leq 2\rho_0$. If there are indices $i : i \leq i^*$ that are not covered by $I_1$, let $i(2)$ be the first of them and let $I_2$ be comprised of all indices $i : i(2) \leq i \leq i^*$ with $\rho_i \leq 2\rho_{i(2)}$. If there are indices in $1, ..., i^*$ that are not covered by $I_1$ and $I_2$, then let $i(3)$ be the first of them and let $I_3$ be comprised of all indices $i : i(3) \le i \le i^*$ such that $\rho_i \le 2\rho_{i(3)}$, and so on. After a finite number of steps, the set $\{1, ..., i^*\}$ will be partitioned into groups $I_j$, $j = 1, ..., k$, in such a way that

$$I_j = \{i(j), i(j)+1, ..., i(j+1)-1\}, \quad (i(1) = 0,\ i(k+1) = i^*+1),$$

$$\rho_i \le 2\rho_{i(j)}, \qquad i \in I_j,$$

and $\rho_{i(j+1)} > 2\rho_{i(j)}$. From the latter relation, it follows that

$$k \le \log_2 \frac{1-\lambda_*}{1-\lambda_0} + 1 \le \log_2 \frac{1}{1-\lambda_0} + 1. \tag{2.2.50}$$

Furthermore, from (2.2.49), we conclude that the number $l_j$ of indices in $I_j$ does not exceed $8\rho_{i(j)}^{-1}$, so that $\sum_{i \in I_j} \rho_i \le 16$ (indeed, we have at most $8\rho_{i(j)}^{-1}$ terms, at most $2\rho_{i(j)}$ each). Thus,

$$\begin{aligned} F(x_0) - F(x_{i^*+1}) &\le \sum_{i=1}^{i^*} \rho_i \le \sum_{j=1}^{k} \sum_{i \in I_j} \rho_i \\ &\le 16k \le 16\left(1 + \log_2 \frac{1}{1-\lambda(F,x)}\right). \end{aligned} \tag{2.2.51}$$

Now, $\lambda(F, x_{i^*+1}) \le \lambda_* = 0.2679... < \frac{1}{3}$, and therefore Theorem 2.2.2(iii) tells us that

$$F(x_{i^*+1}) - \min_Q F \le \frac{1}{2}\omega^2(\lambda_*)\frac{1+\omega(\lambda_*)}{1-\omega(\lambda_*)} < 0.215,$$

which, combined with (2.2.51), leads to (2.2.47). $\square$

## 2.3 Self-concordant barriers

### 2.3.1 Definition of a self-concordant barrier

Recall that a self-concordant function is a smooth convex function with the second-order differential being Lipschitz continuous with respect to the local Euclidean metric defined by this differential. The functions that, in addition, are themselves Lipschitz continuous with respect to the above local metric, are of special interest to us. The latter requirement means precisely that the Newton decrement of the function is bounded from above. We call these functions *self-concordant barriers*. As we shall see later, these barriers play the central role in the interior point machinery. In this section, we introduce the notion of a self-concordant barrier and study the basic properties of these barriers.

As already mentioned, a self-concordant barrier $F$ should satisfy the following relations:

$$| D^3F(x)[h,h,h] | \le \text{const}_1 \{D^2F(x)[h,h]\}^{3/2},$$

$$| DF(x)[h] | \leq \text{const}_2\{D^2F(x)[h,h]\}^{1/2}.$$

Thus, generally speaking, there are two Lipschitz constants, $\text{const}_1$ and $\text{const}_2$, responsible for the barrier. Note that a scaling $F \to cF$ updates these constants as follows:

$$\text{const}_1 \to c^{-1/2}\,\text{const}_1, \qquad \text{const}_2 \to c^{1/2}\,\text{const}_2.$$

In particular, we can enforce one of the constants to be equal to a prescribed value; from a technical viewpoint, it is convenient to provide $\text{const}_1 = 2$. We now give the precise definition.

**Definition 2.3.1** *Let $G$ be a closed convex domain in a finite-dimensional real vector space $E$ and let $\vartheta \geq 0$. A function $F : \operatorname{int} G \to \mathbf{R}$ is called a $\vartheta$-self-concordant barrier for $G$ (notation: $F \in \mathcal{B}(G,\vartheta)$) if $F$ is 1-strongly self-concordant on* int $G$ *and*

$$\vartheta(F) \equiv \sup\{\lambda^2(F,x) \mid x \in \operatorname{int} G\} \leq \vartheta.$$

*The value $\vartheta$ is called the parameter of the barrier $F$.*

**Example 0.** The function $F(x) \equiv$ const is a 0-self-concordant barrier for $\mathbf{R}^n$.

**Example 1.** The function $F(t) = -\ln t$ is a 1-self-concordant barrier for the nonnegative half-axis.

**Remark 2.3.1** *As we shall see later,* $F(x) =$ const *is the only possible self-concordant barrier for the whole $\mathbf{R}^n$ (see Corollary* 2.3.1, *below), and this is the only possible self-concordant barrier with the parameter less than* 1*: If a closed convex domain $G \subseteq \mathbf{R}^n$ differs from the whole space and $F$ is a $\vartheta$-self-concordant barrier for $G$, then $\vartheta \geq 1$ (see Corollary* 2.3.3, *below). Henceforth, we always (unless the opposite is explicitly indicated) deal with barriers for proper subsets of $\mathbf{R}^n$; so the reader should note that the value of the parameter in question is $\geq 1$.*

The following statement is almost evident (compare with Proposition 2.1.1).

**Proposition 2.3.1** (i) Stability under affine substitutions of argument. *Let $G$ be a closed convex domain in $E$, $F$ be a $\vartheta$-self-concordant barrier for $G$, and let $x = \mathcal{A}(y) \equiv Ay + b$ be an affine transformation from a space $E^+$ into $E$ such that $\mathcal{A}(E^+) \bigcap \operatorname{int} G \neq \emptyset$. Let*

$$G^+ = \mathcal{A}^{-1}(G), \quad F^+(y) = F(\mathcal{A}(y)) : \operatorname{int} G^+ \to \mathbf{R}.$$

*Then $F^+$ is a $\vartheta$-self-concordant barrier for $G^+$.*

(ii) Stability with respect to summation. *Let $G_i$, $1 \leq i \leq m$ be closed convex domains in $E$ and let $F_i$ be $\vartheta_i$-self-concordant barriers for $G_i$. Assume that the set $G^+ = \bigcap_{i=1}^{m} G_i$ has a nonempty interior. Then the function*

$$F^+ = \sum_{i=1}^{m} F_i : \operatorname{int} G^+ \to \mathbf{R}$$

*is a $(\sum_{i=1}^m \vartheta_i)$-self-concordant barrier for $G^+$.*

(iii) Stability with respect to direct products. *Let $F_1(x)$ be a $\vartheta_1$-self-concordant barrier for $G_1 \subset E_1$ and $F_2(y)$ be a $\vartheta_2$-self-concordant barrier for $G_2 \subset E_2$. Then $F_1(x)+F_2(y)$ is a $(\vartheta_1+\vartheta_2)$-self-concordant barrier for $G_1 \times G_2$.*

**Example 2** (the standard logarithmic barrier for a polyhedral set). As we have already seen, the function $(-\ln t)$ is a 1-self-concordant barrier for the ray $\{t \geq 0\}$. From Proposition 2.3.1(i), it follows that the function $-\ln(b - a^T x)$ is a 1-self-concordant barrier for the half-space $\{x \in \mathbf{R}^n \mid a^T x \leq b\}$. Now, from Proposition 2.3.1(ii), it follows that, if $G$ is a convex polytope $\{x \in \mathbf{R}^n \mid a_i^T x \leq b_i,\ 1 \leq i \leq m\}$ and the linear inequalities defining $G$ satisfy the Slater condition, then the function

$$F(x) = -\sum_{i=1}^{m} \ln(b_i - a_i^T x)$$

is an $m$-self-concordant barrier for $G$. This *standard logarithmic barrier* for a polytope underlies the pioneer interior-point methods for linear programming.

In what follows (Chapter 5), we shall present a number of self-concordant barriers for various standard domains arising in nonlinear convex programming.

### 2.3.2 Basic properties

Self-concordant barriers possess a number of properties that are heavily exploited by the interior-point machinery. The following statement summarizes most important of them. Denote by

$$\pi_y(x) \equiv \inf\{t \geq 0 \mid y + t^{-1}(x - y) \in G\}$$

the Minkowsky function of $G$ whose pole is at $y$. Recall that

$$W_r(x) = \{z \in E \mid D^2F(x)[z - x, z - x] \leq r^2\}$$

denotes the Dikin ellipsoid of $F$ centered at $x$ of the radius $r$.

**Proposition 2.3.2** *Let $G$ be a closed convex domain in $E$ and $F$ be a $\vartheta$-self-concordant barrier for $G$. Then*

(i) *Let $x, y \in \operatorname{int} G$. Then*

(i.1) *The unit Dikin ellipsoid $W_1(x)$ belongs to $G$;*

(i.2) *The following inequalities hold:*

$$DF(x)[x - y] \leq \frac{\vartheta \pi_y(x)}{1 - \pi_y(x)}; \tag{2.3.1}$$

$$DF(x)[y - x] \leq \vartheta; \tag{2.3.2}$$

$$F(x) \leq F(y) + \vartheta \ln \frac{1}{1 - \pi_y(x)}; \tag{2.3.3}$$

$$F(x) \geq F(y) + DF(y)[x-y] + \ln \frac{1}{1-\pi_y(x)} - \pi_y(x); \tag{2.3.4}$$

$$| DF(x)[h] | \leq \frac{\vartheta}{1-\pi_y(x)} \{D^2F(y)[h,h]\}^{1/2}, \qquad h \in E; \tag{2.3.5}$$

$$D^2F(x)[h,h] \leq \left( \frac{1+3\vartheta}{1-\pi_y(x)} \right)^2 D^2F(y)[h,h]. \tag{2.3.6}$$

*Moreover, if* $z \in \partial G$ *and* $\pi_z(x) \leq (\vartheta^{1/2}+1)^{-2}$, *then*

$$DF(x)[z-x] \geq 1 - (\vartheta^{1/2}+1)^2 \pi_z(x). \tag{2.3.7}$$

(ii) $G = G + E_F$ *(cf. Corollary* 2.1.1) *and* $F$ *is constant along the directions parallel to* $E_F$.

*Moreover,* $F$ *is bounded from below on* $\operatorname{int} G$ *if and only if the image of* $G$ *in the factor-space* $E/E_F$ *is bounded.*

*If* $F$ *is bounded from below, then it attains its minimum over* $\operatorname{int} G$ *at any point of the set* $X_F$ *that is a translation of* $E_F$ : $X_F = x(F) + E_F$, *where* $x(F)$ *is (any) minimizer of* $F$, *and the* $(1+3\vartheta)$*-enlargement of the Dikin ellipsoid* $W_1(x(F))$ *contains* $G$

$$\begin{aligned} &\{x \in E \mid D^2F(x(F))[x-x(F), x-x(F)] \leq 1\} \subseteq G \\ &\subseteq \{x \in E \mid D^2F(x(F))[x-x(F), x-x(F)] \leq (1+3\vartheta)^2\}. \end{aligned} \tag{2.3.8}$$

(iii) *Let* $x \in \operatorname{int} G, h \in E$ *and* $q_x(h) = \sup\{t \mid x \pm th \in G\}$. *Then*

$$\{D^2F(x)[h,h]\}^{-1/2} \leq q_x(h) \leq (1+3\vartheta)\{D^2F(x)[h,h]\}^{-1/2}, \tag{2.3.9}$$

*so that the diameter of the unit Dikin ellipsoid* $W_1(x)$ *in a direction* $h$ *is at least twice and at most* $2(1+3\vartheta)$ *times the smaller of the distances from* $x$ *to the boundary of* $G$ *in the directions* $\pm h$.

(iv) *The function*

$$f(x) = -\exp\left\{-\frac{1}{\vartheta}F(x)\right\}$$

*is convex on* $\operatorname{int} G$.

**Proof.** (i) Relation (i.1) follows from Theorem 2.1.1(ii). It remains to verify (i.2). Let

$$\Delta \equiv \{t \in \mathbf{R} \mid y + t(x-y) \in \operatorname{int} G\} = (-T', T), \quad T' > 0, \quad T > 1.$$

Let

$$\phi(t) = F(y + t(x-y)) : \Delta \to \mathbf{R}.$$

By Proposition 2.3.1(i), we have $\phi \in \mathcal{B}(\operatorname{cl}\Delta, \vartheta)$. It is possible that $\phi$ is a constant function; then relations (2.3.1)–(2.3.3) for $x$ and $y$ under consideration are evident. Moreover, in this case, $x - y \in E_F$; thus, either $x = y$, or the whole line $(x, y)$ is contained in $W_1(x)$ and therefore in $G$ ((i) is already proved!). In both cases, we have $\pi_y(x) = 0$, so that (2.3.4) holds.

Now assume that $\phi$ is not a constant function. Since $\phi$ is a barrier for $\mathrm{cl}\,\Delta$, we have $\phi''(t) > 0$ (Corollary 2.1.1) and $(\phi'(t))^2/\phi''(t) \leq \vartheta$, $t \in \Delta$, or $\phi''(t) \geq \vartheta^{-1}(\phi'(t))^2$. Let $\psi(t) = \phi'(t)$ and assume that $\psi(t_0) > 0$ for some $t_0 \in \Delta$. By the comparison theorem for differential equations as applied to

$$\eta(t) = \frac{\vartheta\psi(t_0)}{\vartheta - (t - t_0)\psi(t_0))}$$

(note that $\eta' \equiv \vartheta^{-1}\eta^2$, $\eta(t_0) = \psi(t_0)$), we have $\psi(t) \geq \eta(t)$ for each $t \geq t_0$ such that $\psi$ and $\eta$ are well defined at $t$. Thus, $T - t_0 \leq \vartheta/\psi(t_0)$. We obtain the following result:

*If $t_0 \in \Delta$ is such that $\psi(t_0) \equiv \phi'(t_0) > 0$, then $T - t_0 \leq \frac{\vartheta}{\psi(t_0)}$.*

Let us verify (2.3.1). This relation is evident in the case of $DF(x)[x-y] \leq 0$. Assume that $DF(x)[x - y] > 0$. Since $DF(x)[x - y] = \phi'(1) = \psi(1)$, we have $T - 1 \leq \vartheta/\psi(1)$, so that (2.3.1) holds. It is clear that (2.3.1) holds also for $y \in \partial G$.

Now let us prove (2.3.3). Relation (2.3.1) as applied to the barrier $\phi$ for $\mathrm{cl}\,\Delta$ leads to inequality $\phi'(t) \leq \vartheta/(T - t)$ for $0 \leq t < T < \infty$. Thus,

$$F(x) = \phi(1) \leq \phi(0) + \int_0^1 \frac{\vartheta}{T - t} dt = F(y) + \vartheta \ln \frac{T}{T - 1},$$

which implies (2.3.3). If $T = \infty$, or if $\pi_y(x) = 0$ (which is the same), we have, by applying (2.3.1) to $\phi$, $\phi'(t) \leq 0$, $t \in \Delta$, so that (2.3.3) is evident.

Now let us prove (2.3.2). Since $\phi$ is convex, we have $\phi'(t) \geq \phi'(0), 0 < t < T$, and, as it was already proved, in the case of $T < \infty$, we have $\phi'(t) \leq \vartheta/(T - t)$. Thus,

$$\phi'(0) \leq \frac{\vartheta}{T} = \vartheta\pi_y(x) \leq \vartheta,$$

or $DF(y)[x - y] \leq \vartheta$. If $T = \infty$, then, by (2.3.1), $\phi'(t) \leq 0$, $t > 0$, so, in this case again, $DF(y)[x - y] \leq \vartheta$. The inequality obtained differs from (2.3.2) only in notation.

Let us prove (2.3.4). The function $\phi$ is a barrier for $\mathrm{cl}\,\Delta$; thus, in the case of $T < \infty$, the relation $0 \leq t < T$ implies, by (i.1), the inequality $t + (\phi''(t))^{-1/2} \leq T$, or the inequality $\phi''(t) \geq (T - t)^{-2}$. Thus,

$$\begin{aligned} F(x) = \phi(1) &= \phi(0) + \phi'(0) + \int_0^1 \phi''(t)(1 - t)dt \\ &\geq F(y) + DF(y)[x - y] + \int_0^1 (1 - t)(T - t)^{-2} dt \\ &= F(y) + DF(y)[x - y] - \ln(1 - \pi_y(x)) - \pi_y(x), \end{aligned}$$

which is required in (2.3.4). In the case of $T = \infty$, we have $\pi_y(x) = 0$, and (2.3.4) is an immediate consequence of the convexity of $F$.

Let us prove (2.3.5). The situation in an evident way can be reduced to the case of $E_F = \{0\}$. Let us introduce a scalar product on $E$ as $\langle h, s\rangle = D^2F(y)[h,s]$ and let $\|\cdot\|$ be the corresponding norm. Let us identify the first- and second-order differentials with the gradients and the Hessians. We have $F''(y) = I$; thus, the open unit ball $V$ centered at $y$ is contained in $\operatorname{int} G$ ((i.1)). Let $y'$ be the point on the ray $[y, x)$ such that $x$ lies between $y$ and $y'$ and let $V'$ be the image of $V$ under the dilatation with the center at $y'$ and the coefficient $\alpha = \| x - y' \| / \| y - y' \|$. Then $V' \subset \operatorname{int} G$ is an open ball centered at $x$ with radius $\alpha$. Let $\alpha'$ be such that $0 < \alpha' < \alpha$, $h$ be the unit normalization of $F'(x)$ and let $z = x - \alpha' h$. Then $z \in \operatorname{int} G$ and $\pi_z(x) \le \frac{1}{2}$, which, by (2.3.1), gives $\langle F'(x), x - z\rangle \le \vartheta$, or $\| F'(x) \| \le \vartheta/\alpha'$. The quantity $\vartheta/\alpha'$ tends to $\vartheta/(1 - \pi_y(x))$ along an appropriately chosen sequence of values of $y'$ and $\alpha'$; thus, $\| F'(x) \| \le \vartheta/(1 - \pi_y(x))$. The latter inequality, by virtue of choice of the scalar product, is (2.3.5).

Let us prove (2.3.6). Since $W_1(y) \subset G$, the set

$$V = \{z \in E \mid D^2F(y)[z - x, z - x] < (1 - \pi_y(x))^2\},$$

which is a union of the images of $\operatorname{int} W_1(y)$ under dilatations with the centers in $\operatorname{int} G$, is contained in $\operatorname{int} G$. It suffices to prove (2.3.6) under the assumption that $D^2F(x)[h,h] = 1$; moreover, we can assume that $DF(x)[h] \ge 0$ (otherwise, we can replace $h$ by $-h$). Under the notation

$$x(t) = x + th, \quad \phi(t) = DF(x(t))[h], \quad 0 \le t < T \equiv \sup\{t \mid x(t) \in \operatorname{int} G\},$$

we have $T \ge 1$ ((i.1)) and $\phi'(t) \ge (1 - t)^2$, $0 \le t < 1$ (the latter relation holds by virtue of Theorem 2.2.1 and the equality $D^2F(x)[h,h] = 1$). These relations combined with the inequality $\phi(0) \ge 0$ lead to the inequality $\phi(t) \ge t(3 - 3t + t^2)/3$, $0 < t < 1$, or to the relation

$$DF(x(t))[x(t) - x] = t\phi(t) \ge t^2(1 - t + \tfrac{1}{3}t^2) \equiv \alpha(t).$$

By (2.3.1), the latter inequality means that $\pi_x(x(t)) \ge \alpha(t)/(\vartheta + \alpha(t))$. Taking limit as $t$ tends to 1, we obtain $\pi_x(x + h) \ge (1 + 3\vartheta)^{-1}$, so that the point $x + (1 + 3\vartheta)h$ does not belong to $\operatorname{int} G$ and hence does not belong to $V$. The latter statement means that

$$(1 + 3\vartheta)^2 D^2F(y)[h,h] \ge (1 - \pi_y(x))^2 = (1 - \pi_y(x))^2 D^2F(x)[h,h],$$

which is required in (2.3.6).

Let us prove (2.3.7). Let $\Delta = \{t \mid z + t(x - z) \in \operatorname{int} G\}$. Then

$$\Delta = (0, T), \quad T = \pi_z^{-1}(x) \ge (1 + \vartheta^{1/2})^2.$$

Let $\phi(t) = F(z + t(x - z))$. Then $\phi(t)$ is a barrier for $\mathrm{cl}\,\Delta$, so that $\phi''(t) \geq t^{-2}$, $t \in \Delta$ ((i.1)). When $1 < t < T$, we have, by (2.3.1),

$$t\phi'(t) \leq \frac{\vartheta t}{T - t}$$

(we set $1/(T-t) = 0$ when $T = \infty$). Thus,

$$\phi'(1) + \int\limits_1^t \frac{1}{\tau^2} d\tau \leq \frac{\vartheta}{T - t}, \qquad 1 \leq t < T$$

or

$$\phi'(1) \leq \frac{\vartheta}{T-t} - 1 + \frac{1}{t}, \qquad 1 \leq t < T.$$

If $T < \infty$, then we can take in the above inequality $t = (1 + \vartheta^{1/2})^{-1}\pi_z^{-1}(x)$ (under the premise of (2.3.7) this value of $t$ is $\geq 1$), which leads to

$$\phi'(1) \leq -1 + (1 + \vartheta^{1/2})^2 \pi_z(x).$$

In the case of $T = \infty$, the latter relation follows from the above inequality as $t \to \infty$. Thus,

$$DF(x)[z - x] = -\phi'(1) \geq 1 - (1 + \vartheta^{1/2})^2 \pi_z(x).$$

Thereby (i) is proved.

(ii) The fact that $F$ is constant along the intersections of $\mathrm{int}\, G$ with $E_F + x$, $x \in \mathrm{int}\, G$, is evident because $\lambda(F, \cdot) < \infty$; since $F$ tends to $\infty$ as the argument approaches a boundary point of $G$, the sets $x + E_F$, with $x \in \mathrm{int}\, G$ are contained in $\mathrm{int}\, G$. To prove the remaining statements of (ii), we can assume that $E_F = \{0\}$ (otherwise, we can consider the restriction of $F$ onto the intersection of $G$ and a subspace complement to $E_F$). Since in the case of $E_F = \{0\}$ the function $F$ is strongly convex, it is clear that in the case of bounded $G$ the function $F$ attains its minimum over $\mathrm{int}\, G$ at a unique point. Now assume that $G$ is unbounded; let us prove then that $F$ is unbounded from below. Indeed, $\mathrm{int}\, G$ contains a ray $L = [y, x)$, $y, x \in \mathrm{int}\, G$. By (2.3.1), we have $DF(z)[x - y] \leq 0, z \in L$; if $\pi_x(y) = 0$, then $\mathrm{int}\, G$ contains the ray $[x, y)$; hence $DF(z)[y - x] \leq 0, z \in L$, or $D^2F(z)[y - x] = 0$. The latter relation contradicts the assumption that $E_F = \{0\}$. Thus, $\pi_x(y) > 0$, and hence, for $z_t = y + t(x - y)$, we have $\lim_{t\to\infty} \pi_{z_t}(y) = 1$. By (2.3.4), we have

$$F(y) \geq F(z_t) + DF(z_t)[y - z_t] - \ln(1 - \pi_{z_t}(y)) - \pi_{z_t}(y) \geq F(z_t) - \ln(1 - \pi_{z_t}(y)) - 1$$

(since $DF(z_t)[y - z_t] \geq 0$ by the above arguments), which implies that

$$F(z_t) \leq F(y) + \ln(1 - \pi_{z_t}(y)) + 1 \to -\infty, \qquad t \to \infty.$$

It remains to verify (2.3.8). The left inclusion follows from (i.1). To prove the right inclusion, it suffices to show that, if $x(F)$ is the minimizer of $F$ over

$\operatorname{int} G$ and $h \in E$ is such that $D^2F(x(F))[h,h] = 1$, then the point $x(F) + \rho h$, $\rho = 1 + 3\vartheta$, does not belong to $\operatorname{int} G$. Let $x(t) = x(F) + th$ and let $\phi(t) = DF(x(t))[h]$; then $\phi(0) = 0$. Then, by choice of $h$ and Theorem 2.2.1, we have $\phi'(t) \geq (1-t)^2$, so that $\phi(t) \geq t(3 - 3t + t^2)/3$, $0 \leq t < 1$. At the same time, by (2.3.1), we have

$$t\phi(t) \leq \frac{\vartheta \pi_{x(F)}(x(t))}{1 - \pi_{x(F)}(x(t))}.$$

These results imply that

$$\pi_{x(F)}(x(1)) \geq \frac{1}{1 + 3\vartheta},$$

which leads to the right inclusion in (2.3.8). Thereby (ii) is proved.

Let us prove (iii). The left inequality in (iii) follows from (i.1). To prove the right inequality, it suffices to consider the case of $D^2F(x)[h,h] = 1, DF(x)[h] \geq 0$ and to verify that the point $x + (1 + 3\vartheta)h$ does not belong to $\operatorname{int} G$. The latter statement can be proved in the same way as used in the proof of the right inclusion in (2.3.8).

To prove (iv), it suffices to note that

$$D^2 f(x)[h,h] = \frac{1}{\vartheta^2} f(x)(\vartheta D^2F(x)[h,h] - \{DF(x)[h]\}^2) \geq 0. \qquad \square$$

**Corollary 2.3.1** *Let $F$ be a $\vartheta$-self-concordant barrier for a closed convex domain $G \subset E$ and let $h$ be an element of the recessive cone of $G$, i.e., $x + th \in G$ for all $x \in G$ and all $t > 0$. Then, for all $x \in \operatorname{int} G$, we have*

$$\{D^2F(x)[h,h]\}^{1/2} \leq -DF(x)[h].$$

*In particular, if $\operatorname{int} G$ contains a line $x_0 + \mathbf{R}h$, then $F$ is constant along this line.*

**Proof.** For $x \in \operatorname{int} G$, let $L = \{t \mid x + th \in G\}$. Then $L$ is a ray of the form $[-a, \infty)$, where $\infty \geq a > 0$, and the function $f(t) = F(x + th)$, $-a < t < \infty$, is a $\vartheta$-self-concordant barrier for $L$ (Proposition 2.3.1(i)). We should prove that

$$\{f''(0)\}^{1/2} \leq -f'(0).$$

It is possible that $f''(0) = 0$. In this case, $E_f = \mathbf{R}$, and therefore (see Proposition 2.3.2(ii)) $L = \mathbf{R}$ and $f$ is constant, so that the statement is evident. Now let $f''(0) > 0$, so that $E_f = \{0\}$. In view of Proposition 2.3.2(ii), $f$ does not attain its minimum over $\operatorname{int} L$. At the same time, (2.3.2) means that $f'(0)\tau \leq \vartheta$ for all $\tau$ from the interval $(-a, \infty)$, and therefore $f'(0) \leq 0$. To complete the proof, it suffices to note that $\lambda(f, 0) = (f'(0))^2/f''(0) \geq 1$, since otherwise $f$ would attain its minimum on $\operatorname{int} L$ (Theorem 2.2.3(i)). $\square$

### 2.3.3 Logarithmically homogeneous barriers

Now let us introduce a special class of self-concordant barriers for convex *cones*. This class is closely related to potential reduction interior-point methods (see Chapter 4).

**Definition 2.3.2** *Let $K$ be a closed convex and proper (i.e., $K \neq E$,) cone with a nonempty interior in a finite-dimensional real vector space $E$, let $\vartheta \geq 1$, and let $F : \operatorname{int} K \to \mathbf{R}$ be a function. $F$ is called $\vartheta$-*logarithmically homogeneous barrier for $K$ *(notation: $F \in B_\vartheta(K)$) if $F$ is a $C^2$-smooth convex function on $\operatorname{int} K$ such that $F(x_i) \to \infty$ for each sequence $\{x_i \in \operatorname{int} K\}$ that converges to a boundary point of $K$, and, for each $x \in \operatorname{int} K$ and each $t > 0$, we have*

$$F(tx) = F(x) - \vartheta \ln t. \tag{2.3.10}$$

*If $F \in B_\vartheta(K)$ and $F$ is 1-self-concordant on $\operatorname{int} K$, then $F$ is called $\vartheta$-*normal barrier for $K$ *(notation: $F \in NB_\vartheta(K)$).*

**Example 3.** The standard logarithmic barrier (see Example 2) for the nonnegative orthant $\mathbf{R}^n_+$

$$F(x) = -\sum_{i=1}^{n} \ln(x_i)$$

is an $n$-normal barrier for this cone.

The following statement is evident (compare with Proposition 2.3.1).

**Proposition 2.3.3** (i) Stability with respect to linear homogeneous substitutions of argument. *Let $K \subset E$, $F \in B_\vartheta(K)$ and let $A : H \to E$ be a linear homogeneous mapping from a finite-dimensional linear space $H$ into $E$ such that $A(H) \bigcap \operatorname{int} K \neq \emptyset$ and $A^{-1}(K) \neq H$. Then the set $A^{-1}(K)$ is a closed and proper convex cone with a nonempty interior in $H$, and the function $\Phi(u) = F(Au) : \operatorname{int} A^{-1}(K) \to \mathbf{R}$ belongs to $B_\vartheta(A^{-1}(K))$. If $F \in NB_\vartheta(K)$, then also $\Phi \in NB_\vartheta(A^{-1}(K))$.*

(ii) Stability with respect to summation. *Let $K_i \subset E$, $1 \leq i \leq m$ be such that the cone $K = \bigcap_{i=1}^m K_i$ has a nonempty interior and let $F_i \in B_{\vartheta_i}(K_i)$, $1 \leq i \leq m$. Then the function $F(x) = \sum_{i=1}^m F_i(x) : \operatorname{int} K \to \mathbf{R}$ belongs to $B_\vartheta(K)$, $\vartheta = \sum_{i=1}^m \vartheta_i$. If $F_i \in NB_{\vartheta_i}(K_i)$ for all $i$, then also $F \in NB_\vartheta(K)$.*

(iii) Stability with respect to direct products. *Let $K_1$ and $K_2$ be closed convex cones with nonempty interiors in the spaces $E_1$, $E_2$, respectively, $K_1 \neq E_1$, $K_2 \neq E_2$. Let $F_i \in B_{\vartheta_i}(K_i)$, $i = 1, 2$. Then $F_1(x) + F_2(y) \in B_{\vartheta_1+\vartheta_2}(K_1 \times K_2)$, and, if $F_i \in NB_{\vartheta_i}(K_i)$, $i = 1, 2$, then also $F_1(x) + F_2(y) \in NB_{\vartheta_1+\vartheta_2}(K_1 \times K_2)$.*

A logarithmically homogeneous barrier satisfies a number of useful identities.

**Proposition 2.3.4** *Let $F$ be a $\vartheta$-logarithmically homogeneous barrier for a cone $K$. Then for each $x \in \operatorname{int} K$, $h \in E$ and each $t > 0$,*

$$DF(tx) = t^{-1} DF(x), \tag{2.3.11}$$

$$DF(x)[h] = -D^2F(x)[x, h], \tag{2.3.12}$$

$$DF(x)[x] = -\vartheta, \tag{2.3.13}$$

$$D^2F(x)[x, x] = \vartheta, \tag{2.3.14}$$

$$\lambda^2(F, x) = \vartheta. \tag{2.3.15}$$

**Proof.** Differentiation of (2.3.10) in $x$ leads to (2.3.11), and differentiation of (2.3.10) in $t$ at the point $t = 1$ leads to (2.3.13). Differentiation of (2.3.13) in $x$ in the direction $h$ leads to (2.3.12); substituting $h = x$ into (2.3.12), we obtain (2.3.14).

It remains to prove (2.3.15). We have

$$\begin{aligned}\lambda^2(F, x) &= \sup\{2DF(x)[h] - D^2F(x)[h, h] \mid h \in E\} \\ &= \sup\{2D^2F(x)[-x, h] - D^2F(x)[h, h]\} = D^2F(x)[x, x] = \vartheta\end{aligned}$$

(we have used (2.3.12) and (2.3.14)). □

**Corollary 2.3.2** *A $\vartheta$-normal barrier for $K$ is a $\vartheta$-self-concordant barrier for $K$.*

This is an immediate consequence of the definitions and relation (2.3.15). □

**Proposition 2.3.5** *Let $K \subset E$ be a closed convex cone and let $F$ be a $\vartheta$-logarithmically homogeneous barrier for $K$. Let $E(K)$ be the maximal subspace of $E$ contained in $K$. Then $F$ is constant at each set of the form $x + E(K)$, $x \in \operatorname{int} K$. If $F$ is normal, then also $D^2F(x)[h, h] \neq 0$ for each $h \in E \backslash E(K)$.*

**Proof.** Let $h \in E(K)$ and $x \in \operatorname{int} K$. We wish to prove that $DF[x][h] = 0$. Otherwise, we can assume that $DF(x)[h] > 0$ (replacing, if necessary, $h$ by $-h$). Let $\phi(s) = DF(x + sh)[h]$. Since the line $\{x + sh\}$ belongs to $\operatorname{int} K$ (the definition of $E(K)$), the function $\phi$ is well defined and continuously differentiable on the half-axis $\{s \geq 0\}$. We have $\phi(0) > 0$ and

$$\phi'(s) = D^2F(x + sh)[h, h] \geq \lambda^{-2}(F, x + sh)\{DF(x + sh)[h]\}^2$$

(the definition of $\lambda(F, \cdot)$). By (2.3.15), we conclude from the above inequality that $\phi'(s) \geq \vartheta^{-1}\phi^2(s)$, whence $\phi(s) \geq \phi(0) > 0$, $s > 0$, and $(\phi^{-1}(s))' \leq -\vartheta^{-1}$, or $\phi^{-1}(s) \leq \phi^{-1}(0) - \vartheta^{-1}s$, $s \geq 0$; the resulting relation

$$0 < \phi^{-1}(s) \leq \phi^{-1}(0) - \vartheta^{-1}s, \qquad s > 0,$$

is impossible. Thus, $DF(x)[h] = 0$, $x \in \operatorname{int} K, h \in E(K)$, and $F$ is constant along $x + E(K)$.

It remains to prove that, if $F \in NB_\vartheta(K)$, then $D^2F(x)[h,h] > 0$ for each $x \in \operatorname{int} K$ and each $h \in E \backslash E(K)$. Since $F$ is a $\vartheta$-self-concordant barrier for $K$ (Corollary 2.3.1), the statement under consideration follows from Proposition 2.3.2(ii). □

### 2.3.4 Bounds on the parameter of a self-concordant barrier

It will be demonstrated (and this is the essence of the book) that it is possible to associate a self-concordant barrier for a closed convex domain with a number of interior-point methods minimizing linear and convex quadratic forms over the domain. The result is that the parameter of the barrier is responsible for the rate of convergence of these methods (the less the parameter, the better convergence). Therefore, we are interested in possible values of the parameter. As we demonstrate, for an $n$-dimensional domain, there always exists an $O(n)$-self-concordant barrier, and, generally speaking, the parameter cannot be less than $n$. The latter statement will be proved immediately; the former one, i.e., the existence, for an arbitrary closed convex domain in $\mathbf{R}^n$, an $O(n)$-self-concordant barrier, is the subject of §2.5.

The lower bound for the value of the parameter is given by the following proposition.

**Proposition 2.3.6** *Let $G$ be a convex polytope in $\mathbf{R}^n$ such that certain boundary point of $G$ belongs exactly to $k$ of $(n-1)$-dimensional facets of $G$, with the normals to these facets being linearly independent. Then the value of the parameter $\vartheta$ of any $\vartheta$-self-concordant barrier $F$ for $G$ cannot be less than $k$. In particular, the $n$-dimensional nonnegative orthant, simplex, and cube do not admit barriers with the parameter less than $n$.*

**Proof.** Under an appropriate choice of the coordinates, we can assume that the boundary point mentioned in the statement is 0 and that in a neighborhood of this point $G$ is defined by inequalities $x_i \le 0$, $1 \le i \le k$. Let

$$x(t) = -t\sum_{i=1}^{k} e_j,$$

$$x^{(i)}(t) = -t \sum_{1 \le j \le k, j \ne i} e_j, \quad 1 \le i \le k$$

($e_j$ are the unit vectors of the axes). Then, for all sufficiently small $t > 0$, the points $x(t)$ belong to $\operatorname{int} G$, the points $x^{(i)}(t)$ belong to $\partial G$ and $\pi_{x^{(i)}(t)}(x(t)) \to 0$ as $t \to 0$. The latter fact, by virtue of (2.3.7), implies that

$$\liminf_{t \to 0} DF(x(t))[x^{(i)}(t) - x(t)] \ge 1;$$

hence

$$k \leq \liminf_{t \to 0} \sum_{i=1}^{k} DF(x(t))[x^{(i)}(t) - x(t)] = \liminf_{t \to 0} DF(x(t))[0 - x(t)] \leq \vartheta$$

(the latter inequality holds by (2.3.2)). □

**Corollary 2.3.3** *If $G \subseteq \mathbf{R}^n$ is a closed convex domain that differs from $\mathbf{R}^n$ and $F$ is a $\vartheta$-self-concordant barrier for $G$, then $\vartheta \geq 1$.*

**Proof.** Since $G$ is a closed convex domain that differs from $\mathbf{R}^n$, its inverse image under certain affine embedding $\mathcal{A} : \mathbf{R} \to \mathbf{R}^n$ such that $\mathcal{A}(\mathbf{R})$ intersects $\operatorname{int} G$ is the nonnegative ray $\{t \geq 0\}$. From Proposition 2.3.1(i), it follows that $\phi(t) \equiv F(\mathcal{A}(t))$ is a $\vartheta$-self-concordant barrier for the latter ray, so that $\vartheta \geq 1$ in view of Proposition 2.3.6. □

## 2.4 Self-concordance and Legendre transformation

In what follows, we demonstrate that the Legendre transformation of a strongly self-concordant function is strongly self-concordant with the same parameter value. Similarly, the Legendre transformation of a normal barrier for a cone is a normal barrier with the same parameter value for the anti-dual cone. The latter fact will be heavily exploited in Chapter 4 in connection with potential reduction interior-point methods.

### 2.4.1 Legendre transformation of a strongly self-concordant function

Let $E$ be a finite-dimensional real vector space and $E^*$ be the space conjugate to $E$. The value of a form $s \in E^*$ at a vector $x \in E$ will be denoted by $\langle s, x \rangle$. If $F$ is a $C^2$-smooth convex function defined on a convex domain $G \subseteq E$, then $F'(x)$ and $F''(x)$ for $x \in G$ will denote the element of $E^*$ and the linear mapping from $E$ into $E^*$, respectively, uniquely defined by the relations

$$DF(x)[h] = \langle F'(x), h \rangle, \quad D^2F(x)[h,e] = \langle F''(x)h, e \rangle, \quad h, e \in E.$$

Note that, since $F$ is convex, $F''$ is symmetric positive-semidefinite,

$$\langle F''(x)h, e \rangle = \langle F''(x)e, h \rangle; \quad \langle F''(x)h, h \rangle \geq 0.$$

Note also that, if $F''$ nondegenerate, then

$$\lambda^2(F, x) = \left\langle F'(x), (F''(x))^{-1} F'(x) \right\rangle. \tag{2.4.1}$$

Let $a > 0$. A pair $(Q, F)$ comprised of a nonempty open convex set $Q \subset E$ and a nondegenerate (i.e., with a nondegenerate Hessian) strongly $a$-self-concordant on $Q$ function $F$ will be called an *$(a, E)$-pair.* The Legendre transformation of an $(a, E)$-pair $(Q, F)$ is defined as the pair $(Q, F)^* \equiv (Q^*, F^*)$, where

$$Q^* = \Phi(Q), \quad \Phi(x) = DF(x)[\cdot] : Q \to E^*,$$

$$F^*(\xi) = \sup\{\langle \xi, x\rangle - F(x) \mid x \in Q\}.$$

**Theorem 2.4.1** *Let $(Q, F)$ be an $(a, E)$-pair and $(Q^*, F^*)$ be its Legendre transformation. Then $(Q^*, E^*)$ is an $(a, E^*)$-pair and*

$$Q^* = \{\xi \in E^* \mid \text{ the function } F_\xi(x) = F(x) - \langle \xi, x\rangle \quad \text{is below bounded on } Q\}.$$

*Moreover, $(Q^*, F^*)^* = (Q, F)$ (as usually, the space $(E^*)^*$ is identified with $E$).*

**Proof.** Let

$$Q' = \{\xi \in E^* \mid \text{ the function } F_\xi(x) = F(x) - \langle \xi, x\rangle \text{ is below bounded on } Q\}.$$

We will prove that $Q' = Q^*$. Since the inclusion $Q^* \subset Q'$ is evident, we have to establish the inverse inclusion. Let $\xi \in Q'$, so that $F_\xi(x)$ is below bounded on $Q$. It is clear that a linear form is $a'$-self-concordant on $E$ for each $a' > 0$, so that $F_\xi \in S_a^+(Q, E)$ (see Proposition 2.1.1(ii)). Since this function is below bounded on $Q$, it attains its minimum at a point of this set (Theorem 2.2.3(i)), which means that $\xi \in \Phi(Q)$.

Since $D^2F(x)$ is nondegenerate for $x \in Q$, $Q^*$ is open, while $Q'$ is clearly a convex set. Thus, the set $Q^*$ is nonempty, open, and convex. In view of well-known properties of the Legendre transformation, it follows that, if $F$ is a convex function on $Q$ from $C^3$ such that $D^2F$ is nondegenerate, then $F^*$ has the same properties with respect to $Q^*$.

To show that $F^*$ is self-concordant with the parameter value $a$, we fix a point $x \in Q$ and note that, for all $e, h \in E$, we have

$$DF(x)[h] = \langle \Phi(x), h\rangle\,,$$

$$DF^*(\Phi(x))[s] = \langle s, x\rangle\,,$$

$$\langle \Phi'(x)h, e\rangle = D^2F(x)[h, e],$$

$$\langle (\Phi'(x)h)'h, h\rangle = D^3F(x)[h, h, h],$$

and

$$F^*(\Phi(x)) = \langle \Phi(x), x\rangle - F(x)$$

($(\Phi'(x)h)'h$ denotes the derivative of $\Phi'(x)h$ with respect to $x$ in the direction $h$; similar notation is used in the remaining formulas).

Taking the derivatives of these identities along the directions $h$ and $e$, we obtain

$$
\begin{aligned}
DF^*(\Phi(x))[\Phi'(x)h] &= \langle \Phi'(x)h, x\rangle = D^2F(x)[h,x],\\
D^2F^*(\Phi(x))[\Phi'(x)h, \Phi'(x)e] &= (DF^*(\Phi(x))[\Phi'(x)h])'e\\
&\quad -DF^*(\Phi(x))[(\Phi'(x)h)'e]\\
&= \langle (\Phi'(x)h)'e, x\rangle + \langle \Phi'(x)h, e\rangle\\
&\quad - \langle (\Phi'(x)h)'e, x\rangle\\
&= \langle \Phi'(x)h, e\rangle = D^2F(x)[h,e],\\
D^3F^*(\Phi(x))[\Phi'(x)h, \Phi'(x)h, \Phi'(x)h] &= (D^2F^*(\Phi(x))[\Phi'(x)h, \Phi'(x)h])'h\\
&\quad -2D^2F^*(\Phi(x))[\Phi'(x)h, (\Phi'(x)h)'h]\\
&= (\langle \Phi'(x)h, h\rangle)'h\\
&\quad -2D^2F^*(\Phi(x))[\Phi'(x)h, (\Phi'(x)h)'h]\\
&= D^3F(x)[h,h,h]\\
&\quad -2D^2F^*(\Phi(x))[\Phi'(x)h, (\Phi'(x)h)'h]\\
&= D^3F(x)[h,h,h]\\
&\quad -2D^2F(x)[\{\Phi'(x)\}^{-1}\{(\Phi'(x)h)'h\}, h]\\
&= D^3F(x)[h,h,h]\\
&\quad -2\langle \Phi'(x)\{\Phi'(x)\}^{-1}\{(\Phi'(x)h)'h\}, h\rangle\\
&= D^3F(x)[h,h,h]\\
&\quad -2\langle (\Phi'(x)h)'h, h\rangle = -D^3F(x)[h,h,h].
\end{aligned}
$$

Hence, for all $x \in Q$ and $h \in E$, we have

$$
\begin{aligned}
&\mid D^3F^*(\Phi(x))[\Phi'(x)h, \Phi'(x)h, \Phi'(x)h] \mid\\
&\qquad = \mid D^3F(x)[h,h,h] \mid \le 2a^{-1/2}(D^2F(x)[h,h])^{3/2}\\
&\qquad = 2a^{-1/2}(D^2F^*(\Phi(x))[\Phi'(x)h, \Phi'(x)h])^{3/2}.
\end{aligned}
\tag{2.4.2}
$$

As $(x,h)$ runs throughout $Q \times E$, $(\Phi(x), \Phi'(x)h)$ runs throughout $Q^* \times E^*$; therefore (2.4.1) means that $F^* \in S_a(Q^*, E^*)$. It remains to show that $F^*(\xi_i) \to \infty$ for each sequence $\{\xi_i \in \operatorname{int} Q^*\}$ converging to a point $\xi \in \partial Q^*$. To do this, assume that $\{F^*\{\xi_i)\}$ is bounded from above. Then the functions $F_{\xi_i}(x)$ are uniformly in $i$ bounded from below, and the same is true for $F_\xi$. The latter fact, by virtue of the equality $Q' = Q^*$, leads to $\xi \in Q^*$, which is a contradiction (since $Q^*$ is open and $\xi \in \partial Q^*$). Thus, $F^* \in S_a^+(Q^*, E^*)$, which, in view of the above results, means that $(Q^*, F^*)$ is an $(a, E^*)$-pair. The equality $(Q^*, F^*)^* = (Q, F)$ follows immediately from the above results and the standard properties of the Legendre transformation. $\square$

### 2.4.2 Legendre transformation of a self-concordant barrier

Now let us describe the Legendre transformations of self-concordant barriers. For this purpose, let us introduce some notation. For a closed convex domain $G \subset E$, let $\mathcal{R}(G)$ be the recessive cone of $G$

$$\mathcal{R}(G) = \{h \in E \mid x + th \in G, \ \ \forall x \in G \ \forall t \geq 0\}.$$

Also, let

$$\mathcal{R}^*(G) = \{s \in E^* \mid \langle s, h \rangle \leq 0 \ \forall h \in \mathcal{R}(G)\}$$

be the cone anti-dual to $\mathcal{R}(G)$. Note that, if $G$ does not contain any straight line, then $\mathcal{R}^*(G)$ is a closed convex cone with a nonempty interior in $E^*$.

**Theorem 2.4.2** *Let $G$ be a closed convex domain in $E$ that does not contain any straight line and let $F$ be a $\vartheta$-self-concordant barrier for $G$. Then the Legendre transformation $F^*$ of $F$ is defined precisely on the interior of $\mathcal{R}^*(G)$ and satisfies the following relations:*

(a) *$F^*$ is strongly 1-self-concordant on $\operatorname{int}\mathcal{R}^*(G)$ and $D^2F^*(s)$ is nondegenerate, $s \in \operatorname{int}\mathcal{R}^*(G)$;*

(b) $\sup\{D^2F^*(s)[s,s] \mid s \in \operatorname{int}\mathcal{R}^*(G)\} \leq \vartheta$;

(c) *The support function of $G$*

$$\mathcal{S}(s) = \sup\{\langle s, x \rangle \mid x \in G\}$$

*satisfies the inequality*

$$\mathcal{S}(s) - \frac{\vartheta}{t} \leq DF^*(ts)[s] \leq \mathcal{S}(s), \ \ s \in \operatorname{int}\mathcal{R}^*(G).$$

**Proof.** Since $G$ does not contain straight lines, $D^2F(x)$ is nondegenerate, $x \in \operatorname{int} G$ (see Proposition 2.3.2(ii)). Therefore $(\operatorname{int} G, F)$ is a $(1, E)$-pair, so that its Legendre transformation is an $(1, E^*)$-pair (Theorem 2.4.1). Let us prove that the domain $G^*$ of $F^*$ is precisely the interior of $\mathcal{R}^*(G)$. As we have seen in Theorem 2.4.1, $G^*$ is the image of $\operatorname{int} G$ under the transformation $x \to DF(x)$. If $x \in \operatorname{int} G$ and $h \in \mathcal{R}(G)$, then $x + th \in \operatorname{int} G$, $t \geq 0$, and therefore $DF(x)[h] \leq 0$ (see (2.3.2)). Thus, $G^*$ is contained in $\mathcal{R}^*(G)$, and since $G^*$ is open, $G^* \subseteq \operatorname{int}\mathcal{R}^*(G)$. Let us verify the inverse inclusion. Assume that $s \in \operatorname{int}\mathcal{R}^*(G)$. In view of Theorem 2.4.1, to prove that $s \in G^*$, it suffices to establish that the function $f(x) = \langle s, x \rangle - F(x)$ is above bounded on $\operatorname{int} G$. Let $| \cdot |$ be some Euclidean norm on $E$. Since $s \in \operatorname{int}\mathcal{R}^*(G)$, we have for some $\alpha > 0$:

$$\langle s, h \rangle \leq -\alpha \mid h \mid, \qquad h \in \mathcal{R}(G). \tag{2.4.3}$$

Assume that there exists a sequence $\{x_i \in \operatorname{int} G\}$ such that $f(x_i) \to \infty, \ i \to \infty$. In particular,

$$\mid x_i \mid \to \infty, \ i \to \infty \tag{2.4.4}$$

(recall that $F$ is a barrier). Let $y$ be some point of $\operatorname{int} G$. In view of (2.3.3), we have

$$f(x_i) \leq \langle s, x_i \rangle - F(y) - \vartheta \ln(1 - \pi_{x_i}(y)). \tag{2.4.5}$$

From (2.4.3) and (2.4.4) it follows that $\langle s, x_i \rangle \leq -(\alpha/2) \mid x_i \mid$ for all large enough $i$. At the same time, since $y$ is an interior point of $G$, we clearly have $\pi_{x_i}(y) \leq 1 - c \mid x_i \mid^{-1}$ for some positive $c$ and all large enough $i$. Thus, (2.4.5) implies that

$$f(x_i) \leq -\frac{\alpha}{2} \mid x_i \mid - F(y) + \vartheta \ln \frac{\mid x_i \mid}{c}$$

for all large enough $i$, so that $f(x_i) \to -\infty$, $i \to \infty$. The latter relation contradicts the origin of $\{x_i\}$.

Thus, $G^* = \operatorname{int} \mathcal{R}^*(G)$. This combined with Theorem 2.4.1 proves the statement concerning the domain of $F^*$ and (a).

Let us prove (b). By virtue of the standard properties of the Legendre transformation, we have $(F^*)''(F'(x)) = (F''(x))^{-1}$, $x \in \operatorname{int} G$, so that

$$D^2 F^*(F'(x))[F'(x), F'(x)] = \left\langle F'(x), (F''(x))^{-1} F'(x) \right\rangle = \lambda^2(F, x)$$

(see (2.4.1)). Since $F'(\operatorname{int} G) = \operatorname{int} \mathcal{R}^*(G)$, (b) follows.

It remains to prove (c). In view of the standard properties of the Legendre transformation, we have, for $s \in \operatorname{int} \mathcal{R}^*(G)$,

$$\begin{aligned} x(t) \equiv (F^*)'(ts) \in \operatorname{int} G; \qquad F'(x(t)) = ts; \\ DF^*(ts)[e] = \langle e, x(t) \rangle, \quad e \in E^*. \end{aligned} \tag{2.4.6}$$

From the first relation in (2.4.6) and the definition of the support function, it follows that $DF^*(ts)[s] \leq \mathcal{S}(s)$. On the other hand, (2.4.6) implies that, for $x \in G$, we have

$$\begin{aligned} \langle s, x \rangle &= \frac{1}{t}(\langle ts, x(t) \rangle + \langle ts, x - x(t) \rangle) \\ &= \frac{1}{t} DF^*(ts)[ts] + \frac{1}{t} \langle F'(x(t)), x - x(t) \rangle \leq DF^*(ts)[s] + \frac{\vartheta}{t} \end{aligned}$$

(the latter inequality follows from (2.3.2)). Thus,

$$\mathcal{S}(s) = \sup\{\langle s, x \rangle \mid x \in G\} \leq DF^*(ts)[s] + \frac{\vartheta}{t},$$

and (c) follows. □

The arguments used in the proof of (a) and (b) admit a straightforward inversion, which leads to the following result.

**Theorem 2.4.3** *Let $K$ be a closed convex cone with a nonempty interior in $E^*$ and let $F^*$ be strongly* 1*-self-concordant on* $\operatorname{int} K$ *function with nondegenerate second-order differential such that* $\vartheta \equiv \sup\{D^2 F^*(s)[s, s] \mid s \in \operatorname{int} K\} < \infty$.

*Then there exists a closed convex domain $G \subset E$ that does not contain any straight line and a $\vartheta$-self-concordant barrier $F$ for $G$ such that $K = \mathcal{R}^*(G)$ and $F^*$ is the Legendre transformation of $F$. $G$ is uniquely defined by its support function $\mathcal{S}(s) = \sup\{\langle s, x\rangle \mid x \in G\}$, which can be found from the relation*

$$\mathcal{S}(s) = \lim_{t\to\infty} DF^*(ts)[s], \qquad s \in \operatorname{int}\mathcal{R}^*(G).$$

### 2.4.3 Legendre transformation of a self-concordant logarithmically homogeneous barrier

This transformation proves to be a barrier of the same type, as it is demonstrated by the following theorem.

**Theorem 2.4.4** *Let $K$ be a closed convex pointed (i.e., $K \bigcap(-K) = \{0\}$) cone with a nonempty interior in $E$ and let $F$ be a $\vartheta$-logarithmically homogeneous self-concordant barrier for $K$. Then the Legendre transformation $F^*$ of $F$ is a $\vartheta$-logarithmically homogeneous self-concordant barrier for the cone $\mathcal{R}^*(K)$ anti-dual to $K$.*

**Proof.** Since $K$ is a pointed cone, it does not contain any straight line. Therefore, from Theorem 2.4.2 it follows that $F^* \in S_1^+(\operatorname{int}\mathcal{R}^*(K), E^*)$, and, to prove the theorem, it suffices to verify that $F^*$ is $\vartheta$-logarithmically homogeneous as follows:

$$F^*(ts) = F^*(s) - \vartheta \ln t, \quad s \in \operatorname{int}\mathcal{R}^*(K), \quad t > 0.$$

We have

$$\begin{aligned} F^*(ts) &= \sup\{\langle ts, x\rangle - F(x) \mid x \in \operatorname{int} K\} = \sup\{\langle s, x\rangle - F(t^{-1}x) \mid x \in \operatorname{int} K\} \\ &= \sup\{\langle s, x\rangle - F(x) + \vartheta \ln t^{-1} \mid x \in \operatorname{int} K\} = F^*(s) - \vartheta \ln t. \qquad \square \end{aligned}$$

From the definition of the Legendre transformation $F^*$ of a function $F$, it follows that $F(x) + F^*(x) - \langle s, x\rangle \geq 0$, and the infinum of the right-hand side of the latter inequality over $x$, as well as over $s$, equals zero. In the case when $F$ is a self-concordant logarithmically homogeneous barrier for a cone, we can establish another inequality of this type.

**Proposition 2.4.1** *Let $K$ be a closed convex pointed cone with a nonempty interior in $E$, let $F$ be a $\vartheta$-logarithmically homogeneous self-concordant barrier for $K$, let $K^- = \mathcal{R}^*(K)$ be the cone anti-dual to $K$, and let $F^*$ be the Legendre transformation of $K$. Let*

$$V(s, x) = F(x) + F^*(s) + \vartheta \ln(-\langle s, x\rangle) : \operatorname{int} K^- \times \operatorname{int} K \to \mathbf{R}.$$

*Then*

(2.4.7) $$V(\tau s, tx) = V(s, x), \qquad \tau, t > 0,$$

(2.4.8) $$V(s, x) \geq \vartheta \ln \vartheta - \vartheta,$$

*and the above inequality is an equality if and only if $s = tF'(x)$ for some $t > 0$.*

**Proof.** Relation (2.4.7) is evident, since $F^*$ and $F$ are $\vartheta$-logarithmically homogeneous (Theorem 2.4.3). Let us fix $s_0 \in \operatorname{int} K^-, x_0 \in \operatorname{int} K$, and let $p = -\langle s_0, x_0 \rangle$. The function $f(s) = F(x_0) + F^*(s) + \vartheta \ln p : \operatorname{int} K^- \to \mathbf{R}$ is strictly convex. Since $F^*$ is logarithmically homogeneous, we have $f'(ts) = t^{-1} f'(s)$ (see (2.3.11)), and since $F^*$ is the Legendre transformation of $F$, we have $(F^*)'(F'(x_0)) = x_0$. Combining these two observations, we conclude that the derivative of $f$ at the point $s(x_0) = (p/\vartheta)(F)'(x_0) \in \operatorname{int} K$ equals $(\vartheta/p)x_0$; see below:

$$Df(s(x_0))[e] = \frac{\vartheta}{p} \langle e, x_0 \rangle, \qquad e \in E^*. \tag{2.4.9}$$

In addition, in view of (2.3.13), we have

$$-\langle s(x_0), x_0 \rangle = p. \tag{2.4.10}$$

Equations (2.4.9) and (2.4.10) mean that $s(x_0)$ is the minimizer of $f(\cdot)$ over the set $\{s \in \operatorname{int} K^- \mid -\langle s, x_0 \rangle \geq p\}$ as follows:

$$s \in \operatorname{int} K^-, -\langle s, x_0 \rangle \geq p \;\Rightarrow\; f(s) \geq f(s(x_0)). \tag{2.4.11}$$

Moreover, since $f$ is strictly convex, the above minimizer is unique. Now note that $s_0$ satisfies the premise in (2.4.11) (the origin of $p$), so that

$$V(s_0, x_0) \geq f(s(x_0)), \tag{2.4.12}$$

and the equality in (2.4.12) implies that

$$s_0 = s(x_0) = -\frac{\langle s_0, x_0 \rangle}{\vartheta} F'(x_0). \tag{2.4.13}$$

We also have

$$\begin{aligned} f(s(x_0)) &= F(x_0) + F^*((p/\vartheta)F'(x_0)) + \vartheta \ln p \\ &= F(x_0) + F^*(F'(x_0)) - \vartheta \ln(p/\vartheta) + \vartheta \ln p \\ &= \vartheta \ln \vartheta + F(x_0) + F^*(F'(x_0)) \end{aligned}$$

(we have considered that $F^*$ is $\vartheta$-logarithmically homogeneous). In turn, since $F^*$ is the Legendre transformation of $F$, we have $F(x_0) + F^*(F'(x_0)) = \langle F'(x_0), x_0 \rangle$, and the latter quantity, in view of (2.3.13), equals $-\vartheta$. Thus, $f(s(x_0)) = \vartheta \ln \vartheta - \vartheta$. Now (2.4.12) can be rewritten as $V(s_0, x_0) \geq \vartheta \ln \vartheta - \vartheta$, and (2.4.8) is proved. From the above remarks it follows that the equality in (2.4.8) implies that $s = tF'(x)$ for some $t > 0$. Conversely, if $s = tF'(x)$, $t > 0$, then the above calculations immediately prove that (2.4.8) is an equality. □

## 2.5 Universal barrier

In this section, we demonstrate that an arbitrary $n$-dimensional closed convex domain admits an $O(n)$-self-concordant barrier. This barrier is given by certain universal construction and, for this reason, will be called *universal.* In fact, the universal barrier usually is too complicated to be used in interior-point algorithms, so that what follows should be regarded as nothing but an existence theorem. At the same time, this existence theorem is very important theoretically, since it means that the approach we are developing *in principle* can be applied to *any* convex problem.

**Theorem 2.5.1** *There exists an absolute constant $C$ such that each closed convex domain $G$ in $\mathbf{R}^n$ admits a $(Cn)$-self-concordant barrier. If $G$ does not contain any one-dimensional affine subspace of $E$, then we can take as the above barrier the function*

$$F(x) = O(1)\ln \mid G^*(x) \mid : \operatorname{int} G \to \mathbf{R}, \tag{2.5.1}$$

*where $O(1)$ is an appropriately chosen absolute constant;*

$$G^*(x) = \{\phi \in \mathbf{R}^n \mid \phi^T(y - x) \le 1 \ \forall y \in G\}$$

*is the polar of $G$ with respect to the point $x$; and $\mid \cdot \mid$ denotes the Lebesgue $n$-dimensional measure.*

**Remark 2.5.1** *Note that, if $G$ in the above theorem is a cone, then the barrier defined in the theorem is logarithmically homogeneous.*

**Proof.** Without loss of generality, we can restrict ourselves to the case when $G$ does not contain any straight line.

$1^0$. For $x \in \operatorname{int} G$, the polar $G^*(x)$ clearly is a bounded closed convex domain (note that $\operatorname{int} G^*(x) \neq \emptyset$ since $G$ does not contain lines and that the boundedness of $G^*(x)$ follows from the inclusion $x \in \operatorname{int} G$). Hence the function

$$f(x) = \mid G^*(x) \mid$$

is well defined and positive on $\operatorname{int} G$. If $x_i \in \operatorname{int} G$ and $x_i \to x \in \partial G$, then all of the sets $G^*(x_i)$ contain certain fixed open nonempty set (since the sequence $\{x_i\}$ is bounded), and, at the same time, these sets are not uniformly bounded (since $\lim x_i \in \partial G$). Since $G^*(\cdot)$ is a convex set, we have $f(x_i) \to \infty$. Thus, the function

$$\Phi(x) = \ln f(x)$$

is well defined on $\operatorname{int} G$ and tends to $\infty$ as the argument belonging to $\operatorname{int} G$ approaches a point from $\partial G$.

$2^0$. Let $S$ be the unit sphere in $\mathbf{R}^n$ and let

$$p(\phi) = \sup\{\phi^T y \mid y \in G\} : \mathbf{R}^n \to \mathbf{R}\bigcup\{+\infty\}$$

be the support function of $G$. For $x \in \operatorname{int} G$, we have

$$G^*(x) = \left\{ \tau\phi \mid \phi \in S, 0 \le \tau \le r_x(\phi) \equiv \frac{1}{p(\phi) - \phi^T x} \right\},$$

whence

$$f(x) = \frac{1}{n} \int (p(\phi) - \phi^T x)^{-n} dS(\phi)$$

($\int$ denotes the integral over the unit sphere in $\mathbf{R}^n$; $dS(\phi)$ means the element of the Lebesgue area; of course, $(+\infty)^{-n} = 0$). It is clear that $f$ (and hence $\Phi$) is $C^\infty$-smooth on $\operatorname{int} G$; moreover,

$$\begin{aligned} D^l f(x)[h, \ldots, h] &= \frac{(-1)^l (n+l-1)!}{n!} \int \{\phi^T h\}^l (p(\phi) - \phi^T x)^{-l-n} dS(\phi) \\ &= \frac{(-1)^l (n+l)!}{n!} \int\limits_{G^*(x)} \{y^T h\}^l dy \equiv \frac{(-1)^l (n+l)!}{n!} I_l(h) \end{aligned}$$

(we have used the description of $G^*(x)$ in terms of $r_x(\phi)$). A straightforward computation leads to the following expressions (where $x \in \operatorname{int} G$ and $I_0$ does not depend on $h$):

$$\begin{aligned} D\Phi(x)[h] &= -(n+1) I_1(h) I_0^{-1}, \\ D^2\Phi(x)[h,h] &= (n+1)(n+2) I_2(h) I_0^{-1} - (n+1)^2 \{I_1(h) I_0^{-1}\}^2, \\ D^3\Phi(x)[h,h,h] &= -(n+3)(n+2)(n+1) I_3(h) I_0^{-1} \\ &\quad +3(n+2)(n+1)^2 I_2(h) I_1(h) I_0^{-2} - 2(n+1)^3 I_1^3(h) I_0^{-3}. \end{aligned}$$

Let us fix $x \in \operatorname{int} G$ and $h \in \mathbf{R}^n$ such that $\| h \|_2 = 1$ and let

$$\Delta = \{t \in \mathbf{R} \mid \exists y \in G^*(x) : y^T h = t\},$$

$$\psi(t) = \left( \operatorname{mes}_{n-1} \{y \in G^*(x) \mid y^T h = t\} \right)^{1/(n-1)},$$

where $\operatorname{mes}_{n-1}$ denotes the $(n-1)$-dimensional Lebesgue measure. Then, clearly,

$$I_l(h) I_0^{-1} = \int\limits_{\Delta} t^l \eta^{n-1}(t) dt,$$

$$\eta(t) = \psi(t) \left( \int\limits_{\Delta} \psi^{n-1}(\tau) d\tau \right)^{-1/(n-1)};$$

hence $\eta(t) \ge 0$, $t \in \Delta$, and $\int \eta^{n-1}(t) dt = 1$. Note that *the function $\eta(t)$ is concave on the segment $\Delta$* (the latter is the Brunn–Minkowsky theorem; see [GR 60].

We see that the differentials of $\Phi$ can be expressed in terms of the quantities $I_l(h) I_0^{-1}$. These quantities can be thought of as the moments of certain random variable $\xi$ (taking values in $\Delta$ with the probability density $\eta^{n-1}(t)$). Let us express the initial moments in terms of central moments; i.e., let us denote ($\mathcal{E}$ means the averaging operator)

$$\mu = I_1(h)I_0^{-1} \equiv \mathcal{E}\xi;$$
$$\sigma^2 = I_2(h)I_0^{-1} - \mu^2 \equiv \mathcal{E}\{\xi - \mathcal{E}\xi\}^2;$$
$$\theta = \mathcal{E}\{\xi - \mathcal{E}\xi\}^3.$$

A straightforward computation leads to

$$D\Phi(x)[h] = -(n+1)\mu;$$
$$D^2\Phi(x)[h,h] = (n+2)(n+1)\sigma^2 + (n+1)\mu^2;$$
$$D^3\Phi(x)[h,h,h] = -(n+3)(n+2)(n+1)\theta - 6(n+2)(n+1)\sigma^2\mu - 2(n+1)\mu^3.$$

Thus, $\Phi$ is convex and

$$| D\Phi(x)[h] | \leq (n+1)^{1/2}\{D^2\Phi(x)[h,h]\}^{1/2}.$$

Considering the results of $1^0$, we see that, to prove the theorem, it suffices to verify that $\Phi$ is self-concordant with an appropriate absolute constant as the parameter value. In other words, it suffices to prove the inequality

$$| (n+3)(n+2)(n+1)\theta + 6(n+2)(n+1)\sigma^2\mu + 2(n+1)\mu^3 |$$
$$\leq O(1)\{(n+2)(n+1)\sigma^2 + (n+1)\mu^2\}^{3/2}.$$

The latter inequality, in turn, would follow from the inequality

$$| \theta | \leq O(1)\sigma^3.$$

$3^0$. Thus, we have reduced our problem to the problem as follows. We are given a segment $\delta = [-a,b] \subset \mathbf{R}$, $a,b > 0$ and a continuous concave nonnegative function $\psi(t)$ on $\delta$ such that

$$\int_{-a}^{b} t\psi^{n-1}(t)dt = 0, \tag{2.5.2}$$

$$\int_{-a}^{b} \psi^{n-1}(t)dt = 1. \tag{2.5.3}$$

Let

$$\theta = \int_{-a}^{b} t^3\psi^{n-1}(t)dt \quad \text{and} \quad \sigma = \left\{\int_{-a}^{b} t^2\psi^{n-1}(t)dt\right\}^{1/2}.$$

We should prove that, under an appropriate choice of an absolute constant $O(1)$, we have $\theta \leq O(1)\sigma^3$. This inequality is evident in the case of $n = 1$; so let us assume that $n > 1$.

First, let

$$\lambda = \frac{1}{\sigma},$$

$$a = \lambda\tilde{a}, \quad b = \lambda\tilde{b},$$

$$\psi(t) = \lambda^{1/(n-1)}\tilde{\psi}\left(\frac{t}{\lambda}\right),$$

$$\tilde{\theta} = \int_{-\tilde{a}}^{\tilde{b}} t^3\tilde{\psi}^{n-1}(t)dt = \lambda^3\theta = \frac{\theta}{\sigma^3},$$

$$\tilde{\sigma} = \left\{\int_{-\tilde{a}}^{\tilde{b}} t^2\tilde{\psi}^{n-1}(t)dt\right\}^{1/2} = \lambda\sigma = 1.$$

Note that

$$\int_{-\tilde{a}}^{\tilde{b}} \tilde{\psi}^{n-1}(t)dt = 1, \qquad \int_{-\tilde{a}}^{\tilde{b}} t\tilde{\psi}^{n-1}(t)dt = 0.$$

Thus, our problem can be reduced to the case when $a, b, \psi$ satisfy (2.5.2), (2.5.3) and the condition

$$\int_{-a}^{b} t^2\psi^{n-1}(t)dt = 1; \tag{2.5.4}$$

under these assumptions, we wish to prove that

$$|\,\theta\,| \leq O(1). \tag{2.5.5}$$

It is convenient to introduce the body

$$G^* = \{(t,u) \in \mathbf{R} \times \mathbf{R}^n \mid t \in \delta,\ \|\,u\,\|_2 \leq \psi(t)\}.$$

Taking the volume of the unit ball in $\mathbf{R}^{n-1}$ as the unit of volume in this space, we have the facts that $G^*$ is a convex compact body of unit volume (see (2.5.3)), and the center of gravity of this body is at the origin (see (2.5.2)).

Without loss of generality, we can assume that

$$|\,\theta\,| \leq \int_{0}^{b} t^3\psi^{n-1}(t)dt \equiv \theta^*.$$

Thus, it suffices to evaluate from above the quantity $\theta^*$.

Each hyperplane passing through the gravity center of a convex compact body of unit volume divides this body into parts with the volumes of the parts being more than $1/e$ [GR 60]. In particular,

$$1 - \frac{1}{e} \le \int_0^b \psi^{n-1}(t)dt \equiv V \le \frac{1}{e}. \tag{2.5.6}$$

Let $\tau$ be such that

$$V' \equiv \int_\tau^b \psi^{n-1}(t)dt = \left(\frac{n-1}{n}\right)^n V. \tag{2.5.7}$$

In view of (2.5.4), (2.5.6), and (2.5.7), we have ("the Tchebyshev inequality")

$$\tau \le O(1) \tag{2.5.8}$$

(henceforth, all the constant factors in $O(\cdot)$ are absolute constants). Therefore (2.5.4) implies that

$$\int_0^\tau t^3\psi^{n-1}(t)dt \le O(1). \tag{2.5.9}$$

Now let us introduce a *linear* function $\phi(t)$ and a positive real $h$ satisfying the relations

$$\phi(h) = 0; \quad \phi(\tau) = \psi(\tau); \quad \int_\tau^h \phi^{n-1}(t)dt = V'; \tag{2.5.10}$$

i.e., let us replace the part of $G^*$ situated to the right of the hyperplane $t = \tau$ by the cone of the same volume and of the same intersection with this hyperplane. It is clear that the graph of $\phi$ is a secant of the graph of $\psi$, and the $t$-coordinates of the intersection points of these graphs are $\tau$ and $\tau' > \tau$. In addition, $h \ge b$. Note that

$$\int_\tau^b t^3\psi^{n-1}(t)dt - \int_\tau^h t^3\phi^{n-1}(t)dt = \int_\tau^h t^3\gamma(t)dt,$$

where $\gamma$ is a function with the zero value of the integral over the segment $[\tau, h]$, such that $\gamma$ is nonnegative on $[\tau, s]$ and nonpositive on $[s, h]$ for an appropriate $s$ (we have considered the convexity of $\psi$ and the linearity of $\phi$). In view of these properties, we have

$$\int_\tau^h t^3\gamma(t)dt \le 0.$$

Thus,

$$\theta^* \le \int_\tau^h t^3\phi^{n-1}(t)dt + \int_0^\tau t^3\psi^{n-1}(t)dt = \theta^{**} + O(1),$$

$$\theta^{**} \equiv \int_{\tau}^{h} t^3 \phi^{n-1}(t)dt$$

(see (2.5.9)). Thus, it suffices to prove that $\theta^{**} \leq O(1)$.

Let us verify that $h \leq n\tau$. Indeed, consider the cone

$$K = \{(t,u) \mid 0 \leq t \leq h, \ \| u \|_2 \leq \phi(t)\}.$$

The part of this cone situated between the hyperplanes $t = 0$ and $t = \tau$ contains similarly defined part of $G^*$, and the part $K'$ of the cone $K$, which is situated to the right of the hyperplane $t = \tau$, has the same volume $V'$ as the corresponding part of $G^*$. Therefore

$$V \leq \mid K \mid = \left(\frac{h}{h-\tau}\right)^n \mid K' \mid = \left(\frac{h}{h-\tau}\right)^n V' = \left(\frac{h}{h-\tau}\right)^n \left(\frac{n-1}{n}\right)^n V$$

(the latter equality holds by virtue of the definition of $V'$), which implies that $h \leq n\tau$. Thus, we have

$$h = \eta n \tau, \quad \eta \leq 1, \quad \tau \leq O(1) \tag{2.5.11}$$

and

$$\theta^{**} = S \int_{\tau}^{h} t^3 \left(\frac{h-t}{h-\tau}\right)^{n-1} dt, \qquad V' = \frac{1}{n} S(h-\tau), \tag{2.5.12}$$

where $S = \psi^{n-1}(\tau)$. We have

$$\begin{aligned}
\theta^{**} &= S\tau^4 \int_{1}^{n\eta} l^3 \left(\frac{\eta n - l}{\eta n - 1}\right)^{n-1} dl \\
&= \frac{S\tau^4}{(n\eta-1)^{n-1}} \int_{0}^{n\eta-1} \{n^3\eta^3 - 3n^2\eta^2 s + 3n\eta s^2 - s^3\} s^{n-1} ds \\
&= S\tau^4(n\eta-1)\left(n^2\eta^3 - \frac{3n^2\eta^2(n\eta-1)}{n+1} + \frac{3n\eta(n\eta-1)^2}{n+2} - \frac{(n\eta-1)^3}{n+3}\right) \\
&= V'\tau^3 n \left(\left[n^2\eta^3 - \frac{(n\eta-1)^3}{n+3}\right]_1 - \left[\frac{3n^2\eta^2(n\eta-1)}{n+1} - \frac{3n\eta(n\eta-1)^2}{n+2}\right]_2\right).
\end{aligned}$$

Since $V' \leq 1$ and $\tau \leq O(1)$, it suffices to verify that the expression denoted by $[\cdot]_1 - [\cdot]_2$ does not exceed $O(n^{-1})$ (this will lead to the desired estimate $\theta^{**} \leq O(1)$). A straightforward computation, which takes into account the relation $0 \leq \eta \leq 1$, leads to

$$[\cdot]_1 = \frac{3n^3\eta^2(\eta+1) - 3\eta n^2}{n^2+3n} + O\left(\frac{1}{n}\right),$$

$$[\cdot]_2 = \frac{3n^3\eta^2(\eta+1) - 3n^2\eta}{n^2+3n+2} + O\left(\frac{1}{n}\right);$$

hence $[\cdot]_1 - [\cdot]_2$ is of the desired order. $\square$

*This page intentionally left blank*

# Chapter 3
# Path-following interior-point methods

In this chapter, we develop a general approach to the design of polynomial-time interior-point methods. In the previous chapter, we demonstrated that, to effectively minimize a (strongly) self-concordant function, we can use the Newton minimization method. In what follows, we explain how to exploit this observation to solve a general type convex programming problem. The general strategy is as follows.

**A.** Consider a minimization problem in the form

$$(f): \quad f(x) \to \min \mid x \in G \subset E,$$

where $f$ is a convex function and $G$ is a closed convex subset of $E$. It is well known that, without loss of generality, we can assume $f$ to be *linear*. Of course, we can think of $G$ as being full-dimensional (otherwise, we could replace $E$ by the affine span of $G$). Thus, from the theoretical viewpoint, we can restrict ourselves to *standard* problems only, i.e., to problems $(f)$ generated by a linear $f(\cdot)$ and a closed convex domain $G$.

**B.** The linear objective involved into a standard problem $(f)$ is, of course, self-concordant. Nevertheless, we cannot minimize it straightforwardly with the aid of the Newton method, since the objective is not *strongly* self-concordant on $G$, while our machinery works only for strongly self-concordant functions. To overcome this difficulty, we can use the (quite traditional) penalty approach; i.e., we can regularize the objective to obtain a strongly self-concordant function that converges in a proper sense to the objective when the parameter responsible for the regularization is varied in an appropriate manner. In other words, we can associate with $(f)$ a parameterized family of problems

$$(F_t): \quad F_t(x) \to \min \mid x \in G_t \subset E,$$

such that the trajectory $x^*(t)$ of minimizers of $F_t$ converges to the set of solutions of $(f)$ as the penalty parameter $t$ tends, say, to $+\infty$. Then we can try to approximate the trajectory $x^*(t)$ in an appropriate manner along some sequence of parameter values tending to $+\infty$, which gives approximate solutions. The approximation of the trajectory usually is formed according to the *path-following scheme*: Given, for the current value $t$ of the penalty parameter, a "close" to $x^*(t)$ approximation $x(t)$, we replace $t$ by a larger parameter value $t'$ and regard $x(t)$ as an approximation of a new minimizer $x^*(t')$. Then we improve this latter approximation by some numerical optimization method to restore the closeness of the improved approximation $x(t')$ to the new point $x^*(t')$ of the trajectory.

In this chapter, we study the above scheme under the assumption that the family under consideration consists of self-concordant functions and that current approximations are improved by the Newton method. In §3.1 we indicate conditions on the family that ensure the polynomiality of the resulting method. The remaining sections are devoted to four path-following methods associated with concrete families satisfying the conditions described in §3.1. Namely, §3.2 contains the results on the barrier-generated family and barrier-generated path-following method; the path-following method of centers and the underlying family are studied in §3.3. Sections 3.4 and 3.5 present one additional family and two additional methods (the *dual* and the *primale parallel trajectories* methods).

## 3.1 Self-concordant families

In this section, we describe rather general conditions on a family that allow us to relate the rate of varying the penalty parameter and the amount of Newton steps sufficient to maintain closeness to the trajectory $x^*(\cdot)$. The conditions might look rather technical; nevertheless, they can be easily implemented in all our applications and, in these particular cases, save a lot of highly technical considerations. We therefore believe that the patience of the reader following the analysis of a general self-concordant family will be more than compensated by reducing his total effort.

### 3.1.1 Definition and basic properties

**Definition 3.1.1** *Let $E$ be a finite-dimensional real vector space,*

$$\mathcal{F} = \{Q_t, F_t, E\}_{t\in\Delta}$$

*be a family of functions defined on nonempty open convex subsets $Q_t \subset E$, $\Delta$ be an open nonempty interval on the real axis, and $Q_* = \{(t,x) \in E_* \equiv \mathbf{R}\times\mathbf{R}^n \mid t \in \Delta,\ x \in Q_t\}$. Let $\alpha(t)$, $\gamma(t)$, $\mu(t)$, $\xi(t)$, $\eta(t)$ be continuous positive scalar functions on $\Delta$, where $\alpha$, $\gamma$, $\mu$ are assumed continuously differentiable, and let $\kappa \in (0, \lambda_*)$.*[1] *The family $\mathcal{F}$ is called self-concordant with the parameters $\alpha$, $\gamma$, $\mu$, $\xi$, $\eta$, $\kappa$ (notation: $\mathcal{F} \in \Sigma_\Delta(\alpha,\gamma,\mu,\xi,\eta,\kappa)$), if the following conditions hold:*

($\Sigma$.1) Convexity and regularity. *$Q_*$ is an open subset of $E_*$; $F_t(x)$ is convex in $x$, continuous in $(t,x) \in Q_*$ and has three derivatives in $x$, $D^iF_t(x)$, continuous in $(t,x) \in Q_*$ for $i = 1,2,3$ and continuously differentiable in t for $i = 1,2$;*

($\Sigma$.2) Self-concordance of members. *For any $t \in \Delta$, the function $F_t : Q_t \to \mathbf{R}$ is self-concordant with the parameter value $\alpha(t)$;*

[1] Recall that $\lambda_* = 2 - \sqrt{3}$.

($\Sigma$.3) Compatibility of neighbours. *The set* $X_*(\kappa) \equiv \{(t,x) \in Q_* \mid \lambda(F_t,x) \leq \kappa\}$ *is closed in* $E_\Delta \equiv \Delta \times E$, *and there exists a neighbourhood (in* $\Delta \times E$*) of this set* $X^+(\kappa)$ *such that, for each* $(t,x) \in X^+(\kappa)$ *and* $h \in E$, *the following inequalities hold:*

$$(3.1.1) \quad \mid \{DF_t(x)[h]\}'_t - \{\ln \mu(t)\}'_t DF_t(x)[h] \mid \leq \xi(t)\alpha^{1/2}(t)\{D^2F_t(x)[h,h]\}^{1/2},$$

$$(3.1.2) \quad \mid \{D^2F_t(x)[h,h]\}'_t - \{\ln \gamma(t)\}'_t D^2F_t(x)[h,h] \mid \leq 2\eta(t)D^2F_t(x)[h,h]$$

*(henceforth,* $D$ *and* $\{\cdot\}'_t$ *mean the derivatives in* $x$ *and* $t$*, respectively).*

*The family* $\mathcal{F}$ *is called strongly self-concordant with the parameters* $\alpha, \gamma, \mu, \xi, \eta$ *(notation:* $\mathcal{F} \in \Sigma^+_\Delta(\alpha,\gamma,\mu,\xi,\eta)$*) if it satisfies conditions* ($\Sigma$.1), ($\Sigma$.2), *and the following:*

($\Sigma^+$.3) *Inequalities* (3.1.1) *and* (3.1.2) *hold for each* $(t,x) \in Q_*$ *and* $h \in E$, *and the set* $X^*(a) = \{(t,x) \in Q_* \mid F_t(x) \leq a\}$ *is closed in* $\Delta \times E$ *for each* $a \in \mathbf{R}$.

Note that the essence of the definition is in inequalities (3.1.1) and (3.1.2); these inequalities, roughly speaking, state that the first- and the second-order derivative of $F_t$ in $x$, taken along a direction $h$ of the unit local length $\{D^2F_t(x)[h,h]\}^{1/2}$, vary with $t$ at a rate that is "almost proportional" to the derivative itself.

Let us outline the basic properties of self-concordant families.

**Proposition 3.1.1** *Let* $\mathcal{F} = \{Q_t, F_t, E\}_{t\in\Delta}$ *be a family. Then the following assertions are true:*

(i) *Strong self-concordance of a family implies its self-concordance,*

$$\mathcal{F} \in \Sigma^+_\Delta(\alpha,\gamma,\mu,\xi,\eta) \Rightarrow \mathcal{F} \in \Sigma_\Delta(\alpha,\gamma,\mu,\xi,\eta,\kappa) \quad \forall \kappa \in (0,\lambda_*);$$

(ii) Stability with respect to affine substitutions of "spatial" argument. *Let* $x = \mathcal{A}(y) = Ay + b$ *be an affine transformation of a finite-dimensional real vector space* $E^+$ *into* $E$, $Q^+_t = \{y \in E^+ \mid Ay + b \in Q_t\}$ *and* $F^+_t(y) = F_t(Ay+b) : Q^+_t \to \mathbf{R}$. *Then the following implications hold:*

(ii.1)

$$\mathcal{F} \in \Sigma_\Delta(\alpha,\gamma,\mu,\xi,\eta,\kappa), \;\; \mathcal{A}(E^+) = E$$
$$\Rightarrow \mathcal{F}^+ \equiv \{Q^+_t, F^+_t, E^+\}_{t\in\Delta} \in \Sigma_\Delta(\alpha,\gamma,\mu,\xi,\eta,\kappa)$$

(ii.2)

$$\mathcal{F} \in \Sigma^+_\Delta(\alpha,\gamma,\mu,\xi,\eta), \qquad (Q^+_t \neq \emptyset \; \forall t \in \Delta)$$
$$\Rightarrow \mathcal{F}^+ = \{Q^+_t, F^+_t, E^+\}_{t\in\Delta} \in \Sigma^+_\Delta(\alpha,\gamma,\mu,\xi,\eta)$$

(iii) Stability with respect to summation. *Let*

$$\mathcal{F} = \{Q_t, F_t, E\}_{t\in\Delta} \in \Sigma^+_\Delta(\alpha,\gamma,\mu,\xi,\eta),$$

$$\mathcal{F}^* = \{Q_t^*, F_t^*, E\}_{t\in\Delta} \in \Sigma_\Delta^+(\alpha^*, \gamma, \mu, \xi^*, \eta^*),$$
$$p,\ p^* > 0.$$

*Let* $Q_t^+ \equiv Q_t \bigcap Q_t^* \neq \emptyset$ *for each* $t \in \Delta$. *Let* $\alpha^+$ *be a positive continuously differentiable function on* $\Delta$ *such that*

$$\alpha^+(t) \leq \min\{p\alpha(t), p^*\alpha^*(t)\}, \qquad t \in \Delta$$

*and let*

$$\eta^+(t) = \max\{\eta(t), \eta^*(t)\},$$
$$\xi^+(t) = \frac{2}{(\alpha^+(t))^{1/2}} \max\{(p\alpha(t))^{1/2}\xi(t), (p^*\alpha^*(t))^{1/2}\xi^*(t)\}.$$

*Then the family*

$$\mathcal{F}^+ = \{Q_t^+, F_t^+ \equiv pF_t + p^*F_t^*, E\}_{t\in\Delta}$$

*belongs to* $\Sigma_\Delta^+(\alpha^+, \gamma, \mu, \xi^+, \eta^+)$.

**Proof.** (i) It suffices to prove that, if $\kappa \in (0, \lambda_*)$, then $X_*(\kappa)$ is closed in $E_\Delta$. For $t \in \Delta$, the function $F_t(\cdot)$, regarded as a function of $x \in Q_t$, clearly belongs to $S^+_{\alpha(t)}(Q_t, E)$; therefore, by Theorem 2.2.2 (see (2.2.15)), we have

$$t \in \Delta, \qquad \lambda(F_t, x) \leq \kappa \Rightarrow F_t(x) - \phi(t) \leq \alpha(t)g(\kappa),$$

where $\phi(t) = \inf\{F_t(y) \mid y \in Q_t\}$ and $g(\kappa) < \infty$ depends on $\kappa$ only. The function $\phi$ is upper semicontinuous by virtue of $(\Sigma.1)$ and hence is bounded from above on each compact set $\Delta^+ \subset \Delta$. It follows that

$$\{(t,x) \in X_*(\kappa) \mid t \in \Delta^+\} \equiv X(\Delta^+, \kappa) \subset X^*(a)$$

for some $a \in \mathbf{R}$. Thus there exists a set $Y(\Delta^+, \kappa)$, contained in $Q_*$ and closed in $E_\Delta$, such that $X(\Delta^+, \kappa) \subset Y(\Delta^+, \kappa)$. By $(\Sigma.1)$, the set $X_*(\kappa)$ is closed in $\Delta^+ \times E$, and therefore $X(\Delta^+, \kappa)$ is closed in $Q_*$. The latter fact is true for each compact set $\Delta^+$ contained in the interval $\Delta$, and hence $X_*(\kappa)$ is closed in $E_\Delta$.

(ii) Under the conditions of (ii.1), as well as those of (ii.2), $\mathcal{F}^+$ clearly satisfies $(\Sigma.1)$ and $(\Sigma.2)$. To verify $(\Sigma.3)$, or, respectively, $(\Sigma^+.3)$, consider the mapping

$$\pi(t,y) \equiv (t, \mathcal{A}(y)) : E_\Delta^+ \to E_\Delta.$$

Evidently, this mapping is continuous. We have

$$X_+^*(a) \equiv \{(t,y) \mid F_t^+(y) \leq a\} = \pi^{-1}(\{(t,x) \mid F_t(x) \leq a\}).$$

Therefore, under the assumptions of (ii.2), the sets $X_+^*(a)$ are closed in $E_\Delta^+$ for each $a \in \mathbf{R}$. It clearly follows that, if $\mathcal{F}$ satisfies (3.1.1) and (3.1.2) for some $(t,x) \in Q_*$, then the corresponding inequalities hold for $\mathcal{F}^+$ at all points $(t,y)$

such that $\pi(t,y) = (t,x)$. Implication (ii.2) is thereby proved. To prove (ii.1) following the same line of argument, it remains to show that, for any $\kappa$, the equality

$$X_*^+(\kappa) \equiv \{(t,y) \in Q_*^+ \mid \lambda(F_t^+, y) \le \kappa\} = \pi^{-1}(X_*(\kappa))$$

is true. The inclusion $X_*^+(\kappa) \supset \pi^{-1}(X_*(\kappa))$ is straightforward. To prove the inverse inclusion, we note that, if $(t,y) \in X_*^+(\kappa)$, then

$$\mid DF_t(\mathcal{A}(y))[Ah] \mid \le \alpha^{1/2}(t)\kappa\{D^2F_t(\mathcal{A}(y))[Ah, Ah]\}^{1/2}.$$

As $h$ runs throughout $E^+$, $Ah$ runs throughout $E$, since $\mathcal{A}$ is an onto mapping. Thus, we have $\pi(t,y) \in X_*(\kappa)$. (ii) is proved.

(iii) All the conditions that must be satisfied by $\mathcal{F}^+, \alpha^+, \gamma, \mu, \xi^+, \eta^+$ according to ($\Sigma$.1), ($\Sigma$.2), ($\Sigma^+$.3) are evident, excluding the closedness of the sets

$$X_+^*(a) \equiv \{(t,x) \in E_* \mid t \in \Delta, x \in Q_t^+, F_t^+(x) \le a\}, \qquad a \in \mathbf{R}$$

in $E_\Delta$. To prove that $X_+^*(a)$ is closed in $E_\Delta$, assume that $(t_i, x_i) \in X_+^*(a)$ and

$$(t_i, x_i) \to (t,x) \in E_\Delta \backslash X_+^*(a).$$

Then $(t,x)$ does not belong to at least one of the sets $Q_*$, $Q_*^*$ (otherwise, $(t,x)$ belongs to $Q_*^+$, and $F^+$ is continuous on this set). If $(t,x) \notin Q_*$, then $F_{t_i}(x_i) \to \infty$ for $i \to \infty$, owing to the closedness of the sets $\{(\tau, u) \in Q_* \mid F_\tau(u) \le \text{const}\}$ in $E_\Delta$, and, if $(t,x) \in Q_*$, then $\{F_{t_i}(x_i)\}$ is bounded from below by virtue of the continuity of $F$ on $Q_*$. By the same reasons, $\{F_{t_i}^*(x_i)\}$ either is bounded or tends to $+\infty$. Since at least one of the sequences $\{F_{t_i}(x_i)\}$, $\{F_{t_i}^*(x_i)\}$ tends to $+\infty$ and since both of these sequences are bounded from below, we have $F_{t_i}^+(x_i) \to \infty$, which contradicts the inclusion $(t_i, x_i) \in X_+^*(a)$, $i \ge 1$. □

### 3.1.2 Metric of self-concordant family

Our final goal is to relate the rate of varying the penalty parameter $t$ in the path-following scheme and the number of Newton steps sufficient to restore closeness to the trajectory after a step in $t$. We are sure about the family consisting of self-concordant functions and we know the main result concerning the Newton method, applied to such a function in terms of the Newton decrement, converging quadratically with objective-independent rate from any point where the Newton decrement is less than an appropriate absolute constant. Thus, the simplest way to control the number of Newton steps in $x$ per step in $t$ is to maintain a fixed upper bound for the Newton decrement of the new function of the family at the previous approximation. It turns out that we can associate with a self-concordant family an explicit metric $\rho$ on the parameter space $\Delta$ with the following property: When $\rho$-distance between two values, $t$ and $t'$, of the penalty is small, then the Newton decrements of $F_t$ and $F_{t'}$ at a point $x$ are close to each other. Thus, the simplest way to control the magnitudes of the Newton decrements is to perform steps in $t$ of small enough $\rho$-length.

In this section, we introduce the metric $\rho$ and, in the next, we establish the relation between the $\rho$-length of a step in $t$ and the associated variation of the Newton decrement.

Assume that $\mathcal{F} = \{Q_t, F_t, E\}_{t\in\Delta} \in \Sigma_\Delta(\alpha, \gamma, \mu, \xi, \eta, \kappa)$. We set

$$\psi(\mathcal{F}; t) = \frac{(\gamma(t)\alpha(t))^{1/2}}{\mu(t)} \tag{3.1.3}$$

and introduce the following metrics on $\Delta$ depending on a parameter $\nu > 0$:

(3.1.4)

$$\rho_\nu(\mathcal{F}; t, \tau) = \max\left\{\left|\ln\left(\frac{\psi(\mathcal{F}, u)}{\psi(\mathcal{F}, v)}\right)\right| \mid u, v \in [t, \tau]\right\} + \frac{1}{\nu}\left|\int_t^\tau \xi(s)ds\right| + \left|\int_t^\tau \eta(s)ds\right|.$$

The following result, which can be proved by a straightforward computation, shows that the property of self-concordance and the metrics associated with a family are invariant with respect to scalings and monotone (nonlinear) substitutions of the parameter.

**Proposition 3.1.2** *Let $\mathcal{F} = \{Q_t, F_t, E\}_{t\in\Delta} \in \Sigma_\Delta(\alpha, \gamma, \mu, \xi, \eta, \kappa)$. Let $\Delta^+$ be an open interval on the real axis, $p(t)$ be a continuously differentiable positive function on $\Delta$, and $\pi(\tau)$ be a continuously differentiable one-to-one mapping from $\Delta^+$ onto $\Delta$. Denote*

$$\mathcal{F}^+ = \{Q_{\pi(\tau)}, p(\pi(\tau))F_{\pi(\tau)}, E\}_{\tau\in\Delta^+}.$$

*Then $\mathcal{F}^+ \in \Sigma_{\Delta^+}(\alpha^+, \gamma^+, \mu^+, \xi^+, \eta^+, \kappa)$, where*

$$\alpha^+(\tau) = \alpha(\pi(\tau))p(\pi(\tau)), \quad \mu^+(\tau) = \mu(\pi(\tau))p(\pi(\tau)), \quad \gamma^+(\tau) = \gamma(\pi(\tau))p(\pi(\tau)),$$

$$\xi^+(\tau) = \xi(\pi(\tau)) \mid \pi'(\tau) \mid, \qquad \eta^+(\tau) = \eta(\pi(\tau)) \mid \pi'(\tau) \mid,$$

*and, for all $\nu > 0$, $\tau$, $\tau' \in \Delta^+$, we have*

$$\rho_\nu(\mathcal{F}^+; \tau, \tau') = \rho_\nu(\mathcal{F}; \pi(\tau), \pi(\tau')).$$

### 3.1.3 Main property of self-concordant families

**Theorem 3.1.1** *Let $\mathcal{F} = \{Q_t, F_t, E\}_{t\in\Delta} \in \Sigma_\Delta(\alpha, \gamma, \mu, \xi, \eta, \kappa)$. Assume that $(t, x) \in Q_*$ satisfies the inequality $\lambda(F_t, x) < \kappa$ and that $t' \in \Delta$ satisfies*

$$\rho_\kappa(\mathcal{F}; t, t') \le 1 - \frac{\lambda(F_t, x))}{\kappa}. \tag{3.1.5}$$

*Then $(t', x) \in Q_*$, and*

$$\lambda(F_{t'}, x) \le \kappa. \tag{3.1.6}$$

**Proof.** $1^0$. Let $\delta = \{\tau \mid \tau \in \Delta,\ (\tau, x) \in X^+(\kappa)\}$ (the set $X_+(\kappa)$ was introduced in item ($\Sigma$.3) of Definition 3.1.1). Then $\delta$ is open in $\Delta$ and contains $t$. Let us denote by $\delta^*$ the connection component of $t$ in $\delta$.

$2^0$. Let us fix some $h \in E$ and consider the following two functions of $\tau \in \delta^*$:

$$a(\tau) = DF_\tau(x)[h], \qquad b(\tau) = D^2F_\tau(x)[h,h]. \tag{3.1.7}$$

By ($\Sigma$.3), we have (with $(\cdot)'$ denoting the derivative in $\tau$)

$$\mid a'(\tau) - (\ln(\mu(\tau)))'a(\tau) \mid \leq \alpha^{1/2}(\tau)\xi(\tau)b^{1/2}(\tau), \tag{3.1.8}$$

$$\mid b'(\tau) - (\ln(\gamma(\tau)))'b(\tau) \mid \leq 2\eta(\tau)b(\tau). \tag{3.1.9}$$

By (3.1.9), we have either $b(\tau) \equiv 0$, $\tau \in \delta^*$ (case $\mathrm{I}_h$), or $b(\tau)$ does not attain the zero value on $\delta^*$ (case $\mathrm{II}_h$). In case $\mathrm{I}_h$, by virtue of (3.1.8) and the inequality

$$\mid a(t) \mid \leq \lambda(F_t, x)\alpha^{1/2}(t)b^{1/2}(t) = 0,$$

we have $a(\tau) \equiv 0$, $\tau \in \delta^*$.

$3^0$. Now assume that $\mathrm{II}_h$ is the case. Set

$$\phi(\tau) = \frac{a(\tau)}{(\alpha(\tau)b(\tau))^{1/2}}, \qquad \tau \in \delta^*.$$

Let $t'' \in \delta^*$ satisfy

$$\rho_\kappa(\mathcal{F}; t, t'') < \rho_\kappa(\mathcal{F}; t, t'). \tag{3.1.10}$$

Denote by $t^*$ the point of the segment $[t, t'']$, nearest to $t''$, at which $\phi$ equals zero, if such a point exists; otherwise, let $t^* = t$. Let $\delta^+$ be the segment with the endpoints $t^*$ and $t''$. We have

$$\rho_\kappa(\mathcal{F}; t^*, t'') < \rho_\kappa(\mathcal{F}; t, t');$$

$$\phi(t^*) \leq \lambda \equiv \lambda(F_t, x); \qquad \phi(\tau) \neq 0,\ \tau \in (t^*, t''] \equiv \delta_0^+. \tag{3.1.11}$$

For $\tau \in \delta^+$, the function $\phi(\tau)$ is continuous, and, for $\tau \in \delta_0^+$, it is continuously differentiable and does not attain zero value.

For $\tau \in \delta_0^+$, we have

$$2\phi'(\tau)\phi(\tau) = \frac{2a^2(\tau)(\ln(\mu(\tau)))'}{\alpha(\tau)b(\tau)} - \frac{a^2(\tau)(\ln(\gamma(\tau)))'}{\alpha(\tau)b(\tau)} - \frac{a^2(\tau)(\ln(\alpha(\tau)))'}{a(\tau)b(\tau)} + \omega(\tau),$$

where

$$\omega(\tau) = \frac{2a(\tau)(a'(\tau) - (\ln(\mu(\tau)))'a(\tau))}{\alpha(\tau)b(\tau)} - \frac{a^2(\tau)(b'(\tau) - (\ln(\gamma(\tau)))'b(\tau))}{\alpha(\tau)b^2(\tau)}.$$

Since $\tau \in \delta^*$, we have $(\tau, x) \in X^+(\kappa)$ and, by (3.1.1) and (3.1.2), we obtain

$$\mid \omega(\tau) \mid \leq \frac{2\alpha^{1/2}(\tau)\xi(\tau)a(\tau)b^{1/2}(\tau)}{\alpha(\tau)b(\tau)} + \frac{2a^2(\tau)\eta(\tau)b(\tau)}{\alpha(\tau)b^2(\tau)} = 2\phi(\tau)\xi(\tau) + 2\eta(\tau)\phi^2(\tau).$$

Thus, for $\tau \in \delta_0^+$, we have

$$| \phi'(\tau) + \{\ln(\psi(\mathcal{F};\tau)\}'_\tau \phi(\tau) | \leq \xi(\tau) + \phi(\tau)\eta(\tau). \tag{3.1.12}$$

Let $\phi^* = \max\{\phi(\tau) \mid \tau \in \delta^+\}$. By (3.1.12),

$$| \phi'(\tau) + \{\ln(\psi(\mathcal{F};\tau)\}'_\tau \phi(\tau) | \leq \xi(\tau) + \phi^*\eta(\tau), \qquad \tau \in \delta_0^+. \tag{3.1.13}$$

Let $\psi_- = \min\{\psi(\tau) \mid \tau \in \delta^+\}$, $\psi_+ = \max\{\psi(\tau) \mid \tau \in \delta^+\}$. Then, by (3.1.13) and the continuity of $\phi$ on $\delta^+$, we have

$$\tau \in \delta^+ \Rightarrow \phi(\tau) \leq \{\phi(t^*) + \kappa\rho_2 + \phi^*\rho_3\}e^{\rho_1}, \tag{3.1.14}$$

where

$$\rho_1 = \ln\frac{\psi_+}{\psi_-}, \quad \rho_2 = \frac{1}{\kappa}\left|\int_{t^*}^{t''} \xi(s)ds\right|, \quad \rho_3 = \left|\int_{t^*}^{t''} \eta(s)ds\right|.$$

Thus,

$$\rho_1 + \rho_2 + \rho_3 = \rho_\kappa(\mathcal{F}; t^*, t'') < 1 - \frac{\lambda}{\kappa}. \tag{3.1.15}$$

By (3.1.15), we have $\rho_1 + \rho_3 < 1$, so that $e^{\rho_1}\rho_3 < 1$, which, by (3.1.14), leads to

$$\phi^* \leq \frac{(\phi(t^*) + \kappa\rho_2)e^{\rho_1}}{1 - \rho_3 e^{\rho_1}}.$$

This inequality, by virtue of (3.1.15) and the first relation in (3.1.11), implies $\phi^* \leq \kappa$.

In view of the definition of $t^*$, the latter inequality means that the implication

$$\{t'' \in \delta^*,\ \rho_\kappa(\mathcal{F}; t, t'') < \rho_\kappa(\mathcal{F}; t, t')\} \Rightarrow \max\{\phi(\tau) \mid \tau \in [t, t'']\} \leq \kappa$$

holds in case $II_h$. By the continuity argument, this proves the implication

$$\{t'' \in \delta^*,\ \rho_\kappa(\mathcal{F}; t, t'') \leq \rho_\kappa(\mathcal{F}; t, t')\} \Rightarrow \max\{\phi(\tau) \mid \tau \in [t, t'']\} \leq \kappa. \tag{3.1.16}$$

Taking into account the definition of $\phi$, we obtain from (3.1.16) that

$$\begin{aligned} &\{t'' \in \delta^*,\ \rho_\kappa(\mathcal{F}; t, t'') \leq \rho_\kappa(\mathcal{F}; t, t')\} \\ &\Rightarrow | DF_{t''}(x)[h] | \leq \alpha^{1/2}(t'')\kappa(D^2F_{t''}(x)[h,h])^{1/2}. \end{aligned} \tag{3.1.17}$$

Relation (3.1.17) has been proved in case $II_h$; in case $I_h$ (where, as we have seen, $DF_{t''}(x)[h] = 0$, $t'' \in \delta^*$), it is evident. Thus, we have

$$\{t'' \in \delta^*,\ \rho_\kappa(\mathcal{F}; t, t'') \leq \rho_\kappa(\mathcal{F}; t, t')\} \Rightarrow \lambda(F_{t''}, x) \leq \kappa. \tag{3.1.18}$$

$4^0$. To complete the proof, it suffices to show that $t' \in \delta^*$; it allows us to take $t'' = t'$ in (3.1.18). If $t' \notin \delta^*$, then there exists $t^+$, which lies between $t$ and

$t'$ and is a boundary point of the interval $\delta^*$. Assume that $t_i$ lies in $\delta^*$ between $t$ and $t^+$ and tend to $t^+$ as $i \to \infty$. Each $t_i$ satisfies the premise in (3.1.18) (since $t_i$ lies between $t$ and $t'$ and belongs to $\delta^*$); hence, by virtue of (3.1.18), the inclusions $(t_i, x) \in X_*(\kappa)$ hold. As $i \to \infty$, the points $(t_i, x)$ converge to $(t^+, x)$. The latter point belongs to $E_\Delta$, since $t$ lies between $t^+ \in \Delta$ and $t' \in \Delta$ and hence itself belongs to $\Delta$. Since $X_*(\kappa)$ is closed in $E_\Delta$, we have $(t^+, x) \in X_*(\kappa)$. Hence, for all $\tau \in \Delta$ that are close enough to $t^+$, the points $(\tau, x)$ belong to $X^+(\kappa)$; so all these $\tau$ belong to $\delta$. This contradicts the assumption that $t^+$ is a boundary point of $\delta^*$. $\square$

Combining Theorem 3.1.1 with the results on Newton's method from §2.2, we obtain the following result.

**Corollary 3.1.1** *Let $\mathcal{F} = \{Q_t, F_t, E\}_{t\in\Delta} \in \Sigma_\Delta(\alpha, \gamma, \mu, \xi, \eta, \kappa)$, let $(t_0, x_{-1}) \in Q_*$ be a point such that*

$$\lambda(F_{t_0}, x_{-1}) \leq \kappa, \tag{3.1.19}$$

*and let $\kappa^+ = \kappa^2(1-\kappa)^{-2}$. Assume that $\{t_i \in \Delta\}_{i\geq 0}$ satisfies the relations*

$$\rho_\kappa(\mathcal{F}; t_i, t_{i+1}) \leq 1 - \frac{\kappa^+}{\kappa}, \qquad i \geq 0. \tag{3.1.20}$$

*Let*

$$x_i = x^*(F_{t_i}, x_{i-1}) \tag{3.1.21}$$

*be the Newton iterate of $x_{i-1}$, the Newton method being applied to $F_{t_i}$.*

*Then $x_i$ are well defined and belong to $Q_{t_i}$, and*

$$\lambda(F_{t_i}, x_i) \leq \kappa \tag{3.1.22}$$

*for all $i \geq 0$.*

Thus, given a sufficiently close approximation, $x_{-1}$, to a $F_{t_0}$-center of $Q_{t_0}$ (that is, to a minimizer of $F_{t_0}$), we can follow the path $x^*(t)$ formed by the minimizers of $F_t$, using a fixed-length steps in $t$ and a single Newton step in $x$ per each step in $t$ (the length of $t$-steps is measured in terms of the metric associated with the family).

In the next sections of this chapter, we describe several self-concordant families and the corresponding polynomial-time algorithms.

## 3.2 Barrier-generated path-following method

In this section, we present a barrier method for the problem

$$(f): \quad f(x) \to \min \mid x \in G \subset E.$$

### 3.2.1 Barrier-generated family

Barrier method is a path-following method associated with the family

$$F_t(x) = tf(x) + F(x),$$

where $F$ a self-concordant barrier for the feasible region $G$.

To provide self-concordance of the latter family, we impose certain restrictions on $f$ (as we see, these restrictions are satisfied at least by linear and convex quadratic objectives, so that our approach covers all standard problems).

**Definition 3.2.1** *Let $G$ be a closed convex domain in $E$, let $F$ be a $\vartheta$-self-concordant barrier for $G$, and let $\vartheta \geq 1$, $\beta \geq 0$. A function $f : G \to \mathbf{R} \bigcup \{+\infty\}$ is called $\beta$-compatible with $F$ (notation: $f \in \mathcal{C}(F, \beta)$) if $f$ is lower semicontinuous convex function on $G$, which is finite and $C^3$-smooth on $\operatorname{int} G$ and such that, for all $x \in \operatorname{int} G$ and $h \in E$, the following inequality holds:*

$$(3.2.1) \qquad | D^3 f(x)[h,h,h] | \leq \beta \{3D^2 f(x)[h,h]\}\{3D^2 F(x)[h,h]\}^{1/2}.$$

The following statement quite straightforwardly follows from the definitions (compare with Proposition 2.3.1).

**Proposition 3.2.1** *Let $G$ be a closed convex domain in $E$ and let $F$ be a $\vartheta$-self-concordant barrier for $G$. Then*

(i) $\beta$-compatibility is stable with respect to affine substitutions of argument. *Let $\mathcal{A}(y) \equiv Ay + b$ be an affine mapping from $E^+$ into $E$ such that $\mathcal{A}(E^+) \bigcap \operatorname{int} G \neq \emptyset$ and let $G^+ = \mathcal{A}^{-1}(G)$, $f \in \mathcal{C}(F, \beta)$, $F^+(y) = F(\mathcal{A}(y)) : \operatorname{int} G^+ \to \mathbf{R}$, $f^+(y) = f(\mathcal{A}(y)) : \operatorname{int} G^+ \to \mathbf{R}$. Then $F^+ \in \mathcal{B}(G^+, \vartheta)$ and $f^+ \in \mathcal{C}(F^+, \beta)$;*

(ii) A convex quadratic function is 0-compatible with the barrier. *If $f$ is a convex quadratic form on $E$, then $f \in \mathcal{C}(F, 0)$,*

Barrier is $O(1)$-compatible with itself. $F \in \mathcal{C}(F, 2/3^{3/2})$ *($F$ is extended onto $G$ by setting $F(x) = +\infty$ for $x \in \partial G$),*

The sum of functions compatible with a barrier is also compatible with it. *If $f_i \in \mathcal{C}(F, \beta_i)$, $p_i \geq 0$, $i = 1, 2$, then $p_1 f_1 + p_2 f_2 \in \mathcal{C}(F, \max\{\beta_1, \beta_2\})$;*

(iii) $\beta$-compatibility is stable with respect to summation of barriers. *Let $G_i$ be closed convex domains in $E$ and let $F_i \in \mathcal{B}(G_i, \vartheta_i)$, $1 \leq i \leq m$ be such that $G^+ = \bigcap_{i=1}^m G_i$ has a nonempty interior. Then a function from $\mathcal{C}(F_i, \beta)$ being reduced onto $G^+$ is $\beta$-compatible with the barrier $F^+ = \sum_{i=1}^m F_i$ for $G^+$.*

The following fact underlies our further considerations.

**Proposition 3.2.2** *Let $G$ be a closed convex domain, let $F$ be a $\vartheta$-self-concordant barrier for $G$, and let $f$ be $\beta$-compatible with $F$. Denote $\Delta = (0, \infty)$ and consider the family*

$$\mathcal{F} = \mathcal{F}(F, f) = \{Q_t \equiv \operatorname{int} G, \;\; F_t(x) = tf(x) + F(x), E\}_{t \in \Delta}.$$

*The family $\mathcal{F}$ is strongly self-concordant with the parameters*

$$\alpha(t) = \frac{1}{(1+\beta)^2}, \quad \mu(t) \equiv \gamma(t) \equiv t,$$

$$\xi(t) = (1+\beta)\frac{\vartheta^{1/2}}{t}, \quad \eta(t) = \frac{1}{2t}. \tag{3.2.2}$$

*In particular,*

$$\psi(\mathcal{F}, t) = \frac{1}{(1+\beta)t^{1/2}}$$

*and*

$$\rho_\nu(\mathcal{F}; t, \tau) = \left\{1 + (1+\beta)\frac{\vartheta^{1/2}}{\nu}\right\}\left|\ln\frac{t}{\tau}\right|. \tag{3.2.3}$$

**Proof.** Let us verify that, under the choice of the parameters described in (3.2.2), relations ($\Sigma$.1), ($\Sigma$.2), and ($\Sigma^+$.3) hold. Relation ($\Sigma$.1) is evident. To prove ($\Sigma$.2), let us note that, by virtue of $F \in S_1^+(\operatorname{int} G, E)$, we have, for $x \in \operatorname{int} G$, $h \in E$,

$$| D^3F(x)[h,h,h] | \leq 2\{D^2F_t(x)[h,h]\}^{3/2}. \tag{3.2.4}$$

For given $x$ and $h$, let

$$p = \{D^2f(x)[h,h]\}^{1/2}, \qquad q = \{D^2F(x)[h,h]\}^{1/2}.$$

Then, in view of $f \in \mathcal{C}(F, \beta)$, we have

$$\begin{aligned}
| D^3f(x)[h,h,h] | &\leq 3^{3/2}\beta p^2 q = 2\beta\left(\frac{3^{3/2}}{2}p^2 q\right) \\
&= 2\beta\left\{\left(\frac{3}{2}p^2t^{1/3}\right)^{2/3}\left(\frac{3q^2}{t^{2/3}}\right)^{1/3}\right\}^{3/2} \\
&\leq 2\beta\left\{\frac{2}{3}\left(\left(\frac{3}{2}\right)^{2/3}p^{4/3}t^{2/9}\right)^{3/2} + \frac{1}{3}\left(\frac{3^{1/3}q^{2/3}}{t^{2/9}}\right)^3\right\}^{3/2} \\
&= 2\beta\left\{p^2t^{1/3} + \frac{q^2}{t^{2/3}}\right\}^{3/2} \\
&= \frac{2\beta}{t}\{p^2t + q^2\}^{3/2} = \frac{2\beta}{t}\{D^2F_t(x)[h,h]\}^{3/2}.
\end{aligned}$$

This relation, by (3.2.4), implies that

$$t\,| D^3f(x)[h,h,h] | + | D^3F(x)[h,h,h] | \leq 2(1+\beta)\{D^2F_t(x)[h,h]\}^{3/2},$$

and the latter relation combined with (3.2.2) leads to the inequality required in ($\Sigma$.2).

It remains to verify ($\Sigma^+.3$). The closedness of the sets $\{(t,x) \mid t \in \Delta,\ F_t(x) \le a\}$ in $E_\Delta$ is an immediate consequence of the inclusion $F \in S_1^+(\operatorname{int} G, E)$ and the fact that $f$ is below bounded on each bounded subset of $\operatorname{int} G$. Let us prove that, for $x \in \operatorname{int} G$, $h \in E$, relations (3.1.1), (3.1.2) hold. By $F \in \mathcal{B}(G,\vartheta) \subset S_1^+(\operatorname{int} G, E)$, we have

$$\begin{aligned}\left|\{DF_t(x)[h]\}'_t - \frac{1}{t}DF_t(x)[h]\right| &= \frac{1}{t} \mid DF(x)[h] \mid \le \frac{\lambda(F,x)}{t}\{D^2F(x)[h,h]\}^{1/2} \\ &\le \frac{\vartheta^{1/2}}{t}\{D^2F_t(x)[h,h]\}^{1/2} \\ &= \frac{\vartheta^{1/2}}{t}(1+\beta)\alpha^{1/2}(t)\{D^2F_t(x)[h,h]\}^{1/2},\end{aligned}$$

which is required in (3.1.1). Furthermore,

$$\left|\{D^2F_t(x)[h,h]\}'_t - \frac{1}{t}D^2F_t(x)[h,h]\right| = \frac{1}{t}D^2F(x)[h,h] \le \frac{1}{t}D^2F_t(x)[h,h],$$

which leads to (3.1.2). $\square$

### 3.2.2 Barrier method

Let us fix a *bounded* closed convex domain $G$ in $E$, a $\vartheta$-self-concordant barrier $F$ for $G$ and $\beta \ge 0$. Our purpose is to describe a method solving ($f$) under the assumption that the objective $f$ is $\beta$-compatible with $F$.

The method is based on the following ideas.

(A) The trajectory of minimizers

$$x^*(t) = \operatorname{argmin}\{F_t(x) \mid x \in \operatorname{int} G\},$$

i.e., of the points with $\lambda(F_t, x) = 0$, clearly converges to the solution set of ($f$), and the error $\varepsilon(t) = f(x^*(t)) - \min_G f$ is bounded from above by the quantity $\vartheta/t$. Indeed, we have $f'(x^*(t)) = -t^{-1}F'(x^*(t))$ (evident) and $\langle F'(x), y-x\rangle \le \vartheta$, $y \in G$, $x \in \operatorname{int} G$ (see (2.3.2)), so that $\langle f'(x^*(t)), x^*(t) - y\rangle \le \vartheta/t$. Since $f$ is convex, it follows that $\varepsilon(t) \le \vartheta/t$.

Thus, a point of the trajectory $x^*(\cdot)$ associated with large $t$ is a good approximate solution to ($f$). As we see below (Proposition 3.2.4), a point $x$ close enough to $x^*(t)$, namely, such that $\lambda(F_t, x) \le \delta$ with small enough absolute constant $\delta$ (say, $\delta = 0.01$) is also a good approximate solution to ($f$): It turns out that, for the above $x$, we have $f(x) - \min_G f \le 2\vartheta/t$. Thus, it is reasonable to follow the trajectory $x^*(t)$ as $t$ increases.

(B) Theorem 3.1.1 and Proposition 3.2.2 demonstrate that we can follow the trajectory $x^*(t)$ along a sequence of values of $t$ increasing at the ratio

$$\kappa = 1 + \frac{O(1)}{\vartheta^{1/2}(1+\beta)}$$

wherein $O(1)$ is an appropriate absolute constant. Namely, given a pair $(x,t)$ with $x$ close enough to $x^*(t)$, i.e., such that $\lambda(F_t,x) \le 0.01$ and replacing $t$ by $t' = \kappa t$, we obtain "an intermediate" pair $(x,t')$, which is not too far from $x^*(t')$: $\lambda(F_{t'},x) \le 0.02$ (see Theorem 3.1.1 and Proposition 3.2.2). To restore the initial accuracy 0.01 of approximation of $x^*(t')$, it suffices to apply the Newton minimization method to $F_{t'}$. By virtue of Theorem 2.2.2(ii), the Newton iterate $x'$ of $x$ satisfies the inequality $\lambda(F_{t'},x') \le 0.01$, and we are in the same position as before, but with larger value of $t$.

In view of (A), the above procedure produces approximations, $x_i$, to the solution of $(f)$, such that

$$f(x_i) - \min_G f \le \frac{1}{t_0} \exp\left\{-\frac{O(1)}{(1+\beta)\vartheta^{1/2}} i\right\},$$

where $(x_0,t_0)$ is the initial pair that should satisfy the relation

$$\lambda(F_{t_0},x_0) \le \delta.$$

(C) It remains to explain how to find a pair satisfying the latter relation. Note that, if $x_0$ is the minimizer $x(F)$ of $F$ and $t_0 = 0$, then $\lambda(F_0,x_0) = 0$, and therefore $\lambda(F_t,x)$ is small for all $x$ close to $x(F)$ and all small enough positive $t$. Thus, it suffices to approximate the $F$-center $x(F)$ of $G$. It can be done by the same path-following technique, provided that we are given an interior point $w$ of $G$. Namely, consider the family

$$\Phi(x) = t\,\langle -F'(w), x\rangle + F(x),$$

which also is self-concordant. The trajectory of minimizers of the latter family passes through $w$ (the corresponding $t$ is equal to 1) and converges to $x(F)$ as $t \to 0$. We can follow the trajectory starting with the pair $(w,1)$ in the manner described in (A), but now decreasing the parameter at the ratio $\kappa$ instead of increasing it. It can be proved (see Proposition 3.2.3, below) that, after a finite number of steps (which is of order of $\vartheta^{1/2}$, with the constant factor depending on $w$), we obtain a point that, for an appropriately chosen $t_0 > 0$, can be used as the above $x_0$ (another possibility for initialization of the main stage is to use primal-dual conic reformulation of the initial problem; see §§4.3.1 and 4.3.5).

Now let us present the detailed description of the method.

Let the functions $\lambda^+(\lambda)$ and $\zeta(\lambda)$ be defined as

$$\lambda^+(\lambda) = \left(\frac{\lambda}{1-\lambda}\right)^2, \qquad 0 < \lambda < 1,$$

$$\zeta(\lambda) = \frac{\omega^2(\lambda)(1+\omega(\lambda))}{1-\omega(\lambda)}, \qquad 0 < \lambda < \frac{1}{3}$$

(recall that $\omega(\lambda) = 1 - (1-3\lambda)^{1/3}$; see Theorem 2.2.2).

Since $G$ is bounded, $D^2F(x)[\cdot,\cdot]$ is a nondegenerate inner product on $E$ (Proposition 2.3.2(ii)); this product is denoted by $\langle h, e\rangle_{x,F}$, and the corresponding norm, same as in Chapter 2, by $\|\cdot\|_{x,F}$. In these notations we omit the subscript $x$ in the case when $x = x(F)$ is the minimizer of $F$ over $\operatorname{int} G$. Note that this minimizer does exist and is unique (see Proposition 2.3.2(ii)).

Given a norm $\|\cdot\|$ on a finite-dimensional linear space $E$, we denote by $\| y \|^*$ the conjugate norm of a linear functional $\langle y, \cdot\rangle$ on $E$, below:

$$\| y \|^* = \max_{\|x\|\le 1} \langle y, x\rangle .$$

The barrier method is specified by the parameters $\lambda_1'$, $\lambda_1$, $\lambda_2$, $\lambda_3'$, $\lambda_3$, such that

$$0 < \lambda^+(\lambda_1) \le \lambda_1' < \lambda_2 < \lambda_3 < \lambda_*;$$

$$\lambda_1' < \lambda_1 < \lambda_*; \qquad \lambda^+(\lambda_3) \le \lambda_3' < \lambda_3; \tag{3.2.5}$$

$$\zeta(\lambda_1') \le \frac{1}{9}, \qquad (1+\beta)\lambda_2 < \lambda_3;$$

$$\frac{\omega(\lambda_3')}{1-\omega(\lambda_3')} < 1, \qquad \frac{\omega(\lambda_2)}{1-\omega(\lambda_2))} \le \frac{1}{3}, \tag{3.2.6}$$

and by a *starting point*

$$w \in \operatorname{int} G. \tag{3.2.7}$$

The method consists of two stages, *the preliminary stage* and *the main stage.*

### 3.2.3 Preliminary stage

This stage results in an approximation $u$ of $x(F)$ such that $\lambda(F,u) \le \lambda_2$. To find such an $u$, the method follows the trajectory of minimizers of the family

$$\mathcal{F}^{(1)} \equiv \mathcal{F}(F,g) = \{\operatorname{int} G,\ F_t^{(1)}(x) = tg(x) + F(x), E\}_{t>0}$$

as $t \to 0$; herein

$$g(x) = -DF(w)[x-w]. \tag{3.2.8}$$

Clearly, $g$ is 0-compatible with $F$, so that the family $\mathcal{F}^{(1)}$ is strongly self-concordant (Proposition 3.2.2). It is important that the trajectory of minimizers associated with the family passes through the point $w$ (the corresponding $t$ equals to 1). Note that, for this family, we have

$$\alpha(t) \equiv 1; \qquad \rho_\nu(\mathcal{F}^{(1)}; t, t') = \left(1 + \frac{\vartheta^{1/2}}{\nu}\right)\left|\ln \frac{t}{t'}\right|. \tag{3.2.9}$$

The desired approximation is constructed as follows. We set

$$t_i = \kappa_1^{-i}, \quad i \ge 0, \quad \kappa_1 = \exp\left\{\frac{\lambda_1 - \lambda_1'}{\lambda_1 + \vartheta^{1/2}}\right\}, \tag{3.2.10}$$

so that

(3.2.11) $$\rho_{\lambda_1}(\mathcal{F}^{(1)}; t_i, t_{i+1}) \leq 1 - \frac{\lambda_1'}{\lambda_1}, \qquad i \geq 0$$

and then compute the points $x_i$ according to the relations

(3.2.12) $$x_{-1} = w; \qquad x_i = x^*(F_{t_i}^{(1)}, x_{i-1}), \quad i \geq 0$$

(recall that $x^*(H, x)$ denotes the Newton iterate of $x$, the Newton minimization method being applied to $H$; see §2.2).

Process (3.2.12) is terminated at the first moment $i = i^*$ at which the relation

(3.2.13) $$\lambda(F, x_{i-1}) \leq \lambda_2$$

holds. The result of the preliminary stage is the point

(3.2.14) $$u = x_{i^*-1}.$$

**Proposition 3.2.3** (i) *The preliminary stage is well defined; i.e., all $x_i$ are well defined and belong to* $\operatorname{int} G$, $-1 \leq i \leq i^*$. *Furthermore, $i^* < \infty$, and the following relations hold:*

(3.2.15) $$(\mathrm{I}_i): \qquad \lambda(F_{t_i}^{(1)}, x_{i-1}) \leq \lambda_1,$$

(3.2.16) $$(\mathrm{J}_i): \qquad \lambda(F_{t_i}^{(1)}, x_i) \leq \lambda_1';$$

(ii) *The vector $u$ obtained at the preliminary stage satisfies the relations*

(3.2.17) $$\lambda(F, u) \leq \lambda_2,$$

(3.2.18) $$u \in W_{1/3}(x(F))$$

*(recall that $W_r(x) = \{y \mid D^2F(x)[y - x, y - x] \leq r^2\}$ is the Dikin's ellipsoid of $F$);*

(iii) *The number $i^*$ of the iterations at the preliminary stage satisfies the inequality*

(3.2.19) $$i^* \leq 1 + \frac{\lambda_1 + \vartheta^{1/2}}{\lambda_1 - \lambda_1'} \left( \ln \frac{31}{\lambda_2 - \lambda_1'} + \ln \frac{\vartheta}{1 - \pi_{x(F)}(w)} \right)$$

*(recall that $x(F)$ is the minimizer of $F$ over* $\operatorname{int} G$ *and that $\pi_u(\cdot)$ is the Minkowsky function of $G$ with the pole at $u$).*

**Proof.** $1^0$. Let us prove (3.2.15) and (3.2.16) by induction. By the definitions of $g$ and $t_0$, we have $\lambda(F_{t_0}^{(1)}, x_{-1}) = 0$, so that $(\mathrm{J}_0)$ holds. Assume that $x_{i-1}$ are well defined and belong to $\operatorname{int} G$ and that relations $(\mathrm{I}_i)$, $0 \leq i \leq k$ and $(\mathrm{J}_i)$, $0 \leq i < k$ hold. Relation $(\mathrm{I}_k)$, by virtue of $\lambda_1 \in (0, \lambda_*)$ and Theorem 2.2.2(ii) (the theorem is applied with $F = F_{t_k}^{(1)}$, $x = x_{k-1}$), implies that $x_k$ is well defined and $(\mathrm{J}_k)$ holds. Furthermore, by (3.2.14), (3.2.11), Proposition 3.2.2 and Theorem 3.1.1 (applied with $\kappa = \lambda_1$, $x = x_k$, $t = t_k$, $t' = t_{k+1}$, $F. = F.^{(1)}$), relation $(\mathrm{I}_{k+1})$ also holds. The induction is completed.

$2^0$. Let us prove that $i^* < \infty$ and that (iii) holds. Let us fix $i < i^* - 1$ and denote $F_{t_i}^{(1)}$ by $\Phi$. Then $\Phi \in S_1^+(\operatorname{int} G, E)$ and $\lambda(\Phi, x_i) \leq \lambda_1' < \lambda_* < \frac{1}{3}$ by virtue of $(\mathrm{I}_i)$. In view of (3.2.17), we have

$$\Phi(x_i) - \Phi(x(F)) \leq \Phi(x_i) - \inf\{\Phi(x) \mid x \in \operatorname{int} G\} \leq \tfrac{1}{2}\zeta(\lambda_1') \leq \tfrac{1}{18} \tag{3.2.20}$$

(we have considered (3.2.5), (3.2.6)). By Theorem 2.1.1 and Proposition 2.3.2(i.1), we have

$$\{\| e \|_F = 1,\ t \in [0,1)\} \Rightarrow \left\{x(F) + te \in \operatorname{int} G,\ \frac{d^2}{dt^2}F(x(F) + te) \geq (1-t)^2\right\}$$

or, in view of $(d/dt)F(x(F) + te) \mid_{t=0} = 0$,

$$\begin{aligned} &\| e \|_F = 1, \qquad t \in [0,1) \Rightarrow x(F) + te \in \operatorname{int} G, \\ &F(x(F) + te) - F(x(F)) \geq (6 - 4t + t^2)\frac{t^2}{12}, \end{aligned} \tag{3.2.21}$$

and hence

$$\begin{aligned} &\| e \|_F = 1,\ t \in [0,1) \Rightarrow \Phi(x(F) + te) - \Phi(x(F)) \\ &\geq t^2\frac{6 - 4t + t^2}{12} - t \cdot t_i \| Dg \|_F^* \end{aligned} \tag{3.2.22}$$

(recall that, for a norm $\| \cdot \|$ on $E$, $\| y \|^*$ denotes the conjugate norm of a linear functional $y : E \to \mathbf{R}$). By virtue of (2.3.5), we have

$$\| Dg \|_F^* \leq \frac{\vartheta}{1 - \pi_{x(F)}(w)} \equiv \Omega,$$

and hence we obtain

$$\begin{aligned} &\| e \|_F = 1,\ t \in [0,1) \Rightarrow \Phi(x(F) + te) - \Phi(x(F)) \\ &\geq (6 - 4t + t^2)\tfrac{t^2}{12} - t \cdot t_i\Omega. \end{aligned} \tag{3.2.23}$$

Let us verify that

$$t_i \geq \min\left\{\frac{1}{288\Omega}, \frac{\lambda_2 - \lambda_1'}{2\Omega}\right\}.$$

Indeed, otherwise for $x \in \partial W_{1/2}(x(F))$, by virtue of (3.2.23), we would have

$$\Phi(x) - \Phi(x(F)) > \tfrac{17}{192} - \tfrac{19}{576} = \tfrac{1}{18},$$

and (3.2.20) would lead to $x_i \in W_{1/2}(x(F))$. Hence by Theorem 2.1.1 we would have

$$\| Dg \|_{x_i,F}^* \leq 2 \| Dg \|_F^* \leq 2\Omega$$

or

$$\lambda(F, x_i) \le \lambda(\Phi, x_i) + t_i \parallel Dg \parallel^*_{x_i,F} \le \lambda'_1 + 2t_i\Omega \le \lambda_2,$$

which contradicts the assumption that $i < i^* - 1$.

The established lower bound for $t_i$, $i < i^*$, combined with (3.2.10), leads to (3.2.19).

It remains to note that (3.2.17) is equivalent to (3.2.13); relation (3.2.17), by (2.2.17), (3.2.6), implies (3.2.18). □

Note that the logarithmic term in (3.2.19) involves the quantity $(1 - \pi_{x(F)}(w))$, which is responsible for the quality of the starting point (the less it is; i.e., the "closer" to the boundary $w$ is, the worse the efficiency estimate (3.2.19)). Note that this quantity can be evaluated via the *asymmetry coefficient* of $G$ with respect to $w$. This coefficient $\alpha(G : w)$ is, by definition, the largest $\alpha$ such that every chord of $G$ passing through $w$ is divided by this point in the ratio not less than $\alpha$,

$$\alpha(G : w) = \max\{\alpha \mid w + \alpha(w - G) \subset G\}.$$

Evidently, $1 - \pi_x(w) \ge \alpha(G : w)$ for each $x \in G$, so that (3.2.19) implies the estimate

$$i^* \le 1 + \frac{\lambda_1 + \vartheta^{1/2}}{\lambda_1 - \lambda'_1}\left(\ln \frac{31}{\lambda_2 - \lambda'_i} + \ln \frac{\vartheta}{\alpha(G : w)}\right). \tag{3.2.24}$$

### 3.2.4 Main stage

The main stage consists in minimizing the function $f$. At this stage, we follow the trajectory of minimizers of the family

$$\mathcal{F}^{(2)} = \mathcal{F}(F, f) \equiv \{\operatorname{int} G, F_t^{(2)}(x) = tf(x) + F(x), E\}_{t>0}$$

along a sequence $t_i \to \infty$. Note that this family is strongly self-concordant (Proposition 3.2.2) with the parameters

$$\alpha(t) = \frac{1}{(1+\beta)^2},$$

$$\rho_\nu(\mathcal{F}^{(2)}; t, t') = \left(1 + (1+\beta)\frac{\vartheta^{1/2}}{\nu}\right)\left|\ln \frac{t}{t'}\right|. \tag{3.2.25}$$

Let

$$t_0 = \frac{\lambda_3 - (1+\beta)\lambda(F, u)}{(1+\beta)/ \parallel Df(u) \parallel^*_{u,F}}. \tag{3.2.26}$$

We assume that $Df(u) \ne 0$, since otherwise $u$ is a solution to $(f)$. We set

$$t_i = \kappa_2^i t_0, \; i \ge 0; \quad \kappa_2 = \exp\left\{\frac{\lambda_3 - \lambda'_3}{\lambda_3 + (1+\beta)\vartheta^{1/2}}\right\}, \tag{3.2.27}$$

so that

$$\rho_{\lambda_3}(\mathcal{F}^{(2)}; t_i, t_{i+1}) = 1 - \frac{\lambda_3}{\lambda_3'}, \qquad i \geq 0 \tag{3.2.28}$$

and then compute the points $x_i$ given by

$$x_{-1} = u; \qquad x_i = x^*(F_{t_i}^{(2)}, x_{i-1}), \quad i \geq 0. \tag{3.2.29}$$

These points $x_i$ are regarded as approximate solutions produced by the barrier method.

**Proposition 3.2.4** (i) *The main stage is well defined, i.e., $x_i$, $i \geq -1$, are well defined and belong to* $\operatorname{int} G$, *and the following inequalities hold:*

$$(\mathcal{I}_i): \qquad \lambda(F_{t_i}^{(2)}, x_{i-1}) \leq \lambda_3, \tag{3.2.30}$$

$$(\mathcal{J}_i): \qquad \lambda(F_{t_i}^{(2)}, x_i) \leq \lambda_3'. \tag{3.2.31}$$

(ii) *For every $i \geq 0$, we have*

$$\begin{aligned} &f(x_i) - f(x^*) \\ &\leq \frac{9(1+\beta)}{\lambda_3 - (1+\beta)\lambda_2}\left\{2\vartheta + \frac{\zeta(\lambda_3')}{2(1+\beta)^2}\right\} \\ &\cdot V_F(f) \exp\left\{-\frac{\lambda_3 - \lambda_3'}{\lambda_3 + (1+\beta)\vartheta^{1/2}} i\right\}, \end{aligned} \tag{3.2.32}$$

*where $x^*$ is a minimizer of $f$ over $G$ and*

$$V_F(f) = \sup\{f(x) \mid x \in W_{1/2}(x(F))\} - \inf\{f(x) \mid x \in W_{1/2}(x(F))\}. \tag{3.2.33}$$

**Proof.** $1^0$. By (3.2.5), we have $\lambda_3(1+\beta)^{-1} > \lambda_2 \geq \lambda(F, u)$ (the latter relation is a consequence of (3.2.17)), so that $t_0 > 0$. Let us verify that

$$t_0 \geq \frac{\lambda_3 - (1+\beta)\lambda_2}{9(1+\beta)V_F(f)}. \tag{3.2.34}$$

Indeed, by (3.2.18), we have $u \in W_{1/3}(x(F))$; thus, by Theorem 2.1.1,

$$\| e \|_{u,F} \leq \frac{\| e \|_F}{1 - \frac{1}{3}} = \frac{3}{2} \| e \|_F, \qquad e \in E.$$

Hence the ellipsoid $W_{1/9}(u)$ is contained in $W_{1/2}(x(F))$, which leads to $\| Df(u) \|_{u,F}^* \leq 9V_F(f)$, and (3.2.34) follows.

$2^0$. Let us prove $(\mathcal{I}.)$, $(\mathcal{J}.)$ by induction. We have

$$\begin{aligned} \lambda(F_{t_0}^{(2)}, u) &= \frac{1}{\alpha^{1/2}(t_0)} \sup\{\mid DF_{t_0}^{(2)}(u)[h] \mid \{D^2F_{t_0}^{(2)}(u)[h,h]\}^{-1/2} \mid h \neq 0\} \\ &\leq (1+\beta)\sup\{\mid DF_{t_0}^{(2)}(u)[h] \mid \{D^2F(u)[h,h]\}^{-1/2} \mid h \neq 0\} \\ &\leq (1+\beta)\sup\{\{\mid DF(u)[h] \mid + t_0 \mid Df(u)[h] \mid\} \langle h,h\rangle_{u,F}^{-1/2} \mid h \neq 0\} \\ &\leq (1+\beta)\{\lambda(F,u) + t_0 \| Df(u) \|_{u,F}^*\} = \lambda_3 \end{aligned}$$

(the latter equality holds in view of (3.2.26)), so that $(\mathcal{I}_0)$ holds. Assume that $k \geq 0$ is such that relations $(\mathcal{I}_i)$ hold for $0 \leq i \leq k$ and that relations $(\mathcal{J}_i)$ hold for $0 \leq i < k$. Relation $(\mathcal{I}_k)$, by Theorem 2.2.2(ii), implies that $x_k$ is well defined, belongs to $\operatorname{int} G$, and that $(\mathcal{J}_k)$ holds. Furthermore, relation $(\mathcal{J}_k)$, combined with Theorem 3.1.1 and (3.2.28), leads to $(\mathcal{I}_{k+1})$. Thereby (i) is proved.

$3^0$. Let us prove (ii). Let us fix $i$ and denote $t_i$ by $t$, $F^{(2)}_{t_i}$ by $\Phi$, $x_i$ by $z$. In view of $(\mathcal{J}_i)$, we have $\lambda(\Phi, z) \leq \lambda_3' < 1$; moreover, $\Phi \in S_a^+(\operatorname{int} G, E)$ with $a = (1+\beta)^{-2}$. The function $\Phi$ attains its minimum over $\operatorname{int} G$ (Theorem 2.2.2) at certain point $v$, and (2.2.15), (2.2.17), and (3.2.6) imply that

$$\text{(3.2.35)} \quad \Phi(z) - \Phi(v) \leq \frac{a}{2}\zeta(\lambda_3') \equiv \tau, \qquad D^2\Phi(v)[z-v, z-v] < a.$$

Let us verify that

$$\text{(3.2.36)} \qquad DF(v)[z-v] \geq -\vartheta.$$

Indeed, by virtue of the second relation in (3.2.35) and Theorem 2.1.1(ii), the point $z' = v + (v - z)$ belongs to $\operatorname{int} G$, and (2.3.2), as applied to $x = v$, $y = z'$, implies (3.2.36).

Let $x^*$ be the minimizer of $f$ over $G$ (the minimizer does exist, since $G$ is bounded and $f$ is lower semicontinuous on $G$). We have

$$f(x^*) \geq f(v) + Df(v)[x^* - v];$$

furthermore, by definition of $v$, we have

$$Df(v)[h] = -\frac{1}{t}DF(v)[h],$$

so that

$$f(x^*) \geq f(v) - \frac{1}{t}DF(v)[x^* - v] \geq f(v) - \frac{\vartheta}{t}$$

(the latter inequality holds by virtue of (2.3.2)). At the same time, by (3.2.35), we have

$$f(z) \leq f(v) + \frac{1}{t}\{F(v) - F(z) + \tau\} \leq f(v) + \frac{1}{t}\{DF(v)[v-z] + \tau\} \leq f(v) + \frac{\vartheta + \tau}{t}$$

(the latter inequality holds by virtue of (3.2.36)). The above inequalities imply that

$$\text{(3.2.37)} \qquad f(z) \leq f(x^*) + \frac{2\vartheta + \tau}{t}.$$

Relations (3.2.37) and (3.2.34) prove (3.2.32). $\square$

### 3.2.5 Efficiency estimate

As a straightforward consequence of the results stated by Proposition 3.2.3 and Proposition 3.2.4, we come to the following main proposition (for simplicity, we restrict ourselves to the case of quadratic $f$).

**Theorem 3.2.1** *Let $G$ be a bounded closed convex domain in $E$, let $F$ be a $\vartheta$-self-concordant barrier for $G$, and let $f$ be a convex quadratic form (i.e., a 0-compatible with $F$ function). Also, let $w \in \operatorname{int} G$ and let $\lambda_1$, $\lambda_1'$, $\lambda_2$, $\lambda_3$, $\lambda_3'$ satisfy (3.2.5), (3.2.6) for $\beta = 0$. Let the barrier method specified by the parameters $\beta = 0$, $\lambda_1$, $\lambda_1'$, $\lambda_2$, $\lambda_3$, $\lambda_3'$ and by the starting point $w$ be applied to problem $(f)$. Then, for each $\varepsilon \in (0,1)$, the total number of iterations of the preliminary and the main stages $N(\varepsilon)$ required to obtain an approximate solution $x_\varepsilon \in \operatorname{int} G$ such that*

$$f(x_\varepsilon) - \min_G f \le \varepsilon V_F(f)$$

*satisfies the inequality*

$$(3.2.38)\quad N(\varepsilon) \le O(1)\vartheta^{1/2} \ln \frac{2\vartheta}{\varepsilon(1-\pi_{x(F)}(w))} \le O(1)\vartheta^{1/2} \ln \frac{2\theta}{\varepsilon\alpha(G:w)}$$

*(the constant factor $O(1)$ depends only on the parameters $\lambda_1$, $\lambda_1'$, $\lambda_2$, $\lambda_3$, $\lambda_3'$).*

*Each of the iterations reduces to a single Newton step, as applied to a convex combination of $f$ and $F$ (or that of $F$ and a linear form).*

A theoretically good (approximately optimal for large $\vartheta$) choice of the parameters $\lambda_1$, $\lambda'$, $\lambda_2$, $\lambda_3$, $\lambda'$ in the case of $\beta = 0$ is

$$\lambda_1 = \lambda_3 = 0.193; \quad \lambda_1' = \lambda_3' = \lambda^+(\lambda_1) = 0.057; \quad \lambda_2 = 0.150.$$

Under this choice, for sufficiently large $\vartheta$, the principal term in the asymptotic $(\varepsilon \to 0)$ representation of the right-hand side of (3.2.38) is $7.36\vartheta^{1/2}\ln(1/\varepsilon)$.

### 3.2.6 Large-step strategy

Consider the case of a standard problem

$$(3.2.39)\qquad c^T x \to \min \mid x \in G,$$

$G$ being a closed and bounded convex domain in $E = \mathbf{R}^n$, and let $F$ be a $\vartheta$-self-concordant barrier for $G$. All we are interested in when solving the problem by an associated with $F$ path-following method is maintaining closeness to the path, i.e., to guarantee the inequality of the type

$$(3.2.40)\qquad \lambda(F_{t_i}, x_i) \le \bar\lambda$$

along a sequence $\{t_i\}$ of values of the penalty parameter $t$ varying in certain ratio $\kappa$; here the *path tolerance* $\bar\lambda$ is a given (and not too large) absolute constant, and $F_t(x) = td^T x + F(x)$ ($d = c$ at the main stage, and $-d$ is the gradient of $F$ at the starting point at the preliminary stage). Now "not too large $\bar{\lambda}$" in our previous context was defined, along with other parameters of the method, by (3.2.5), (3.2.6); since we now deal with linear objective only, in what follows, it is sufficient to assume that $\bar{\lambda} < 1$. The above considerations demonstrate that, under proper choice of absolute constant $\bar{\lambda}$, we can ensure the ratio

$$\kappa_* = 1 + \frac{O(1)}{\vartheta^{1/2}} \tag{3.2.41}$$

by a single Newton step per a step in $t$. In the large-scale case ($\vartheta$ is large), (3.2.41) results in "small" steps in $t$, and the method will certainly be slow. From a practical viewpoint, it seems more attractive to use a larger rate of varying the penalty parameter, say, to choose $\kappa > 1$ as an absolute constant, and to use several Newton steps in $x$ per a step in $t$. In this section, we demonstrate that the *worst-case* number of Newton steps per step in $t$ in this scheme is bounded from above by $O(\vartheta)$ with the constant factor in $O(\cdot)$ depending on $\kappa$ and $\bar{\lambda}$ only.

**Proposition 3.2.5** *Let $F$ be a $\vartheta$-self-concordant barrier for a closed and bounded convex domain $G$, let $\bar{\lambda} \in (0,1)$, let $d$ be a vector, and let $\kappa > 0$. Consider the family*

$$F_t(x) = td^T x + F(x) \tag{3.2.42}$$

*and let $u \in \operatorname{int} G$ and $t > 0$ be such that*

$$\lambda(F_t, u) \le \bar{\lambda}. \tag{3.2.43}$$

*Set*

$$T = \kappa t \tag{3.2.44}$$

*and let $x_i$ be defined as*

$$x_0 = u, \qquad x_{i+1} = x_i - \frac{[F''(x_i)]^{-1}\nabla_x F_T(x_i)}{1 + \lambda(F_T, x_i)}. \tag{3.2.45}$$

*Let $i^*$ be the smallest value of $i$ for which*

$$\lambda(F_T, x_i) \le \bar{\lambda}. \tag{3.2.46}$$

*Then*

$$i^* \le O(1)(\mid \kappa - 1 \mid \vartheta^{1/2} + \vartheta(\kappa - 1 - \ln \kappa)) + 1,$$

*with $O(1)$ depending on $\bar{\lambda}$ only.*

**Proof.** Let

$$x(\tau) = \operatorname{argmin}\{F_\tau(x) \mid x \in \operatorname{int} G\},$$

so that

$$F'(x(\tau)) = -\tau d. \tag{3.2.47}$$

**Lemma 3.2.1** *We have*

$$F_T(x(t)) - F_T(x(T)) \le \vartheta\left(\frac{T}{t} - 1 - \ln\frac{T}{t}\right). \tag{3.2.48}$$

**Proof.** Let

$$f(\tau) = F_\tau(x(t)) - F_\tau(x(\tau)).$$

Since $G$ is bounded, $F''$ is nondegenerate (Proposition 2.3.2(iii)), whence, by the implicit function theorem as applied to (3.2.47), $x(\tau)$ is continuously differentiable and

$$x'(\tau) = -[F''(x(\tau))]^{-1}d = \frac{1}{\tau}[F''(x(\tau))]^{-1}F'(x(\tau)). \tag{3.2.49}$$

We have

$$f'(\tau) = d^T x(t) - d^T x(\tau) - \tau d^T x'(\tau) - (F'(x(\tau)))^T x'(\tau) = d^T x(t) - d^T x(\tau)$$

(the latter equality follows from (3.2.47)), whence

$$f'(t) = 0 \tag{3.2.50}$$

and

$$\begin{aligned} f''(\tau) &= -d^T x'(\tau) = \frac{1}{\tau} d^T [F''(x(\tau))]^{-1} F'(x(\tau)) \\ &= \frac{1}{\tau^2}(F'(x(\tau)))^T [F''(x(\tau))]^{-1} F'(x(\tau)) = \frac{1}{\tau^2}\lambda^2(F, x(\tau)) \end{aligned}$$

(we sequentially used (3.2.49) and (3.2.47)). Since $F$ is a $\vartheta$-self-concordant barrier, we conclude that

$$0 \le f''(\tau) \le \frac{\vartheta}{\tau^2},$$

which, combined with (3.2.50) and evident relation $f(t) = 0$, implies relation (3.2.48). $\square$

From (3.2.43) and items (iv) and (i) of §2.2.4, it follows that

$$\| u - x(t) \|_{x(t),F} \le O(1) \tag{3.2.51}$$

(here and henceforth in the proof, all $O(1)$ depend on $\bar\lambda$ only), whence

$$\begin{aligned} F_T(u) - F_T(x(t)) &= (T - t)d^T(u - x(t)) + F_t(u) - F_t(x(t)) \\ &\le | T - t | \| d \|^*_{x(t),F} \| u - x(t) \|_{x(t),F} \\ &\quad + F_t(u) - F_t(x(t)), \end{aligned} \tag{3.2.52}$$

where $\| \cdot \|^*_{x(t),F}$ is the norm conjugate to $\| \cdot \|_{x(t),F}$,

$$\| p \|^*_{x(t),F} = \sup\{p^T h \mid \| h \|_{x(t),F} \le 1\}.$$

By definition of the Newton decrement, we have

$$\lambda(F_t, x) = \| \nabla F_t(x) \|^*_{x,F} .$$

We have $0 = \nabla F(x(t)) + td$, and, at the same time,

$$\vartheta^{1/2} \geq \lambda(F, x(t)) = \| \nabla F(x(t)) \|^*_{x(t),F},$$

whence

$$\| d \|^*_{x(t),F} \leq \frac{\vartheta^{1/2}}{t}. \tag{3.2.53}$$

Furthermore, from (3.2.43) and item (iv) of §2.2.4, it follows that

$$F_t(u) - F_t(x(t)) \leq O(1). \tag{3.2.54}$$

Relations (3.2.51)–(3.2.54) imply that

$$\begin{aligned} F_T(u) - F_T(x(t)) &= F_t(u) - F_t(x(t)) + (T - t)d^T(u - x(t)) \\ &\leq O(1) + | T - t | \| d \|^*_{x(t),F} \| u - x(t) \|_{x(t),F} \\ &\leq O(1) \left( \left| \frac{T}{t} - 1 \right| \vartheta^{1/2} + 1 \right), \end{aligned}$$

which, combined with (3.2.48), implies that

$$F_T(u) - F_T(x(T)) \leq O(1) \left( \left| \frac{T}{t} - 1 \right| \vartheta^{1/2} + 1 \right) + \vartheta \left( \frac{T}{t} - 1 - \ln \frac{T}{t} \right). \tag{3.2.55}$$

Now, from Proposition 2.2.2, it follows that, while (3.2.46) is not satisfied, iterations (3.2.45) decrease $F_T$ at least by $O(1)$, so that (3.2.55) means that the number of these iterations before termination is bounded from above by

$$O(1)(| \kappa - 1 | \vartheta^{1/2} + \vartheta(\kappa - 1 - \ln \kappa) + 1). \qquad \Box$$

Now consider the method for solving (3.2.39) in which the penalty parameter $t$ is varied in a ratio $\kappa$, $| \kappa - 1 | \geq 1 + \vartheta^{1/2}$ and a step in $t$ is accompanied by the number of steps (3.2.45) sufficient to restore closeness to the trajectory (i.e., to provide (3.2.40)). Following the same line of argument as in the proofs of Proposition 3.2.3 and Proposition 3.2.4, we can derive from Proposition 3.2.5 that the method finds an $\varepsilon$-solution to (3.2.39) at the total cost of no more than

$$O(1)\vartheta \ln \left( \frac{2\vartheta}{\varepsilon(1 - \pi_{x(F)}(w))} \right) \tag{3.2.56}$$

Newton steps. The constant factor in the latter $O(\cdot)$ depends on $\kappa$ and $\bar{\lambda}$ only in such a way that, if $\kappa = 1 \pm \vartheta^{-1/2}$, then $O(1)$ is of order of $\vartheta^{-1/2}$, and (3.2.56) becomes the same as (3.2.38); if $\kappa$ and $\bar{\lambda} < 1$ are absolute constants, then this factor is an absolute constant. Thus, from the viewpoint of the worst-case behavior, the "optimistic" strategy, where $\kappa$ is chosen as an absolute constant, is $O(\vartheta^{1/2})$ times worse than the strategy with "small" rate (3.2.41) of updating the penalty. Nevertheless, the results corresponding to the case of "large" rate of updating $t$ are useful: They demonstrate that the corresponding strategy still results in a procedure with polynomial worst-case behaviour (i.e., in procedure converging with an objective-independent linear rate). On the other hand, this strategy seems to be more flexible than the basic (and optimal in the worst case) strategy (3.2.41). Indeed, we could hope that standard technique like line search in the Newton direction would restore the closeness to the path at a significantly less than $O(\vartheta)$ number of Newton steps, so that the total effort would be less than that one that is *for sure* required by the basic strategy.

## 3.3 Method of centers

In this section, we describe another path-following method generated by a $\vartheta$-self-concordant barrier $F$ for a closed convex *bounded* domain $G$ in $E$. In what follows, we assume that $f$ is a convex quadratic form (so that the method can solve at least all standard problems).

The method is associated with the following family. Let us choose a constant $\zeta \geq 1$ and set

$$t^* = \min\{f(x) \mid x \in G\}, \qquad \Delta = (t^*, +\infty),$$

$$Q_t = \{x \in \operatorname{int} G \mid f(x) < t\},$$

$$F_t(x) = \zeta \ln(1/(t - f(x))) + F(x) : Q_t \to \mathbf{R}, \qquad (t \in \Delta).$$

Thus, we have defined the family

$$\mathcal{F}^* = \mathcal{F}^*(F, f) = \{Q_t, F_t, E\}_{t \in \Delta}.$$

Note that the path of minimizers for this family is the same path as for the $F$-generated barrier family from the previous section, but the parameterization of the path differs from that one considered in §3.2.

It turns out that the family is self-concordant, as demonstrated in the following theorem.

**Theorem 3.3.1** *For each $\lambda \in (0, \lambda_*)$ and $\kappa' \in (\lambda, \lambda_*)$, and for $\alpha$, $\gamma$, $\mu$, $\xi$ $\eta$, $\kappa$ chosen as*

$$\alpha(t) \equiv 1; \quad \gamma(t) \equiv 1; \quad \mu(t) \equiv 1; \quad \kappa = \lambda;$$

$$\xi(t) = \frac{\zeta^{1/2}\Omega}{t - t^*}; \qquad \eta(t) = \frac{\Omega}{t - t^*}, \tag{3.3.1}$$

*where*[2]

$$\Omega = 1 + \frac{\delta + \vartheta}{\zeta}, \qquad \delta = \frac{\kappa'}{1 - \beta}\left(1 + 3\vartheta^* + \frac{\beta}{1 - \beta}\right),$$

[2]Recall that $\omega(\lambda) = 1 - (1 - 3\lambda)^{1/3}$.

$$\vartheta^* = \vartheta + \zeta, \qquad \beta = \omega(\kappa'), \tag{3.3.2}$$

*the family $\mathcal{F}^*(F, f)$ is self-concordant with the above parameters.*
*In particular, we have*

$$\psi(t) \equiv 1, \qquad t \in \Delta, \tag{3.3.3}$$

*and*

$$\rho_\lambda(\mathcal{F}^*; t, t') = \Omega\left(1 + \frac{\zeta^{1/2}}{\lambda}\right) \mid \ln \frac{t - t^*}{t' - t^*} \mid . \tag{3.3.4}$$

*Moreover, the following implication holds:*

$$t \in \Delta, \quad x \in Q_t, \quad \lambda(F_t, x) \le \lambda \Rightarrow \frac{1}{t - f(x)} \le \frac{\Omega}{t - t^*}. \tag{3.3.5}$$

**Proof.** Let us verify that relations ($\Sigma$.1), ($\Sigma$.2), ($\Sigma$.3) hold. Relation ($\Sigma$.1) is evident.

Let us verify that the function

$$f_t(x) = \ln \frac{1}{t - f(x)}$$

for each $t \in \Delta$ is, as a function of $x$, a 1-self-concordant barrier for the set $R_t = \{x \in E \mid f(x) \le t\}$. This is an immediate corollary of the following result.

**Lemma 3.3.1** *Let $s$ be a convex quadratic form on $E$ such that the set $Q = \{x \mid s(x) < 0\}$ is nonempty. Then the function*

$$S(x) = -\ln(-s(x))$$

*is a 1-self-concordant barrier for the set* cl $Q$.

**Proof of lemma**. We should prove that $S$ is strongly 1-self-concordant on $Q$ and $\lambda(S, x) \le 1$. Clearly, $S$ is $C^\infty$ smooth and tends to infinity along every sequence of points of $Q$ converging to a boundary point of this set; so it remains to verify 1-self-concordance of $S$ and to evaluate the Newton decrement. For $x \in Q, h \in E$, we have

$$DS(x)[h] = -\frac{Ds(x)[h]}{s(x)},$$

$$D^2S(x)[h, h] = -\frac{D^2s(x)[h, h]}{s(x)} + \left(\frac{Ds(x)[h]}{s(x)}\right)^2,$$

$$D^3S(x)[h, h, h] = \frac{3Ds(x)[h]D^2s(x)[h, h]}{s(x)^2} - 2\left(\frac{Ds(x)[h]}{s(x)}\right)^3;$$

thus, under the notation $q = -s^{-1}(x)Ds(x)[h], p = (-s^{-1}(x)D^2s(x))^{1/2}$ (recall that $s(\cdot)$ is negative on $Q$), we obtain

$$DS(x)[h] = q, \quad D^2S(x)[h,h] = q^2 + p^2, \quad D^3S(x)[h,h,h] = 3p^2q + 2q^3,$$

and the desired relations between the derivatives of $S$ follows immediately from the evident inequalities

$$\begin{aligned} 2(p^2+q^2)^{3/2} &= 2 \mid q \mid^3 \left(1 + \left(\frac{p}{\mid q \mid}\right)^2\right)^{3/2} \\ &\geq 2 \mid q \mid^3 \left(1 + \frac{3}{2}\left(\frac{p}{\mid q \mid}\right)^2\right) \\ &= 2 \mid q \mid^3 + 3p^2 \mid q \mid \geq \mid 3p^2q + 2q^3 \mid \end{aligned}$$

(the resulting inequality in the case of $q = 0$, when the latter computation is not valid, evidently is also true) and $q^2 \leq p^2 + q^2$. □

Now we can continue the proof of the theorem.

Since $\zeta \geq 1$ and $f_t$ is a 1-self-concordant barrier for $R_t$, the function $\zeta f_t$ is a $\zeta$-self-concordant barrier for the latter set. Therefore, by virtue of Proposition 2.3.1(ii), $F_t$ is a $\vartheta^*$-self-concordant barrier for $R_t$. In particular, $F_t \in S_1^+(Q_t, E)$, as it is required in ($\Sigma$.2) for $\alpha$ chosen in accordance with (3.3.1).

Let us verify ($\Sigma$.3). Let

$$X^+(\kappa) = \{(t,x) \in Q_* \mid \lambda(F_t, x) < \kappa'\}$$

(we henceforth use the notation from §3.1.1). It is clear that $X^+(\kappa)$ is a neighbourhood of $X_*(\kappa)$ in $Q_*$.

Let us verify that the set $X_*(\kappa)$ is closed in $E_\Delta$. Indeed, let $(t_i, x_i) \in X_*(\kappa)$ and $(t_i, x_i) \overset{i\to\infty}{\to} (t, x)$, where $t \in \Delta$. By Theorem 2.2.2(iii) and in view of the fact that $F_\tau$ is strongly 1-self-concordant on $Q_\tau$, we have

$$F_{t_i}(x_i) \leq \phi(t_i) + c$$

for certain constant $c$, where

$$\phi(\tau) = \min\{F_\tau(x) \mid x \in Q_\tau\}, \qquad \tau \in \Delta.$$

The function $\phi$ clearly is bounded along the sequence $\{t_i\}$ (since this sequence converges to a point from $\Delta$), so that $\{F_{t_i}(x_i)\}$ is bounded. The latter fact, in view of definition of $F_\tau$, implies the inclusion $x \in Q_t$, or $(t,x) \in Q_*$. We see that the closure of $X_*(\kappa)$ in $E_\Delta$ is contained in $Q_*$; since $X_*(\kappa)$ is evidently closed in $Q_*$, $X_*(\kappa)$ is closed in $E_\Delta$.

It remains to verify that, under the choice of the parameters in accordance with (3.3.1), relations (3.1.1), (3.1.2) hold for our $X^+(\kappa)$.

Let us fix $(t,x) \in X^+(\kappa)$. Then

$$\lambda(F_t, x) < \kappa'. \tag{3.3.6}$$

Let

$$x^* = \operatorname{argmin}\{F_t(y) \mid y \in Q_t\}$$

(the existence and the uniqueness of $x^*$ follow from Proposition 2.3.2(ii), since $Q_t$ is bounded and $F_t$ is strongly 1-self-concordant on $Q_t$; note that, by the same reason, $D^2F_t$ is nondegenerate on $Q_t$). Let us introduce an Euclidean structure on $E$ by the inner product

$$\langle h, s\rangle = D^2F_t(x^*)[h,s].$$

Denote the corresponding norm by $\| \cdot \|$. Let $W$ be the open unit ball centered at $x^*$. By Proposition 2.3.2(ii) and since $F_t$ is a $\vartheta^*$-self-concordant barrier for cl $Q_t$, we have

$$W \subset Q_t \subset \{y \mid \| y - x^* \| \le (1 + 3\vartheta^*)\}. \tag{3.3.7}$$

Furthermore, in view of (3.3.6) and Theorem 2.2.2(iii), we have ($\beta = \omega(\kappa') < 1$)

$$D^2F_t(x)[x^* - x, x^* - x] < \beta^2, \qquad \| x - x^* \| \le \frac{\beta}{1-\beta}, \tag{3.3.8}$$

whence, by Theorem 2.1.1,

$$D^2F_t(x^*)[h,h] \ge (1-\beta)^2 D^2F_t(x)[h,h]. \tag{3.3.9}$$

In view of (3.3.6) and (3.3.9), we have

$$\| \nabla F_t(x) \| \le \frac{\kappa'}{1-\beta} \tag{3.3.10}$$

($\nabla$ denotes the gradient with respect to the Euclidean structure $\langle \cdot, \cdot \rangle$). Let $u^*$ be the minimizer of $f$ on cl $Q_t$ (or, which is the same, on $G$). Then, taking into account (3.3.7) and (3.3.6), we obtain

$$\begin{aligned}\delta = \frac{\kappa'}{1-\beta}\left(1 + 3\vartheta^* + \frac{\beta}{1-\beta}\right) &\ge \langle \nabla F_t(x), x - u^*\rangle \\ &= \frac{\zeta}{t - f(x)}\langle \nabla f(x), x - u^*\rangle + \langle \nabla F(x), x - u^*\rangle \ge \frac{\zeta(f(x) - t^*)}{t - f(x)} - \vartheta.\end{aligned}$$

Thus,

$$\frac{f(x) - t^*}{t - f(x)} \le \frac{\delta + \vartheta}{\zeta} = \Omega - 1 \tag{3.3.11}$$

or

$$\frac{1}{t - f(x)} \le \frac{\Omega}{t - t^*}. \tag{3.3.12}$$

Thereby (3.3.5) is proved.

Furthermore, we have

$$\mid \{DF_t(x)[h]\}'_t \mid = \frac{\zeta \mid Df_t(x)[h] \mid}{t - f(x)} \leq \frac{\zeta (D^2 f_t(x)[h,h])^{1/2}}{t - f(x)}$$
$$\leq \frac{\zeta^{1/2}(D^2 F_t(x)[h,h])^{1/2}}{t - f(x)},$$

which, combined with (3.3.1) and (3.3.12), implies (3.1.1) (we have taken into account that $f_t \in \mathcal{B}(\{x \mid f(x) \leq t\}, 1)$).

By the same arguments, we have

$$\mid \{D^2 F_t(x)[h,h]\}'_t \mid \leq \frac{2\zeta \mid D^2 f_t(x)[h,h] \mid}{t - f(x)} \leq \frac{2D^2 F_t(x)[h,h]}{t - f(x)},$$

which, in view of (3.3.1) and (3.3.12), implies (3.1.2). □

Now we can present the *generated by F method of centers* for problem $(f)$ with convex quadratic $f$. We restrict ourselves to the description of the main stage of the method.

Let us fix constants $\lambda$, $\lambda'$ such that

(3.3.13) $$\lambda^+(\lambda) \leq \lambda' < \lambda < \lambda_*.$$

Assume that we are given $t_0 \in \Delta$ and $x_{-1} \in Q_{t_0}$ satisfying

(3.3.14) $$\lambda(F_{t_0}, x_{-1}) \leq \lambda.$$

Let us define a sequence of points $x_i$ and numbers $t_i \in \Delta$ as follows:

- Given $t_i$ and $x_{i-1}$ such that

(3.3.15) $$(\mathrm{I}_i): \quad t_i \in \Delta; \quad x_{i-1} \in Q_{t_i}; \quad \lambda(F_{t_i}, x_{i-1}) \leq \lambda,$$

we find a point $x_i$ satisfying the relations

(3.3.16) $$(\mathrm{J}_i): \quad x_i \in Q_{t_i}, \quad \lambda(F_{t_i}, x_i) \leq \lambda'.$$

Note that, under condition $(\mathrm{I}_i)$, the Newton iterate of $x_{i-1}$ (the Newton method is applied to $F_{t_i}$), i.e., the point $x_i = x^*(F_{t_i}, x_{i-1})$ satisfies $(\mathrm{J}_i)$ (see Theorem 2.2.2(ii));

- After $x_i$ is found, we define $t_{i+1}$ in accordance with the equation

(3.3.17) $$\ln \frac{t_{i+1} - f(x_i)}{t_i - f(x_i)} = -\frac{\lambda - \lambda'}{\Omega(\zeta^{1/2} + \lambda)} \equiv -\chi$$

(since $x_i \in Q_{t_i}$, it is clear that $t_i > f(x_i)$; thus, $t_{i+1}$ is well defined).

The following lemma demonstrates that the method is well defined.

**Lemma 3.3.2** *Relation* $(J_i)$ *implies the inclusion* $t_{i+1} \in \Delta$, *relation* $(I_{i+1})$ *and the inequality*

$$f(x_i) - t^* < t_{i+1} - t^* \le (t_i - t^*)e^{-\nu}, \tag{3.3.18}$$

*where*

$$\nu = \ln \frac{\Omega}{\Omega - 1 + e^{-\chi}}.$$

**Proof.** We have $t_i > f(x_i) \ge t^*$, so that $t_{i+1} \in \Delta$. We also have

$$1 > \frac{t_{i+1} - t^*}{t_i - t^*} > \frac{t_{i+1} - f(x_i)}{t_i - f(x_i)},$$

which, by (3.3.17) and Theorem 3.3.1, leads to

$$\rho_\lambda(\mathcal{F}^*; t_i, t_{i+1}) \le 1 - \frac{\lambda'}{\lambda}.$$

The latter relation, by Theorem 3.1.1, implies $(\mathrm{I}_{i+1})$.

To verify (3.3.18), we note that (3.3.17) and (3.3.5) imply that

$$1 - e^{-\chi} = \frac{t_i - t_{i+1}}{t_i - f(x_i)} \le \frac{\Omega(t_i - t_{i+1})}{t_i - t^*},$$

so that

$$\frac{t_i - t_{i+1}}{t_i - t^*} \ge \frac{1 - e^{-\chi}}{\Omega},$$

which leads to the second inequality in (3.3.18). The first inequality in (3.3.18) follows from the inclusion in $(\mathrm{I}_{i+1})$. □

Relation (3.3.18) leads to the accuracy estimate for the method

$$f(x_i) - \min_G f < (t_0 - t^*) \exp\left\{-\frac{i+1}{\nu}\right\}. \tag{3.3.19}$$

The value of $\nu$ depends on $\lambda$, $\lambda'$, and $\zeta$ only. Assume that $\lambda$ and $\lambda'$ are absolute constants satisfying (3.3.13). Then, maximizing $\nu$ over $\zeta$, we obtain

$$\zeta = O(\vartheta) \quad \text{and} \quad \nu = O\left(\frac{1}{\vartheta^{1/2}}\right),$$

with absolute constant factors in both of the $O(\cdot)$. Thus, the rate of convergence of the method under consideration is the same as that one of the associated with $F$ barrier method.

A reasonable choice of the parameters for the method is

$$\lambda = 0.136, \quad \lambda' = \lambda^+(\lambda) = 0.025, \quad \zeta = 3\vartheta.$$

For large $\vartheta$, this leads to

$$\nu \cong \frac{0.011}{\vartheta^{1/2}}.$$

To initialize the method of centers, we need a pair $(t_0, x_{-1})$ such that $x_{-1} \in Q_{t_0}$ and $\lambda(F_{t_0}, x_{-1}) \le \lambda$. To find such a pair, we can first approximate the $F$-center of $G$, for example, with the aid of the preliminary stage of the barrier method. The stage is terminated when a point $x$ with $\lambda(F, x) \le \lambda/2$ is found. This point can be taken as $x_{-1}$. Then, clearly, $\lambda(F_t, x_{-1}) \le \lambda$ for all sufficiently large $t$, which allows us to choose an appropriate $t_0$.

As we already have mentioned, the path of minimizers of the family $\mathcal{F}^*$ associated with the method of centers is, geometrically, the same path as in the previous section; nevertheless, the method of centers and the barrier-generated method use different parameterizations of the path and form different approximations to it.

## 3.4 Dual parallel trajectories method

The method we describe here is associated with the *homogeneous* self-concordant family. To define the family, let us introduce a special class of self-concordant functions.

Let $E^*$ be the space conjugate to $E$. Let $\mathcal{L}(E^*, \vartheta)$, $\vartheta \ge 1$ be the set of all 1-self-concordant on the whole $E^*$ functions $F^*$ such that

$$\vartheta^*(F^*) \equiv \sup\{D^2F^*(\phi)[\phi, \phi] \mid \phi \in E^*\} \le \vartheta. \tag{3.4.1}$$

As we have seen in §2.4, functions of that type with nondegenerate second-order derivative are precisely the Legendre transformations of $\vartheta$-self-concordant barriers for *bounded* closed convex domains in $E$.

The following statement is evident.

**Proposition 3.4.1** *Let $F_i^* \in \mathcal{L}(E^*, \vartheta_i)$, $p_i \ge 1$, $i = 1, 2$. Let $x(\phi)$ be an affine form on $E^*$ and let $\phi = A\psi$ be a homogeneous linear transformation from $H^*$ into $E^*$. Then*

(i) $p_1F_1^*(\phi) + p_2F_2^*(\phi) \in \mathcal{L}(E^*, p_1\vartheta_1 + p_2\vartheta_2)$;

(ii) $F_1^*(\phi) + x(\phi) \in \mathcal{L}(E^*, \vartheta_1)$;

(iii) $F_1^*(A\psi) \in \mathcal{L}(H^*, \vartheta_1)$.

Let $E_0^*$ be a hyperplane in $E^*$, codim $E_0^* = 1$; let $b \in E^*\backslash E_0^*$, $\Delta = (0, \infty)$ and $F^* \in \mathcal{L}(E^*, \vartheta)$. This data define a family of functions on $E_0^*$,

$$\mathcal{F}^{**} \equiv \mathcal{F}^{**}(F^*, E_0^*, b) = \{Q_t \equiv E_0^*,\ F_t(\phi) = F^*(t\phi + tb),\ E_0^*\}_{t\in\Delta}. \tag{3.4.2}$$

**Proposition 3.4.2** *For every collection of data $\{\vartheta \ge 1,\ F^* \in \mathcal{L}(E^*, \vartheta),\ E_0^*,\ b\}$ and, for each $\kappa \in (0, \lambda_*)$, the family $\mathcal{F}^{**}(F^*, E_0^*, b)$ is self-concordant with the parameters*

$$\alpha(t) \equiv 1; \quad \mu(t) \equiv t; \quad \gamma(t) \equiv t^2; \quad \xi(t) \equiv \eta(t) \equiv \vartheta^{1/2}/t; \quad \kappa. \tag{3.4.3}$$

*In particular, $\psi(\mathcal{F}^{**}, t) \equiv 1$ and*

$$\rho_\nu(\mathcal{F}^{**}; t, \tau) = \vartheta^{1/2}\left(1 + \frac{1}{\nu}\right)\left|\ln\frac{t}{\tau}\right|. \tag{3.4.4}$$

**Proof**. Relation ($\Sigma$.1) is evident; relation ($\Sigma$.2) immediately follows from the inclusion $F^* \in S_1(E^*, E^*)$ and Proposition 2.1.1(i). Let us verify ($\Sigma$.3); namely, let us prove that (3.1.1) and (3.1.2) hold for $X^+(\kappa) = Q_* \equiv \Delta \times E_0^*$. Indeed, let us fix $\psi \in E_0^*$, $t \in \Delta$ and let $\phi = t\psi + tb$. Then, for $\zeta \in E_0^*$, we have

$$DF_t(\psi)[\zeta] = tDF^*(\phi)[\zeta], \qquad D^2F_t(\psi)[\zeta,\zeta] = t^2D^2F^*(\phi)[\zeta,\zeta],$$

$$\begin{aligned} | \{DF_t(\psi)[\zeta]\}'_t - \{\ln(t)\}'_t DF_t(\psi)[\zeta] | &= | D^2F^*(\phi)[\phi,\zeta] | \\ &\le \frac{1}{t}(D^2F^*(\phi)[\phi,\phi])^{1/2}(t^2D^2F^*(\phi)[\zeta,\zeta])^{1/2} \\ &\le \frac{1}{t}\vartheta^{1/2}(D^2F_t(\psi)[\zeta,\zeta])^{1/2} \\ &= \xi(t)\alpha^{1/2}(t)(D^2F_t(\psi)[\zeta,\zeta])^{1/2}. \end{aligned}$$

Furthermore,

$$\begin{aligned} | \{D^2F_t(\psi)[\zeta,\zeta]\}'_t - \{\ln(t^2)\}'_t D^2F_t(\psi) | &= t \mid D^3F^*(\phi)[\zeta,\zeta,\phi] \mid \\ &\le 2tD^2F^*(\phi)[\zeta,\zeta](D^2F^*(\phi)[\phi,\phi])^{1/2} \\ &\le \frac{2\vartheta^{1/2}}{t}D^2F_t(\psi)[\zeta,\zeta] = 2\eta(t)D^2F_t(\psi)[\zeta,\zeta]. \end{aligned}$$

Inequalities (3.1.1) and (3.1.2) are proved. □

Now we describe the *dual parallel trajectories method.*

Let $G$ be a *bounded* closed convex domain in $\mathbf{R}^m$ and let $F$ be a $\vartheta$-self-concordant barrier for $G$. Assume that we know the $F$-center of the $G$ (i.e., the minimizer of $F$ on $\operatorname{int} G$); to simplify the description, let the center be 0. Let $A$, Rank $A = n$ be an $n \times m$ matrix and let $b \in \mathbf{R}^n$. The dual parallel trajectories method solves problems of the type

$$\tau \to \max \mid Ax = \tau b, \; x \in G. \tag{3.4.5}$$

If $G$ is a polytope, then (3.4.5) is an LP problem. Note that the assumption $F'(0) = 0$ is not a severe restriction, which is demonstrated by the following example:

$$\tau \to \max \mid x \in \mathbf{R}^m, \; Ax = \tau b, \; \| x \|_\infty \le 1, \tag{3.4.6}$$

where $G = \{x \in \mathbf{R}^m \mid \| x \|_\infty \le 1\}$, $\vartheta = m$ and $F(x) = -\sum_{i=1}^m \ln(1 - x_i^2)$ (the parameter of this barrier does equal to $m$; see Lemma 3.3.1 and Proposition 2.3.1). Note that (3.4.6) is, in a natural sense, a "universal" format for LP problems.

Without loss of generality, we can assume that

$$A(F''(0))^{-1}A^T = I_n, \tag{3.4.7}$$

(because the system $Ax = \tau b$ can be replaced by an equivalent system with the rows of the matrix being orthonormal with respect to the scalar product $e^T(F''(0))^{-1}h$).

Define the function on $\mathbf{R}^n \times \operatorname{int} G$,

$$L(\phi, x) = -F(x) + \phi^T Ax \tag{3.4.8}$$

and let

$$F^+(\phi) = \max\{L(\phi, x) \mid x \in \operatorname{int} G\}. \tag{3.4.9}$$

$F^+$ is obtained from the Legendre transformation of the barrier $F$ by a homogeneous linear transformation of argument, so that

$$F^+ \in \mathcal{L}(\mathbf{R}^n, m) \tag{3.4.10}$$

and $D^2F^+$ is nondegenerate. Note that, in the case of problem (3.4.6), $F^+$ has an explicit representation,

$$F^+(\phi) = \sum_{i=1}^{m} \left\{ \frac{(a_i^T\phi)^2}{1 + [1 + (a_i^T\phi)^2]^{1/2}} - \ln(1 + [1 + (a_i^T\phi)^2]^{1/2}) \right\}, \tag{3.4.11}$$

where $a_i$, $1 \le i \le m$ are the columns of $A$.

Denote the minimizer of $L(\phi, x)$ over $x \in \operatorname{int} G$ by $X(\phi)$ (this point is well defined). For problem (3.4.6), we have

$$X_i(\phi) = \frac{a_i^T\phi}{1 + [1 + (a_i^T\phi)^2]^{1/2}}, \qquad 1 \le i \le m. \tag{3.4.12}$$

Let $F^\tau(\phi) = F^+(\phi) - \tau b^T\phi$ for $\tau \ge 0$. We use the following result.

**Lemma 3.4.1** *Let $E^+ = \mathbf{R}^n$, $t > 0$, $\phi \in E_t^+ = \{\phi \in \mathbf{R}^n \mid \phi^T b = tb^T b\}$ and let $F_t$ be the restriction of $F^+$ onto $E_t^+$. Also, let $\lambda_\phi \equiv \lambda(F_t, \phi) < \frac{1}{3}$ be such that, for $\zeta_\phi \equiv \omega(\lambda_\phi)(1 - \omega(\lambda_\phi))^{-1}$ and $\xi_\phi = \zeta_\phi(1 - \zeta_\phi)^{-2}$, we have $\xi_\phi < 1$. Then*

(i) *The solution, $\tau_\phi$, to the problem*

$$\tau \to \max \mid \lambda(F^\tau, \phi) \le 1$$

*is well defined and positive, and $\lambda(F^{\tau_\phi}, \phi) = 1$;*

(ii) *The projection, $X^*(\phi)$ of the point $X(\phi)$ onto the plane $E' = \{x \in \mathbf{R}^m \mid Ax = \tau_\phi b\}$, orthogonal with respect to the Euclidean structure on $\mathbf{R}^m$ induced by the inner product $\langle h, e\rangle_\phi = D^2F(X(\phi))[h, e]$, belongs to $G$;*

(iii) *The inequality*

$$t^* - \tau_\phi \le \frac{\vartheta}{tb^T b} \tag{3.4.13}$$

*holds, where $t^*$ is the optimal value of problem (3.4.5).*

**Proof**. Note that $DF^+(0) = 0$ (since $DF(0) = 0$) and that $F^+$ is strongly convex (Theorem 2.4.2; recall that $A$ is a matrix of full row rank). Hence $F^+(\phi)$ tends to $\infty$ as $\| \phi \| \to \infty$, and the minimizers $\phi_t^*$ of $F_t$ are well defined. It is clear that

$$\nabla F^+(\phi_t^*) = \tau^*(t)b$$

for some $\tau^*(t) > 0$ and that the function $\tau^*(t)$ increases on the positive half-axis.

Denote $\phi_t^*$ by $\phi^*$, $\tau^*(t)$ by $\tau^*$, and let $\Phi(\psi) = F^{\tau^*}(\psi)$.

(i) Let us provide $E^+$ with the scalar product $\langle u, v\rangle = D^2\Phi(\phi^*)[u, v]$ and let $\| \cdot \|$ be the corresponding norm, $\Phi'(u)$, $\Phi''(u)$ be the corresponding gradient and Hessian of $\Phi$, respectively. By (3.4.10) and by virtue of the arguments from the beginning of the proof, $\Phi$ is strongly 1-self-concordant on $E^+$. Applying Theorem 2.2.2(iii) to the restriction $\Xi$ of the function $\Phi$ onto $E_t^+$ and taking into account that $\lambda(\Xi, \phi) = \lambda_\phi < \frac{1}{3}$, we obtain $\| \phi - \phi^* \| \leq \zeta_\phi$. Since $\zeta_\phi < 1$, we have

$$\| \Phi''(\phi^* + s(\phi - \phi^*)) \| \leq \frac{1}{(1 - s\zeta_\phi)^2}, \qquad 0 \leq s \leq 1$$

(Theorem 2.1.1). Moreover, $\Phi'(\phi^*) = 0$; thus,

$$\| \Phi'(\phi) \| \leq \frac{\zeta_\phi}{1 - \zeta_\phi}.$$

Applying Theorem 2.1.1, we obtain

$$\lambda(\Phi, \phi) \leq \frac{\zeta_\phi}{(1 - \zeta_\phi)^2}.$$

By assumption of the lemma, the latter quantity is less or equal 1; thus, $\tau_\phi$ is well defined and positive. Moreover, we have

$$\tau_\phi \geq \tau^*(t). \tag{3.4.14}$$

The latter equality in (i) is evident. Thereby (i) is proved.

(ii) Let $F^*(\cdot)$ be the Legendre transformation of $F$; thus, $F^+(\psi) = F^*(A^T\psi)$. Let

$$x = X(\phi), \quad \Xi(\psi) = F^+(\psi) - \tau_\phi b^T\psi, \quad \tau = \tau_\phi.$$

Replacing first- and second-order differentials by gradients and Hessians with respect to the standard Euclidean structure, we obtain, in view of the standard properties of the Legendre transformation,

$$\begin{aligned}
\{1 = \lambda(\Xi, \phi)\} &\Rightarrow \{(\Xi'(\phi))^T u \leq [(\Xi''(\phi)u)^T u]^{1/2},\ u \in \mathbf{R}^n\} \\
&\Rightarrow \{(Ax - \tau b)^T u \leq [(A(F^*)''(A^T\phi)A^T u)^T u]^{1/2},\ u \in \mathbf{R}^n\} \\
&\Longleftrightarrow \{(Ax - \tau b)^T u \leq [(A(F''(x))^{-1}A^T u)^T u]^{1/2},\ u \in \mathbf{R}^n\}.
\end{aligned} \tag{3.4.15}$$

Let $x^*$ be the projection involved into (ii). Then

$$F''(x)(x - x^*) = A^T u^*$$

for certain $u^* \in \mathbf{R}^n$, and

$$Ax - \tau b = A(x - x^*).$$

The latter inequality in (3.4.15), as applied to $u = u^*$, leads to

$$(A(x - x^*))^T u^* \leq [(A(x - x^*))^T u^*]^{1/2}.$$

In view of $F''(x)(x - x^*) = A^T u^*$, we obtain

$$(x - x^*)^T F''(x)(x - x^*) \leq [(x - x^*)^T F''(x)(x - x^*)]^{1/2},$$

whence $(x - x^*)^T F''(x)(x - x^*) \leq 1$. Thus, the ellipsoid

$$\{y \in \mathbf{R}^m \mid D^2 F(x)[y - x, y - x] \leq 1\}$$

contains $x^*$. This ellipsoid is contained in $G$ (Proposition 2.3.2(ii.1)); hence $x^* \in G$. Statement (ii) is proved.

(iii) By the standard duality arguments, $X(\phi^*) \equiv x^*$ belongs to the set

$$G' = \{x \in \operatorname{int} G \mid Ax = \tau^*(t) b\}$$

and minimizes $F$ over this set, so that

$$(3.4.16) \qquad (\forall w \in \mathbf{R}^m) : \; Aw = 0 \Rightarrow w^T F'(x^*) = 0; \; F'(x^*) = A^T \phi^*.$$

Let $y^*$ be a solution to (3.4.5) and let

$$u^* = \frac{t^*}{\tau^*(t)} x^*.$$

Then the premise in (3.4.16) holds for $w = y^* - u^*$, which leads to

$$(3.4.17) \qquad (u^* - x^*)^T F'(x^*) = (y^* - x^*)^T F'(x^*) \leq \vartheta$$

(the latter inequality holds by (2.3.2) and since $F$ is a $\vartheta$-self-concordant barrier for $G$). The equality in (3.4.16), combined with the evident relation

$$A(u^* - x^*) = (t^* - \tau^*(t)) b$$

and with (3.4.17), implies that $(t^* - \tau^*(t)) b^T \phi^* \leq \vartheta$, whence, in view of $\phi^* \in E_t^+$ (or, which is the same, in view of $b^T \phi^* = t b^T b$),

$$t^* - \tau^*(t) \leq \frac{\vartheta}{t b^T b}.$$

This inequality, combined with (3.4.14), proves (iii). $\square$

The above results lead to the following method for (3.4.5). Let us choose $\lambda > 0$ such that

$$0 < \lambda < \lambda_*, \qquad \omega(\lambda) < \tfrac{1}{2},$$

(3.4.18)
$$\frac{\omega(\lambda)(1-\omega(\lambda))}{(1-2\omega(\lambda))^2} < 1$$

and let $t_0$ be the solution of the equation

(3.4.19)
$$\frac{t\delta}{(1-t\delta)^2} = \lambda, \qquad \delta \equiv (b^T b)^{1/2},$$

belonging to $(0, 1/\delta)$.

Set $\phi_{-1} = t_0 b \in E^+ \equiv \mathbf{R}^n$ and let

(3.4.20)
$$t_i = \exp\left\{\frac{\lambda - \lambda^+(\lambda)}{(1+\lambda)m^{1/2}}\right\} t_{i-1}, \qquad i > 0$$

(recall that $\lambda^+(\lambda) = \lambda^2/(1-\lambda)^2$). After $\phi_{i-1} \in E^+_{t_i}$ is found, we find $\phi'_i \in E^+_{t_i}$, the Newton iterate of $\phi_{i-1}$ (the Newton method is applied to the restriction of $F^+$ onto $E^+_{t_i}$), and we then define

$$\phi_i = \frac{t_{i+1}}{t_i}\phi'_i \in E^+_{t_{i+1}}.$$

Then the next iteration is performed. The approximate solution to (3.4.5) produced at $i$th iteration is

$$(x_i = X^*(\phi_i),\ \tau_i = \tau_{\phi_i})$$

(this pair is feasible for (3.4.5); see Lemma 3.4.1(ii)).

By virtue of the above-stated properties of the family $\mathcal{F}^{**}(F^+, E^+_0, b)$ (see Proposition 3.4.2 and (3.4.4)), our standard arguments prove the implication

$$\lambda(F_{t_0}, \phi_{-1}) \le \lambda \quad \Rightarrow \quad (\forall i):\ \{\lambda(F_{t_i}, \phi'_i) \le \lambda^+(\lambda)\} \& \{\lambda(F_{t_{i+1}}, \phi_i) \le \lambda\}.$$

Therefore, by Lemma 3.4.1, we have

(3.4.21)
$$\begin{aligned}&\lambda(F_{t_0}, \phi_{-1}) \le \lambda\\ &\Rightarrow (\forall i):\ x_i \in G,\ Ax_i = \tau_i b,\ \varepsilon_i \equiv 1 - \frac{\tau_i}{t^*} \le \Omega \exp -\frac{\lambda - \lambda^+(\lambda))}{(1+\lambda)\vartheta^{1/2}} i,\end{aligned}$$

where

(3.4.22)
$$\Omega = \frac{\vartheta}{t^* t_0 b^T b}.$$

Let us verify that the premise in (3.4.21) is true.

Indeed, we clearly have $\lambda(F_{t_0}, \phi_{-1}) \le \lambda(F^+, \phi_{-1})$. We have

$$D^2F^+(0)[\zeta,\zeta] = \zeta^T A\{(F^*)''(0)\}A^T\zeta = \zeta^T A[F''(0)]^{-1}A^T\zeta = \zeta^T\zeta,$$

which implies $D^2F^+(0)[b,b] = \delta^2$. Thus, by Theorem 2.1.1, we have, for $0 < t\delta < 1$,

$$| D^2F^+(tb)[b,\zeta] | \leq \frac{\delta}{(1-t\delta)^2}(D^2F^+(0)[\zeta,\zeta])^{1/2},$$

which, combined with the relation $DF^+(0) = 0$, leads to

$$| DF^+(tb)[\zeta] | \leq \frac{t\delta}{1-t\delta}(D^2F^+(0)[\zeta,\zeta])^{1/2}$$

for each $\zeta$ or, by virtue of Theorem 2.1.1, to

$$| DF^+(tb)[\zeta] | \leq \frac{t\delta}{(1-t\delta)^2}(D^2F^+(tb)[\zeta,\zeta])^{1/2}.$$

The latter inequality means that $\lambda(F^+, tb) \leq t\delta(1-t\delta)^{-2}$. This, by virtue of the choice of $t_0$, leads to the relation $\lambda(F_{t_0}, \phi_{-1}) \leq \lambda(F^+, \phi_{-1}) \leq \lambda$. Thus, the premise in (3.4.21) does hold.

To obtain the efficiency estimate for the above method, it remains to evaluate $\Omega$. Let us prove that

$$\Omega \leq 2\frac{\vartheta}{\lambda}.$$

Indeed, since $A(F''(0))^{-1}A^T = I_n$, the point $w = (F''(0))^{-1}A^Tb$ is the nearest to 0 point of the plane $\{x \mid Ax = b\}$ (in the Euclidean metric on $\mathbf{R}^m$ induced by the inner product $\langle h, e\rangle = h^TF''(0)e$). The ellipsoid $W = \{x \in \mathbf{R}^m \mid x^TF''(0)x \leq 1\}$ is contained in $G$ (Theorem 2.1.1(ii)), which implies that

$$t^* \geq \frac{1}{(w^TF''(0)w)^{1/2}} = \frac{1}{\| b \|_2}.$$

Besides this, $t_0 \geq \lambda/(2\delta) = \lambda/(2 \| b \|_2)$ (evidently); hence

$$\Omega = \frac{\vartheta}{t^*t_0 \| b \|_2^2} \leq 2\frac{\vartheta}{\lambda}.$$

Now we obtain from (3.4.21) the following estimate for the relative accuracy of $x_i$:

$$\varepsilon_i \equiv 1 - \frac{\tau_i}{t^*} \leq 2\frac{\vartheta}{\lambda}\exp\left\{-\frac{\lambda - \lambda^+(\lambda)}{(1+\lambda)\vartheta^{1/2}}i\right\}. \tag{3.4.23}$$

The optimal choice of $\lambda$ is

$$\lambda = 0.206...;$$

under this choice, for each $\varepsilon \in (0,1)$, the inequality $\varepsilon_i \leq \varepsilon$ holds for all $i$ such that

$$i \geq N(\varepsilon) \equiv 1 + 12.4\vartheta^{1/2}\ln\frac{7\vartheta}{\varepsilon}. \tag{3.4.24}$$

Note that the implementation of the dual parallel trajectories method requires an explicit representation of the Legendre transformation of $F$. This requirement is satisfied for LP problems in format (3.4.6).

The arithmetic cost per iteration for the above method as applied to (3.4.6) is $O(mn^2)$. A Karmarkar's type speed-up for the method that, in the case of problem (3.4.6), reduces the average (over iterations) arithmetic cost to $O(m^{1/2}n^2 + mn)$ is described in [Ns 88a], [Ns 89].

## 3.5 Primal parallel trajectories method

Consider the following standard problem with *linear* objective (it is more convenient to deal with its maximization formulation instead of the minimization one):

$$c^T x \to \max \mid x \in G, \tag{3.5.1}$$

where $G$ is a *bounded* closed convex domain in $\mathbf{R}^n$. Assume that we are given a $\vartheta$-self-concordant barrier $F$ for $G$ and that we know the $F$-center of $G$ (i.e., the minimizer of $F$ on $\operatorname{int} G$); let this center be 0, that is,

$$0 \in \operatorname{int} G, \;\; F'(0) = 0 \tag{3.5.2}$$

(henceforth, $F'$, $F''$ are the gradient and Hessian of $F$ with respect to the standard Euclidean structure on $\mathbf{R}^n$). Without loss of generality, assume that

$$c^T c = 1.$$

In the barrier path-following method, we follow the path of minimizers

$$x^*(s) = \operatorname{argmin}\{\bar{F}_s(x) = -sc^T x + F(x) \mid x \in \operatorname{int} G\};$$

in the method of centers, we also follow this path but use another parameterization of it. It is important that, in both methods, we use, in a sense, piecewise-constant (i.e., zero-order) approximations of the path. It is natural to follow the path using the first-order information on it. Note that the path is defined by the relation $\nabla F_s(\cdot) = 0$, i.e., by the equation

$$F'(x^*(s)) = sc.$$

Since $G$ is bounded, $F''$ is nondegenerate, and the implicit function theorem implies that the path is differentiable and satisfies the differential equation

$$(x^*)'(s) = [F''(x^*(s))]^{-1}c.$$

In particular, the derivative of the objective $c^T x$ along the path is positive, and we can use the value of the objective at a point of the path as our new parameter,

$$x^*(s) = x^{\#}(t(s)),$$

where

$$c^T x^\#(t) = t$$

and $t$ varies between 0 (the value of the objective at the $F$-center of $G$; see (3.5.2)) and $t^*$, where

$$t^* = \max\{c^T x \mid x \in G\}$$

is the optimal value in (3.5.1). As we just have seen, the derivative $(x^\#)'(t)$ of the path $x^\#(\cdot)$ at a point $t$, $0 < t < t^*$ is proportional to the direction $[F''(x^\#(t))]^{-1}c$ and should satisfy the identity

$$c^T (x^\#)'(t) = 1$$

(the definition of the parameterization), whence

$$(x^\#)'(t) = \Phi(x^\#(t)), \qquad \Phi(x) = \frac{[F''(x)]^{-1}c}{c^T[F''(x)]^{-1}c}.$$

Let us follow the path $x^\#(t)$ using the first-order information. Assume that we are given, for the current value $t$ of the parameter, a "close" to $x^\#(t)$ point $x(t)$. Then the vector $\Phi(x(t))$ is "close" to the derivative of the path at $t$. Having chosen a new value $t^+$ of the parameter, we can regard the point

$$x^+(t^+) = x(t) + (t^+ - t)\Phi(x(t))$$

as a natural approximation to the point $x^\#(t^+)$. Since the errors of approximation may accumulate, we cannot iterate this process straightforwardly and need some correction technique. The simplest way to perform the correction is to accompany each *predictor step* $(t, x(t)) \to (t^+, x^+(t^+))$ by a *corrector* step, which should restore closeness of the updated $x$ to $x^\#(t^+)$. The derivative of $F$ at $x^\#(t^+)$ is proportional to $c$ and $c^T x^\#(t^+) = t^+$, so that $x^\#(t)$ is the minimizer of $F$ on the intersection $(\operatorname{int} G) \bigcap \{x \mid c^T x = t^+\}$. Therefore the simplest way to restore the closeness to the path is to minimize the restriction of $F$ onto the above intersection by a number of Newton steps. In what follows, we describe a method of this type, which requires a single Newton correction step per every predictor step—the *primal parallel trajectories method.* Note that the idea of the method is close to that one of the predictor-corrector methods developed for LP by Mizuno, Todd, and Ye [MTY 89] and Mehrotra [Mh 89].

The primal parallel trajectories method for problem (3.5.1) is defined by parameters $\lambda_1$, $\lambda_2$ such that

$$0 < \lambda_1 < \lambda_*; \qquad 0 < \lambda_2 < \tfrac{1}{3}; \tag{3.5.3}$$

$$\lambda^+(\lambda_1) + \frac{\lambda_2}{1-\lambda_2} \le \lambda_1(1-\lambda_2) \tag{3.5.4}$$

(recall that $\lambda^+(\lambda) = \lambda^2(1-\lambda)^{-2}$).

The method is as follows.

1. *Initialization.* Let

$$x_{-1} = \tau_0 e, \tag{3.5.5}$$

where

$$\tau_0 = \max\left\{\tau \le 1 \mid \frac{\tau}{(1-\tau)^2} \le \lambda_1\right\}, \qquad e = \frac{[F''(0)]^{-1}c}{(c^T[F''(0)]^{-1}c)^{1/2}}.$$

2. *The ith step.* Let $x_{i-1} \in \operatorname{int} G$ be the previous approximate solution. Denote the set $\{y \in \operatorname{int} G \mid c^T(y-x) = 0\}$ by $E(x)$ and the restriction of $F$ onto $E(x)$ by $F_x(y)$. Let (the *corrector step*) $x_i^+ \in E(x_{i-1})$ be the Newton iterate of $x_{i-1}$ (the Newton method is applied to $F_{x_{i-1}}(\cdot)$; it is shown that $x_i^+ \in \operatorname{int} G$). After $x_i^+$ is found, we define (the *predictor step*) $x_i$ as

$$x_i = x_i^+ + \lambda_2 \frac{[F''(x_i^+)]^{-1}c}{(c^T[F''(x_i^+)]^{-1}c)^{1/2}} \tag{3.5.6}$$

(it will be shown that $x_i \in \operatorname{int} G$). The $i$th step is completed.

Let $\Delta = (0, t^*)$ (recall that $t^* > 0$ by (3.5.2)) and let $G^* = \{x \in G \mid c^T x \ge 0\}$. For each $t \in \Delta$, the set $G_t = \{x \in G \mid c^T x = t\}$ is nonempty. The restriction $F_t$ of $F$ onto the relative interior of $G_t$, by virtue of Proposition 2.3.1(i), is a $\vartheta$-self-concordant barrier for $G_t$ (the latter set is regarded as a full-dimensional subset of the corresponding hyperplane). Since $G_t$ is bounded, $F_t$ attains its minimum over the relative interior of $G_t$ at the unique point $x^\#(t)$ (Proposition 2.3.2(ii)). By definition of $x^\#(t)$, we have

$$F'(x^\#(t)) = s(t)c \tag{3.5.7}$$

for certain $(t)$ ($s(t) \ge 0$ by (3.5.2)). The $C^3$-smoothness of $F$ and the nondegeneracy of $D^2F$ imply that $x^\#(t)$ and $s(t)$ are $C^2$-smooth on $\Delta$.

The main result on the primal parallel trajectories method is as follows.

**Proposition 3.5.1** *The primal parallel trajectories method is well defined: For all $i$, the points $x_{i-1}$, $x_i^+$ and $x_i$ are well defined and belong to* $\operatorname{int} G$. *Moreover, for each $i \ge 0$, we have*

$$(\mathrm{I}_i): \quad t_{i-1} \equiv c^T x_{i-1} > 0, \tag{3.5.8}$$

$$(\mathrm{J}_i): \quad \lambda(F_{t_{i-1}}, x_{i-1}) \le \lambda_1, \tag{3.5.9}$$

$$(\mathrm{K}_i): \qquad s(t_i) \ge \left(1 + \frac{\Omega}{\vartheta^{1/2}}\right) s(t_{i-1}), \tag{3.5.10}$$

*where*

$$\Omega = \tfrac{5}{9}\left(1 - (1 - 3\lambda_2\omega(\lambda_1))^{5/3}\right),$$

$$(\mathrm{L}_i): \quad t^* - t_i \le \frac{\vartheta}{s(t_i)}. \tag{3.5.11}$$

*Furthermore, the relative accuracy of ith iterate satisfies the inequality*

$$\varepsilon_i \equiv 1 - \frac{c^T x_i}{c^T x^*} \le \frac{\vartheta}{\gamma}\left(1 + \frac{\Omega}{\vartheta^{1/2}}\right)^{-i}, \tag{3.5.12}$$

*where $\gamma$ depends on $\lambda_1$, $\lambda_2$ only and where $x^*$ is a solution to* (3.5.1).

**Proof.** Let us establish some properties of $x^{\#}(t)$. Taking the derivative with respect to $t$ in (3.5.7) and in the identity $c^T x^{\#}(t) = t$, we obtain

$$s'(t) = c^T[F''(x^{\#}(t))]^{-1}c, \quad (x^{\#}(t))' = \frac{[F''(x^{\#}(t))]^{-1}c}{c^T[F''(x^{\#}(t))]^{-1}c}. \tag{3.5.13}$$

Let us choose $\tau \in \Delta$ and let

$$\| h \|_\tau = \{h^T F''(x^{\#}(\tau))h\}^{1/2}.$$

Equation (3.5.13) implies

$$\| (x^{\#}(\tau))' \|_\tau = \frac{1}{\{c^T[F''(x^{\#}(\tau))]^{-1}c\}^{1/2}} \equiv \phi(\tau). \tag{3.5.14}$$

Moreover, by Theorem 2.1.1 and in view of (3.5.14), we have

(3.5.15)
$$\| x^{\#}(\tau) - x^{\#}(t) \|_t < 1 \ \Rightarrow \| (x^{\#}(\tau))' \|_\tau \le \frac{\phi(t)}{1- \| x^{\#}(\tau) - x^{\#}(t) \|_t)}$$

$$\Rightarrow \| (x^{\#}(\tau))' \|_t \le \frac{\phi(t)}{(1- \| x^{\#}(\tau) - x^{\#}(t) \|_t)^2}.$$

By Theorem 2.1.1, the set $\{y \in \mathbf{R}^n \mid \| y - x^{\#}(t) \|_t < 1\}$ is contained in $\operatorname{int} G$, which, combined with (3.5.15), proves the implication

$$0 \le \tau - t < (3\phi(t))^{-1}$$

$$\Rightarrow \{\tau \in \Delta, \| x^{\#}(\tau) - x^{\#}(t) \|_t \le 1 - \{1 - 3(\tau - t)\phi(t)\}^{1/3}, \tag{3.5.16}$$

$$s'(\tau) \ge \phi^2(t)\{1 - 3(\tau - t)\phi(t)\}^{2/3}\}$$

(we have taken into account (3.5.13) and the implication

$$\| x^{\#}(\tau) - x^{\#}(t) \|_t < 1 \Rightarrow c^T[F''(x^{\#}(\tau)]^{-1}c$$

$$\ge \{1- \| x^{\#}(\tau) - x^{\#}(t) \|_t\}^2 c^T[F''(x^{\#}(t))]^{-1}c$$

(see Theorem 2.1.1)).

Now let us prove $(\mathrm{I}_i)$ – $(\mathrm{L}_i)$. Let

$$J = \{i \geq 0 \mid (\mathrm{I}_j),\ (\mathrm{J}_j) \text{ hold for } 0 \leq j \leq i,\ (\mathrm{K}_j),\ (\mathrm{L}_j) \text{ hold for } 0 \leq j < i,$$

$$x_{j-1} \in \operatorname{int} G,\ 0 \leq j \leq i,\ x_j^+ \in \operatorname{int} G,\ 0 \leq j < i\}.$$

We wish to prove that $J = \{i \geq 0\}$; it suffices to verify that $0 \in J$ and that

$$j \in J \Rightarrow j+1 \in J.$$

Let us first prove that $0 \in J$, i.e., that $x_{-1} \in \operatorname{int} G$, $t_{-1} > 0$, and $(\mathrm{J}_0)$ holds. By (3.5.5), we have $e^T F''(0) e = 1$, whence $\tau e \in \operatorname{int} G$ for $0 \leq \tau < 1$, and $\| F'(\tau e) \|_0 \leq \tau(1-\tau)^{-1}$ (in view of Theorem 2.1.1 and the relation $F'(0) = 0$) or

$$| h^T F'(\tau e) | \leq \frac{\tau}{1-\tau} \{h^T F''(0) h\}^{1/2} \leq \frac{\tau}{(1-\tau)^2} \{h^T F''(\tau e) h\}^{1/2},$$

whence $\lambda(F, \tau_0 e) \leq \lambda_1$. It is clear that $\lambda(F_{t_{-1}}, \tau_0 e) \leq \lambda(F, \tau_0 e)$, which implies $(\mathrm{J}_0)$. Furthermore,

$$t_{-1} = c^T \tau_0 e = \tau_0 \{c^T [F''(0)]^{-1} c\}^{1/2},$$

which, by virtue of the evident inequality $\tau_0 \geq \lambda_1/2$, implies that

$$t_{-1} \geq \frac{\lambda_1}{2} \{c^T [F''(0)]^{-1} c\}^{1/2}. \tag{3.5.17}$$

In particular, $t_{-1} > 0$. Thus, $0 \in J$.

Now let $i \in J$; let us prove then that $i+1 \in J$. First, in view of $\lambda_1 < \lambda_*$ and the fact that $F_{t_{i-1}}$ is a barrier for the set $G_{t_{i-1}}$, $(\mathrm{J}_i)$ implies (see Theorem 2.2.3) the relation

$$x_i^+ \in \operatorname{rint} G_{t_{i-1}} \subset \operatorname{int} G, \qquad \lambda(F_{t_{i-1}}, x_i^+) \leq \lambda^+(\lambda_1) \tag{3.5.18}$$

("rint" denotes the relative interior). Furthermore, let

$$e_i = \frac{[F''(x_i^+)]^{-1} c}{\{c^T [F''(x_i^+)]^{-1} c\}^{1/2}}.$$

Then

$$x_i = x_i^+ + \lambda_2 e_i, \qquad e_i^T F''(x_i^+) e_i = 1, \tag{3.5.19}$$

whence, in view of Theorem 2.1.1(ii) and the inclusion $\lambda_2 \in (0,1)$, $x_i \in \operatorname{int} G$.

We have

$$t_i = c^T x_i = c^T x_i^+ + \lambda_2 c^T e_i$$
$$(3.5.20) \qquad = c^T x_i + \lambda_2 \{c^T [F''(x_i^+)]^{-1} c\}^{1/2}$$
$$= t_{i-1} + \lambda_2 \{c^T [F''(x_i^+)]^{-1} c\}^{1/2}.$$

In particular, $t_i > t_{i-1}$, and $(\mathrm{I}_{i+1})$ holds.

By virtue of (3.5.18) and Theorem 2.2.2(iii), we have

$$(x^{\#}(t_{i-1}) - x_i^+) F''(x_i^+)(x^{\#}(t_{i-1}) - x_i^+) \le \omega^2(\lambda^+(\lambda_1)),$$

whence, by Theorem 2.1.1,

$$\frac{1 - \omega(\lambda^+(\lambda_1))}{\phi(t_{i-1})} \le \{c^T [F''(x_i^+)]^{-1} c\}^{1/2} \le \frac{1}{(1 - \omega(\lambda^+(\lambda_1)))\phi(t_{i-1})}$$

(see (3.5.14)). Therefore (3.5.20) implies

$$(3.5.21) \qquad t_i \ge \tau_i \equiv t_{i-1} + \frac{\lambda_2 (1 - \omega(\lambda^+(\lambda_1)))}{\phi(t_{i-1})}.$$

Since $\lambda_2 < \frac{1}{3}$, relations (3.5.21) and (3.5.16) imply the relations

$$(3.5.22) \quad s'(\tau) \ge \phi^2(t_{i-1}) \{1 - 3(\tau - t_{i-1})\phi(t_{i-1})\}^{2/3}, \qquad t_{i-1} \le \tau \le \tau_i,$$

whence, since $s$ evidently increases,

$$(3.5.23) \qquad \begin{gathered} s(t_i) \ge s(t_{i-1}) + \Omega \phi(t_{i-1}), \\ \Omega = \tfrac{5}{9}(1 - \{1 - 3\lambda_2(1 - \omega(\lambda^+(\lambda_1)))\}^{5/3}). \end{gathered}$$

Since

$$\phi(t) = \frac{1}{(c^T [F''(x^{\#}(t))]^{-1} c)^{1/2}}$$
$$= \frac{s(t)}{\{(F'(x^{\#}(t)))^T [F''(x^{\#}(t))]^{-1} F'(x^{\#}(t))\}^{1/2}} \ge \frac{s(t)}{\vartheta^{1/2}},$$

we obtain

$$s(t_i) \ge s(t_{i-1}) \left(1 + \frac{\Omega}{\vartheta^{1/2}}\right).$$

The latter relation is $(\mathrm{K}_i)$.

Furthermore, let $x^*$ be the solution to (3.5.1). Then we have

$$t^* - t_i = c^T(x^* - x_i) = c^T(x^* - x^{\#}(t_i)) = \frac{[F'(x^{\#}(t_i))]^T (x^* - x^{\#}(t_i))}{s(t_i)} \le \frac{\vartheta}{s(t_i)}$$

(we have used (2.3.2)), which implies $(\mathrm{L}_i)$.

To prove the inclusion $i+1 \in J$, it remains to verify ($\mathrm{J}_{i+1}$). Let $c^T h = 0$ and let

$$x(r) = x_i^+ + r e_i, \qquad 0 \le r \le \lambda_2.$$

Then

$$\begin{aligned}\left|\frac{d}{dr}(F'(x(r))^T h\right| &= \mid e_i^T F''(x(r))h \mid \le \{e_i^T F''(x(r))e_i\}^{1/2}\{h^T F''(x(r))h\}^{1/2} \\ &\le \frac{1}{(1-r)^2}\{e_i^T F''(x(0)e_i\}^{1/2}\{h^T F''(x(0))h\}^{1/2} \\ &= \frac{1}{(1-r)^2}\{h^T F''(x(0))h\}^{1/2}\end{aligned}$$

(we have taken into account Theorem 2.1.1 and the relation $e_i^T F''(x(0))e_i = 1$). By (3.5.18), the relation $c^T h = 0$ implies

$$\mid h^T F'(x(0)) \mid \le \lambda^+(\lambda_1)\{h^T F''(x(0))h\}^{1/2},$$

so that

$$\begin{aligned}\mid h^T F'(x(\lambda_2)) \mid &\le \left(\lambda^+(\lambda_1) + \frac{\lambda_2}{1-\lambda_2}\right)\{h^T F''(x(0))h\}^{1/2} \\ &\le \left(\frac{\lambda^+(\lambda_1)}{1-\lambda_2} + \frac{\lambda_2}{(1-\lambda_2)^2}\right)\{h^T F''(x(\lambda_2))h\}^{1/2}\end{aligned}$$

(see Theorem 2.1.1). The inequality obtained means that

$$\lambda(F_{t_i}, x_i) \le \frac{\lambda^+(\lambda_1)}{1-\lambda_2} + \frac{\lambda_2}{(1-\lambda_2)^2},$$

which, by virtue of (3.5.4), leads to ($\mathrm{J}_{i+1}$). The proposition is proved. □

We see that the rate of convergence of the primal parallel trajectories method is the same as the rate of convergence of our previous methods: It needs no more than $O(\vartheta^{1/2}\ln(2\vartheta/\varepsilon))$ iterations to produce an approximation $x_i$ such that $\varepsilon_i \le \varepsilon \in (0,1)$; the constant factor in $O(\cdot)$ depends on $\lambda_1$, $\lambda_2$ only.

A reasonable choice of the parameters is $\lambda_1 = 0.266$, $\lambda_2 = 0.096$. Under this choice, (3.5.12) leads to

$$\varepsilon_i \le 11.78\vartheta\left(1 + \frac{0.107}{\vartheta^{1/2}}\right)^{-i}.$$

*This page intentionally left blank*

# Chapter 4

# Potential reduction interior-point methods

In the previous chapter, we described a number of polynomial-time interior-point methods based on following certain path. Below, we present another *family of polynomial-time interior-point methods based on explicit Lyapunov's* functions. Namely, in what follows, we generalize onto nonlinear convex problems three potential reduction methods initially developed for LP, specifically, the famous method of Karmarkar [Ka 84], the similar projective method [Nm 87], and the primal-dual method of Todd and Ye [TY 87], [Ye 88a], [Ye 89].

From a theoretical viewpoint, potential reduction interior-point methods have no advantages as compared to the path-following procedures. As we have seen, a path-following method associated with a $\vartheta$-self-concordant barrier for a bounded convex domain minimizes a linear function over this domain to an accuracy $\varepsilon$ in $O(\vartheta^{1/2}\ln(\vartheta/\varepsilon))$ Newton-type steps; the efficiency estimate of the generalized primal-dual method is the same, and the estimates for the generalized method of Karmarkar and the projective method are even worse $(O(\vartheta\ln(\vartheta/\varepsilon)))$. Nevertheless, the potential reduction methods are very important because of the following reason. The accuracy attained by such a method can be estimated in terms of the amount at which the corresponding Lyapunov's function (*the potential*) is reduced during the solution process. At each step of a potential reduction method, the theory prescribes the direction and the stepsize, which ensure certain decreasing of the potential, but we are not forbidden to look for deeper decreasing; say, we can minimize the potential in the above direction with the aid of line search (this is called "large steps"). It is well known that, in the case of LP problems, large steps considerably improve behaviour of the interior-point methods. In contrast, a path-following method must maintain closeness to the corresponding path, and it is not clear how to provide this requirement for large steps without violation of the theoretical efficiency estimates (cf. §3.2.6).

It turns out that, to apply most of the potential reduction methods, we should properly reformulate the problem, i.e., transform it into the so-called conic form (where we should minimize a *linear* objective over the intersection of a closed convex cone and *an affine* subspace). This universal form of a convex programming problem is introduced and studied in §4.1. Section 4.2 is devoted to duality for conic problems; the specifics of conic formulation allows us to develop quite symmetric duality theory for conic problems, which can be regarded as a very natural extension of the standard linear programming duality. This duality can be also regarded as a particular case of general duality theory for problems with "nonnegativity constraints" defined by a convex cone in a Banach space (see, e.g., [ET 76]); the specifics of the presented duality results are caused by linearity of the objective and the equality constraints involved into the problem under consideration. Sections 4.3–4.5 contain the descriptions of the (generalized) method of Karmarkar, the projective method, and the primal-dual method, respectively.

## 4.1 Conic formulation of convex program

### 4.1.1 Motivations

A usual linear programming problem can be written as

$$(*) \qquad \langle d, y \rangle \to \min,$$

$$(**) \qquad F(y) \in K,$$

where $y \in \mathbf{R}^m$, $K = \mathbf{R}^n_+$ is the standard cone $\{x \geq 0\}$ in $\mathbf{R}^n$, and $F$ is an affine mapping from $\mathbf{R}^m$ into $\mathbf{R}^n$; as always, $\langle d, y \rangle$ denotes the value of a linear form $d$ at a vector $y$. The standard formulation of a nonlinear optimization problem has basically the same form, but with a nonlinear mapping $F$ (the objective, of course, always can be chosen to be linear). Thus, in the usual formulation of a nonlinear optimization problem, we deal with the standard "polyhedral" notion of nonnegativity ($K$ is the nonnegative orthant) and a nonlinear manifold $F(\mathbf{R}^m)$. At the same time, we can transform $(*)$, $(**)$ into a nonlinear problem by replacing the nonnegative orthant with a nonpolyhedral cone and preserving the linearity of $F$.

Assume that, in $(*)$, $(**)$, $F$ is an affine mapping and $K$ is a closed convex cone in a finite-dimensional linear space $E$, $\operatorname{int} K \neq \emptyset$. We also assume $K$ to be pointed (i.e., $K \bigcap \{-K\} = \{0\}$). Note that, in $(*)$, $(**)$, we can choose as the unknown the vector $x = F(y)$ instead of $y$ itself; it suffices to add the constraint $x \in L = F(\mathbf{R}^m)$ and to express the objective in terms of $x$ instead of $y$. The latter can be done if and only if $d$ is constant along the kernel of the homogeneous part of $F$; if it is not the case, then $(*)$, $(**)$ is either inconsistent, or below unbounded. Thus, at least every solvable problem $(*)$, $(**)$ with an affine $F$ can be rewritten as

$$(\mathsf{P}): \qquad \langle c, x \rangle \to \min \mid x \in K \bigcap (L + b),$$

where $K$ is a pointed closed convex cone with a nonempty interior in $E$, $c$ belongs to the space $E^*$ conjugate to $E$, $b \in E$, and $L$ is a subspace of $E$.

Under these assumptions on the data involved, we call $(\mathsf{P})$ *a conic problem.*

### 4.1.2 Conic and standard problems

A conic problem, of course, is a special convex programming problem. In turn, each convex problem can be rewritten in the conic form. It clearly suffices to transform into the conic form a standard convex problem (see Chapter 3), i.e., the problem

$$\text{minimize } \langle c', y \rangle \quad \text{s.t. } y \in G,$$

where $G$ is a closed convex domain in $\mathbf{R}^n$. Without loss of generality, we can assume that $G$ does not contain straight lines (otherwise, the problem either is below unbounded, or we can reduce it to an equivalent standard problem with the feasible domain not containing straight lines; it suffices to replace $G$ with its cross section by an appropriate subspace of $E$).

Let us embed $\mathbf{R}^n$ into $E = \mathbf{R}^{n+1}$ as the affine hyperplane $\{x^{(n+1)} = 1\}$ and let $K$ be the closed conic hull of the image of $G$,

$$K = \mathrm{cl}\left\{x \in \mathbf{R}^{n+1} \mid x^{(n+1)} > 0, \ \frac{1}{x^{(n+1)}}(x^{(1)}, ..., x^{(n)})^T \in G\right\}.$$

Clearly, $K$ is a pointed closed cone with a nonempty interior, and $G$ is the intersection of $K$ and the hyperplane $(L + b)$, $L = \{x \in E \mid x^{(n+1)} = 0\}$, $b = (0, ..., 0, 1)^T$ (more rigorously, $G$ is the inverse image of this intersection under the above embedding $F : \mathbf{R}^n \to \mathbf{R}^{n+1}$). The initial objective $c'$ can be thought of as a linear functional $c$ on $\mathbf{R}^{n+1}$ independent on the last coordinate; the conic problem defined by $c$, $K$ and $L$ clearly is equivalent to the initial problem.

Of course, the above reduction is, in fact, the result of a very liberal usage of the (unformal) notion of equivalence between optimization problems. From the algorithmic viewpoint, the possibility to "explicitly" reduce a given convex problem to a conic problem depends on the form in which the initial problem is represented. This question is discussed in more detail in Chapters 5 and 6.

## 4.2 Duality for conic problems

### 4.2.1 Dual problem

Assume for a moment that (P) is solvable and let $x^*$ be an optimal solution to the problem. Assume also that at this solution the Kuhn–Tucker optimality condition holds, so that we can add to the objective a linear functional $y$ constant along $L+b$ in such a way that the resulting linear functional $s^* \equiv c+y$ is nonnegative on the set $K - x^*$. What can we say about $s^*$ ? First, since $K$ is a cone, the inequality

$$(4.2.1) \qquad \langle s^*, x - x^* \rangle \geq 0, \qquad x \in K$$

implies $s^* \in K^*$, where

$$K^* = \{s \in E^* \mid \langle s, x \rangle \geq 0 \ \forall\, x \in K\}$$

is the cone dual to $K$ (under our assumptions on $K$ and $E$, this cone is a pointed closed convex cone with a nonempty interior in $E^*$). Second, (4.2.1) holds for $x = 0$, so that $\langle s^*, x^* \rangle \le 0$, and since $x^* \in K$ and $s^* \in K^*$, we have $\langle s^*, x^* \rangle \ge 0$ as well. It follows that $\langle s^*, x^* \rangle = 0$.

Last, the functional $s^* - c$ is constant along $L + b$, or, which is the same, $s^* - c$ belongs to the subspace $L^\perp \subset E^*$ formed by the functionals vanishing on $L$. Thus, $s^*$ is a feasible solution to the problem

$$\text{minimize } \langle s, x^* \rangle \quad \text{s.t. } s \in K^* \bigcap (L^\perp + c).$$

Moreover, $s^*$ clearly is the optimal solution to this problem (since the objective of the problem is nonnegative on $K^*$ and equals zero at $s^*$). Note that, if we replace the objective $x^*$ of the latter problem with an arbitrary $x \in (L + b)$, say, with $b$, the point $s^*$ remains optimal for the perturbed problem, since the quantity $\langle s, x^* - x \rangle$ clearly is constant when $s$ varies along $(L^\perp + c)$, provided that $x \in (L + b)$.

Thus, any $s^*$ given by the Kuhn–Tucker optimality condition at $x^*$ must belong to the optimal set of the problem

$$\text{minimize } \langle s, b \rangle \quad \text{s.t. } s \in K^* \bigcap (L^\perp + c)$$

as well as must satisfy the complementary slackness relation $\langle s^*, x^* \rangle = 0$.

Note that the resulting problem is again a conic problem.

The above considerations motivate the following definition.

**Definition 4.2.1** *Let $E$ be a finite-dimensional space and $K$ be a closed convex pointed cone in $E$ with a nonempty interior; also let $b \in E$ and $c \in E^*$. The data $E$, $K$, $L$, $b$, $c$ define a pair of conic problems*

$$(\mathsf{P}): \qquad \textit{minimize } \langle c, x \rangle \quad \text{s.t. } x \in K \bigcap (L + b)$$

*and*

$$(\mathsf{D}): \qquad \textit{minimize } \langle s, b \rangle \quad \text{s.t. } s \in K^* \bigcap (L^\perp + c),$$

*wherein $K^* \subset E^*$ is the cone dual to $K$ and $L^\perp \subset E^*$ is the annulator of $L$.*

*(P) and (D) are called, respectively, primal and dual problems associated with the above data.*

Note that, in the polyhedral case ($K$ is the standard $(\dim E)$-facet cone in $E$), the pair (P), (D) is, in fact, the standard primal-dual pair of linear programming problems.

In what follows, we try to extend onto our general conic case the standard relations between the primal and the dual LP problems.

First, our duality is symmetric: (P) can be thought of as the problem dual to (D) (of course, we use the canonical isomorphism $(E^*)^* \equiv E$, so that $(K^*)^* = K$ and $(L^\perp)^\perp = L$). Thus, our duality is quite symmetric, similar to the LP duality; the dual problem remembers everything about the primal problem. Note that the standard Lagrange duality for convex (nonlinear) programming does not possess such a symmetry. For example, for the primal problem of the type

$$f_0(x) \to \min \mid f_i(x) \le 0,\ 1 \le i \le m,\ x \in E$$

where $f_i$, $i \ge 0$, are, say, strongly convex functions, the dual problem is

$$\text{maximize } \phi(s) \equiv \min\left\{ f_0(x) + \sum_{i=1}^{m} s_i f_i(x) \mid x \in E \right\} \quad \text{s.t. } s \ge 0,$$

and it is difficult to extract the primal problem from the dual one.

### 4.2.2 Duality relations

In what follows, we fix a primal-dual conic pair $(\mathsf{P})$, $(\mathsf{D})$. Our aim is to study the relations between properties of the problems comprising the pair. We denote by $\mathsf{P}^*$ the optimal value in the primal problem ($\mathsf{P}^* = +\infty$, if the set $D(\mathsf{P})$ of primal feasible solutions is empty; otherwise, $\mathsf{P}^* = \inf\{\langle c, x\rangle \mid x \in D(\mathsf{P})\}$); the set of feasible dual solutions and the corresponding optimal value are denoted by $D(\mathsf{D})$, $\mathsf{D}^*$, respectively.

*Preliminary results.* Let us start with the following simple statement.

**Lemma 4.2.1** *For every pair of primal and dual feasible solutions $x \in D(\mathsf{P})$, $s \in D(\mathsf{D})$, we have*

$$\langle c, b\rangle \le \langle c, b\rangle + \langle s, x\rangle = \langle c, x\rangle + \langle s, b\rangle\,; \tag{4.2.2}$$

*in particular, if both of the problems are consistent ($D(\mathsf{P}) \ne \emptyset$, $D(\mathsf{D}) \ne \emptyset$), then*

$$\langle c, b\rangle \le \mathsf{P}^* + \mathsf{D}^*. \tag{4.2.3}$$

**Proof.** We have $\langle c, x\rangle + \langle s, b\rangle - \langle c, b\rangle - \langle s, x\rangle = \langle c - s, x - b\rangle$, so that, to establish the equality in (4.2.2), it suffices to prove that $\langle c - s, x - b\rangle = 0$. Since $x$ is primal feasible, $x - b \in L$, so that each functional from $L^\perp$ vanishes at $x - b$. In particular, $\langle c - s, x - b\rangle = 0$, by virtue of $c - s \in L^\perp$ (recall that $s$ is dual feasible). Inequality (4.2.2) follows from $x \in K$, $s \in K^*$. Equation (4.2.3) is an immediate corollary of (4.2.2). □

**Corollary 4.2.1** *Assume that $x^*$ is primal feasible and $s^*$ is dual feasible. Then the following two conditions are equivalent:* $\langle s^*, x^*\rangle = 0$ (complementary slackness), $\langle c, b\rangle = \langle c, x^*\rangle + \langle s, b^*\rangle$ (zero duality gap), *and each of them is sufficient for $x^*$ and $s^*$ to be primal and dual optimal, respectively.*

The above statements motivate the following definitions. Let us say that a primal-dual pair of conic problems is *normal* ($\equiv$ possesses property (N)) if both of the problems are solvable and, for their optimal values, the equality

$$\mathsf{P}^* + \mathsf{D}^* = \langle c, b \rangle \tag{4.2.4}$$

holds. The pair is said to be *weakly normal* ($\equiv$ possessing property (WN)) if both of the problems are consistent and (4.2.4) holds.

*Dual problem and the cost function for the perturbed primal problem.* As in the standard duality theory, we can describe the feasible solutions to the dual problem $(\mathsf{D})$ in terms of support functionals to the optimal value of the primal problem regarded as a function of $b$. Let

$$B = K + L$$

be the set of all $b \in E$ such that the primal problem $(\mathsf{P}(b))$ defined by the data $c$, $K$, $L$, $b$ is feasible; $B$ clearly is convex, and $\operatorname{int} B \neq \emptyset$ (recall that $\operatorname{int} K \neq \emptyset$). Let

$$\mathsf{P}^*(b) = \inf\{\langle c, x\rangle \mid x \in K \bigcap (L + b)\} : B \to \mathbf{R} \bigcup \{-\infty\}.$$

Also, let $\mathsf{D}^*(b)$ be the optimal value in the problem $(\mathsf{D}(b))$ dual to $(\mathsf{P}(b))$.

The function $\mathsf{P}^*(b)$ clearly possesses the following properties:

(i) $\mathsf{P}^*(b)$ is convex;

(ii) For each $b \in B$, we have $b + L \subset B$, and $\mathsf{P}^*(b)$ is constant along $b + L$;

(iii) For each $b \in B$ and each $u \in K$, we have $b + u \in B$ and $\mathsf{P}^*(b + u) \leq \mathsf{P}^*(b) + \langle c, u\rangle$.

**Proposition 4.2.1** *A functional $s \in E^*$ is dual feasible if and only if there exists $\beta = \beta(s) \in \mathbf{R}$ such that*

$$\beta + \langle c - s, b\rangle \leq \mathsf{P}^*(b) \quad \forall b \in B. \tag{4.2.5}$$

*If $\bar{b} \in B$ and $c - s$ is the support on $B$ functional to $\mathsf{P}^*(\cdot)$ at $\bar{b}$, then $s$ solves $(\mathsf{D}(\bar{b}))$ and (WN) holds for the pair $((\mathsf{P}(\bar{b})), (\mathsf{D}(\bar{b})))$. Conversely, if (WN) holds for the latter pair and $s$ solves $(\mathsf{D}(\bar{b}))$, then $c - s$ is a support (on $B$) functional to $\mathsf{P}^*(\cdot)$ at $\bar{b}$.*

**Proof.** If $s$ is dual feasible, then (4.2.5) holds with $\beta = 0$ for each $b \in B$ by virtue of (4.2.2). Conversely, assume that $s$ satisfies (4.2.5). In view of (iii), we have $\mathsf{P}^*(b) \leq \mathsf{P}^*(0) + \langle c, b\rangle$, $b \in K$, which, combined with (4.2.5), implies $\beta - \mathsf{P}^*(0) \leq \langle s, b\rangle$, $b \in K$, so that $s \in K^*$. Furthermore, $\mathsf{P}^*(b) = \mathsf{P}^*(0)$, $b \in L$ (see (ii)), which, combined with (4.2.5), leads to $\langle c - s, b\rangle = 0$, $b \in L$, so that $c - s \in L^{\perp}$. Thus, $s$ is dual feasible.

Now assume that $c - s$ is a support functional to $\mathsf{P}^*(\cdot)$ at $\bar{b} \in B$, so that (4.2.5) holds with $\beta = \mathsf{P}^*(\bar{b}) - \langle c - s, \bar{b}\rangle$. Since $\mathsf{P}^*(\cdot)$ admits a support functional, this function is below bounded on bounded subsets of $B$. We always have $0 \in B$, and clearly $\mathsf{P}^*(0)$ is either zero or $-\infty$. The second alternative in our case is excluded, so that $\mathsf{P}^*(0) = 0$. Relation (4.2.5) with the above $\beta$ and with $b = 0$ therefore leads to $\mathsf{P}^*(\bar{b}) - \langle c - s, \bar{b}\rangle \leq 0$, so that $\mathsf{P}^*(\bar{b}) + \langle s, \bar{b}\rangle \leq \langle c, \bar{b}\rangle$. Since, as we have already proved, $s$ is dual feasible, it follows that $\mathsf{P}^*(\bar{b}) + \mathsf{D}^*(\bar{b}) \leq \mathsf{P}^*(\bar{b}) + \langle s, \bar{b}\rangle \leq \langle c, \bar{b}\rangle$. By virtue of (4.2.3), all the inequalities in this chain are equalities, so that $s$ is dual optimal and (WN) holds for the pair $((\mathsf{P}(\bar{b})), (\mathsf{D}(\bar{b})))$.

It remains to prove that, if (WN) holds for the pair $((\mathsf{P}(\bar{b})), (\mathsf{D}(\bar{b})))$ and $s$ solves $(\mathsf{D}(\bar{b}))$, then $c - s$ is a support functional to $\mathsf{P}^*(b)$ at $\bar{b}$. From (4.2.2), it follows immediately that $\mathsf{P}^*(b) \geq \langle c - s, b\rangle = \langle c - s, b - \bar{b}\rangle + \langle c - s, \bar{b}\rangle$. Since $\langle s, \bar{b}\rangle = \mathsf{D}^*(\bar{b}) = \langle c, \bar{b}\rangle - \mathsf{P}^*(\bar{b})$ (we have taken into account (WN)), we obtain $\mathsf{P}^*(b) \geq \langle c - s, b - \bar{b}\rangle + \mathsf{P}^*(\bar{b})$, and $c - s$ proves to be a support functional to $\mathsf{P}^*$ at $\bar{b}$. □

*Relations between properties of primal and dual problems.* Let $H$ be a vector space and let $\mathsf{T}$ be a conic problem on $H$ defined by the data $r$ (the objective), $Q$ (the cone), and the feasible plane $(M + d)$, where $M$ is the corresponding linear subspace of $H$. Introduce the following predicates:

- (F): Feasibility ($D(\mathsf{T}) \neq \emptyset$);
- (B): Boundedness of the feasible set ($D(\mathsf{T})$ is bounded, e.g., empty);
- (SB): Boundedness of the solution set (the set of optimal solutions to $(\mathsf{T})$ is nonempty and bounded);
- (R): Recessity ($M \bigcap Q \neq \{0\}$);
- (BO): Boundedness of the objective (the objective is below bounded on $D(\mathsf{T})$, e.g., due to $D(\mathsf{T}) = \emptyset$);
- (I): Existence of a feasible interior point ($D(\mathsf{T})$ intersects $\operatorname{int} Q$);
- (S): Solvability ($(\mathsf{T})$ is solvable).

Recall that we have fixed a primal-dual pair $(\mathsf{P})$, $(\mathsf{D})$ of conic problems. In what follows, the properties of the primal and the dual problems are marked with the subscripts $p$ and $d$, respectively; e.g., $(\mathrm{S}_p)$ is the abbreviation for the assertion "the primal problem is solvable." Our aim is to establish some relations between the introduced properties of the primal and the dual problems. Note that, in what follows, the assertions are arranged into "symmetric" pairs (the symmetry is based on the above-mentioned symmetry between the primal and the dual problems, and it suffices to prove only the first statement in each pair).

(1) $(\mathrm{F}_p) \Rightarrow (\mathrm{BO}_d)$; $(\mathrm{F}_d) \Rightarrow (\mathrm{BO}_p)$.

This is an immediate corollary of (4.2.2).

(2) $(\mathrm{F}_p)$ and $(\mathrm{B}_p) \Rightarrow (\mathrm{S}_p)$; $(\mathrm{F}_d)$ and $(\mathrm{B}_d) \Rightarrow (\mathrm{S}_d)$.

It is clear, due to the compactness arguments. □

**Remark 4.2.1** *In the linear programming case ($K$ is a polyhedral cone), item* (2) *can be essentially strengthened:* ($\mathrm{F}_p$) *and* ($\mathrm{BO}_p$) $\Rightarrow$ (N); *this strong statement cannot be extended onto nonlinear case, as it is could be demonstrated by simple examples.*

The following statements are more encouraging.

(3) ($\mathrm{F}_p$) and ($\mathrm{B}_p$) $\Rightarrow$ ($\mathrm{F}_d$) and ($\mathrm{BO}_d$); ($\mathrm{F}_d$) and ($\mathrm{B}_d$) $\Rightarrow$ ($\mathrm{F}_p$) and ($\mathrm{BO}_p$).

In fact, the premise in the first implication implies even (WN); this is demonstrated in item (7).

Implication ($\mathrm{F}_p$) and ($\mathrm{B}_p$) $\Rightarrow$ ($\mathrm{BO}_d$) is a consequence of item (1). To complete the proof of (3), let us first establish the following implication.

(4) $\rceil(\mathrm{I}_p) \Rightarrow (\mathrm{R}_d)$; $\rceil(\mathrm{I}_d) \Rightarrow (\mathrm{R}_p)$.

The premise in the first implication means that the affine subspace $(L+b)$ does not intersect the open nonempty convex set $\operatorname{int} K$. Therefore these two sets can be separated by a nonzero linear functional $s$,

$$x \in (\operatorname{int} K),\ y \in (L+b) \Rightarrow \langle s, x\rangle \geq \langle s, y\rangle .$$

In particular, $s$ is below bounded on $K$ and consequently belongs to $K^*$ and is above bounded on $L$, or, $s \in L^{\perp}$, which is the same. Thus, $0 \neq s \in L^{\perp} \bigcap K^*$, and ($\mathrm{R}_d$) follows. □

Now we can complete the proof of item (3). Assume that the primal problem is consistent and that its feasible set is bounded. We wish to prove that, under these assumptions, the dual problem is consistent. Otherwise, $\rceil(\mathrm{I}_d)$, and consequently (see (4)) ($\mathrm{R}_p$) would take place. That means that $\xi \in K \bigcap L$ for certain $\xi \neq 0$. The latter relation, combined with ($\mathrm{F}_p$), implies unboundedness of the primal feasible set, which contradicts the premise in (3).

**Remark 4.2.2** *In connection with implication* (3) *it is worth noting that, in the case of linear programming, this implication can be strengthened:* ($\mathrm{F}_p$) *and* ($\mathrm{B}_p$) $\Rightarrow$ ($\mathrm{S}_d$) *(since, in the case of* LP *we have* ($\mathrm{F}_p$) *and* ($\mathrm{B}_p$) $\Rightarrow$ ($\mathrm{S}_p$) $\Leftrightarrow$ ($\mathrm{S}_d$)). *In the nonlinear case, it seems to be difficult to strengthen* (3), *since it can happen that* ($\mathrm{F}_p$) *and* ($\mathrm{B}_p$) *and* ($\rceil$,$\mathrm{S}_d$).

The following result seems to be closest to the LP Duality Theorem.

(5) ($\mathrm{I}_p$) and ($\mathrm{BO}_p$) $\Rightarrow$ ($\mathrm{S}_d$) and (WN); ($\mathrm{I}_d$) and ($\mathrm{BO}_d$) $\Rightarrow$ ($\mathrm{S}_p$) and (WN).

Under the premise of the first implication, the function $\mathsf{P}^*(\cdot)$ is finite at the point $b$ and $b \in \operatorname{int} B$. Since $E$ is finite-dimensional and $\mathsf{P}^*(\cdot)$ is convex on the convex set $B$, it follows that the function is finite and continuous at least on $\operatorname{int} B$ and that it admits support functionals at $b$. It remains to use Proposition 4.2.1. □

To make the references more convenient, let us present (5) in "verbal" form.

**Theorem 4.2.1 (Duality Theorem)** *Let* ($\mathsf{P}$), ($\mathsf{D}$) *be a primal-dual pair of conic problems defined by the data $K$, $L$, $c$, $b$ and let the pair be such that the* *primal feasible set intersects* $\operatorname{int} K$ *and the objective of the primal problem is below bounded on the primal feasible set. Then the dual problem is solvable, and the optimal values of the primal problem* $\mathsf{P}^*$ (min *or* inf *of the primal objective over the primal feasible set) and the dual one* $\mathsf{D}^*$ *satisfy the relation*

$$\mathsf{P}^* + \mathsf{D}^* = \langle c, b\rangle \tag{4.2.6}$$

(*"zero duality gap"*).

*If, in addition, the dual feasible set intersects* $\operatorname{int} K^*$, *then both of the problems are solvable,* (4.2.6) *holds, and a pair of feasible primal and dual solutions* $(x^*, s^*)$ *is comprised of optimal solutions to the problems if and only if*

$$\langle s^*, x^*\rangle = 0 \tag{4.2.7}$$

(*"complementary slackness condition"*).

The first part of the theorem is item (5). The second part is an immediate consequence of the first part and (4.2.2). □

(6) ($\mathrm{S}_p$) and ($\mathrm{I}_p$) $\Rightarrow$ (N); ($\mathrm{S}_d$) and ($\mathrm{I}_d$) $\Rightarrow$ (N).

This is an immediate corollary of (5).

Note that the premise in (6) is precisely the standard assumption that justifies the use of the Kuhn–Tucker optimality condition in the above motivation of the "conic" duality.

The following statement extends implication (3).

(7) ($\mathrm{SB}_p$) $\Rightarrow$ (WN); ($\mathrm{SB}_d$) $\Rightarrow$ (WN).

Before proving this implication, we should mention that its conclusion, generally speaking, cannot be strengthened to (N).

To prove the first implication in (7), assume that the set $X^*$ of optimal solutions of the primal problem is bounded and nonempty and let $x_0 \in X^*$ and $x_1 \in \operatorname{int} K$. Without loss of generality, we can take $b = x_0$. Indeed, this substitution does not vary the primal problem and updates only the objective of the dual problem. The new objective at the dual feasible set is the old one plus a constant. Clearly, the initial and the updated primal-dual pairs either both possess, or both do not possess, property (WN).

For $\varepsilon \in (0,1)$, let ($\mathsf{P}_\varepsilon$), ($\mathsf{D}_\varepsilon$) be the primal-dual pair defined by the same data as ($\mathsf{P}$), ($\mathsf{D}$), excluding the value of $b$, the latter quantity now being replaced by $x(\varepsilon) = x_0 + \varepsilon x_1$. Note that ($\mathsf{P}_0$), ($\mathsf{D}_0$) clearly is the initial pair, while, for $\varepsilon > 0$, problem ($\mathsf{P}_\varepsilon$) possesses property (I) (since $x(\varepsilon)$ is a feasible interior solution to the problem). Let us prove that, for all small enough $\varepsilon > 0$, problem ($\mathsf{P}_\varepsilon$) is solvable. Otherwise, we could find a sequence $\{\varepsilon_i > 0\}$ tending to 0 and such that each of the problems ($\mathsf{P}_{\varepsilon_i}$) would admit an unbounded minimizing sequence with the first element $x(\varepsilon_i)$, so that we could find a sequence $\{x_i \in K\}$, $\| x_i \| \to \infty$, $i \to \infty$, with $x_i$ being a feasible solution of ($\mathsf{P}_{\varepsilon_i}$) such that $\langle c, x_i\rangle \le \langle c, x(\varepsilon_i)\rangle$. Without loss of generality, we can assume that $x_i = t_i e_i + x(\varepsilon_i)$, $t_i \to \infty$, where $e_i \to e \neq 0$ as $i \to \infty$. For each positive $t$ and all large enough $i$, the points $x(\varepsilon_i) + t e_i$ belong to $K$, together with $x(\varepsilon_i)$ and

$x_i$, and therefore the point $x_0 + te$ also belongs to $K$. Since $x_i$ and $x(\varepsilon_i)$ are feasible solutions of $(\mathsf{P}_{\varepsilon_i})$, we have $e_i \in L$ and, consequently, $e \in L$. Last, $\langle c, x_i \rangle \le \langle c, x(\varepsilon_i)\rangle$ implies $\langle c, e_i \rangle \le 0$, and therefore $\langle c, e \rangle \le 0$. We see that the ray $\{x_0 + te \mid t \ge 0\}$ is feasible for $(\mathsf{P})$, and the objective is nonincreasing along this ray. It follows that this ray is contained in the set of optimal solutions to $(\mathsf{P})$, which contradicts the boundedness of the latter set. Note that we have proved something more strong than the solvability of $(\mathsf{P}_\varepsilon)$ for all small enough $\varepsilon$; in fact, we have established that there exist a bounded set $Q$ and a positive $\varepsilon_0$, such that all the sets

$$Q_\varepsilon = \{x \mid x \text{ is feasible for } (\mathsf{P}_\varepsilon) \text{ and } \langle c, x\rangle \le \langle c, x(\varepsilon)\rangle\}$$

for $\varepsilon \le \varepsilon_0$ are contained in $Q$. From the latter statement, it immediately follows that, if $\mathsf{P}^*_\varepsilon$ is the optimal value in $(\mathsf{P}_\varepsilon)$, then $\liminf_{\varepsilon\to+0} \mathsf{P}^*_\varepsilon \ge \mathsf{P}^*$. Also, since

$$\limsup_{\varepsilon\to+0} \mathsf{P}^*_\varepsilon \le \lim_{\varepsilon\to 0} \langle c, x(\varepsilon)\rangle = \langle c, x_0\rangle = \mathsf{P}^*,$$

we conclude that $\mathsf{P}^*_\varepsilon \to \mathsf{P}^*$ as $\varepsilon \to 0$.

For all small enough $\varepsilon > 0$, problems $(\mathsf{P}_\varepsilon)$ possess property (I) (since $x(\varepsilon) \in \operatorname{int} K$, $\varepsilon > 0$) and, as we already have proved, property (S). By virtue of item (6), it follows that, for the above $\varepsilon$, problems $(\mathsf{D}_\varepsilon)$ are solvable, and their solutions $s(\varepsilon)$ satisfy the relation

$$\mathsf{P}^*_\varepsilon + \langle s(\varepsilon), x(\varepsilon)\rangle = \langle c, x(\varepsilon)\rangle\,. \tag{4.2.8}$$

The feasible sets of all the problems $(\mathsf{D}_\varepsilon)$, $\varepsilon \ge 0$ coincide, so that all $s(\varepsilon)$ are feasible for $(\mathsf{D})$. Since $x(\varepsilon) - x_0 \in K$ and $s(\varepsilon) \in K^*$, we have $\langle s(\varepsilon), x_0\rangle \le \langle s(\varepsilon), x(\varepsilon)\rangle$, and (4.2.8) implies $\mathsf{D}^* \le \langle s(\varepsilon), x_0\rangle \le \langle c, x(\varepsilon)\rangle - \mathsf{P}^*_\varepsilon$. The right-hand side of the latter inequality tends to $\langle c, x_0\rangle - \mathsf{P}^*$ as $\varepsilon \to +0$, which results in $\mathsf{D}^* + \mathsf{P}^* \le \langle c, x_0\rangle$. Since we have reduced the situation to the case where $x_0 = b$, the concluding relation combined with (4.2.2) proves that the pair $((\mathsf{P}),(\mathsf{D}))$ possesses property (WN). □

To conclude this section, let us note that, in the nonlinear case, even solvability of the both primal and dual problems does not ensure (N).

## 4.3 Karmarkar method for nonlinear problems

### 4.3.1 Formulation of the problem. Assumptions

Consider a conic problem $(\mathsf{P})$. For our purposes, it is convenient to represent the feasible plane $(L + b)$, which is an affine subspace, as an intersection of a linear subspace $M$ in $E$ and an affine hyperplane *not passing through the origin.* Of course, it can be done quite straightforwardly, provided that $(L + b)$ does not contain the origin. The case when the latter condition is not satisfied can be excluded, since, in this case, the feasible set $K \bigcap (L + b)$ is a cone, so that the problem is either below unbounded or has trivial optimal solution 0. Thus, in this section (as well as in the next), we deal with the problem

$$\langle c, x\rangle \to \min \mid x \in K \bigcap M,\ \langle e, x\rangle = 1, \tag{4.3.1}$$

where $c$, $e \in E^*$, $K$ is a closed convex pointed cone in $E$ with a nonempty interior and $M$ is a linear subspace in $E$. In the case of $K = \mathbf{R}^n_+$, (4.3.1) is precisely the Karmarkar format of an LP problem.

Assume that we are given a $\vartheta$-self-concordant logarithmically homogeneous barrier $F$ for the cone $K$. This is the only representation of $K$ used by the below method.

In the case of $K = \mathbf{R}^n_+$, we could take as $F$ the standard logarithmic barrier

$$F(x) = -\sum_{i=1}^{n} \ln x_i \tag{4.3.2}$$

for the nonnegative orthant ($\vartheta = n$; see §2.3, Example 2).

In what follows, we describe the (generalized) method of Karmarkar for (4.3.1). This method, as in the LP case, requires the following three additional assumptions about the problem:

(K.1) The feasible set $K_e = \{x \in K \bigcap M \mid \langle e, x\rangle = 1\}$ is bounded and intersects $\operatorname{int} K$;

(K.2) An interior feasible solution, $x_0 \in \operatorname{rint} K_e$, is known (as usual, "rint" denotes the relative interior);

(K.3) It is known that the optimal value of the objective is 0.

Note that (K.3), in fact, is no more than the assumption that we know the optimal value of the objective; indeed, if this value is $f^*$, then, replacing the objective $c$ by $c - f^*e$, we obtain an equivalent problem with zero optimal value.

Note also that, under assumptions (K.1)–(K.3), we could solve (4.3.1) with the aid of a path-following method from Chapter 3, with the preliminary stage of the latter method being eliminated. Indeed, let us replace the functional $e$ involved into (4.3.1) by the functional $e' = -\vartheta^{-1}F'(x_0)$. The resulting problem can be rewritten as

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

subject to

$$x \in G \equiv K \bigcap M \bigcap \{y \mid \langle F'(x_0), y - x_0\rangle = 0\}. \tag{4.3.3}$$

The restriction $F^{\#}$ of $F$ onto the (relative) interior of $G$ is a $\vartheta$-self-concordant barrier for $G$, and $x_0$ is the minimizer of $F^{\#}$, so that we can immediately apply to (4.3.3) the main stage of the path-following method associated with $F^{\#}$. On the other hand, it is easily seen that the optimal value in (4.3.3), the same as in (4.3.1), is zero, and the transformation $x \to \langle e, x\rangle^{-1} x$ maps strictly feasible solutions to (4.3.1) into strictly feasible solutions to (4.3.1) of the same (up to a constant factor depending on the problem) accuracy.

### 4.3.2 Generalized Karmarkar method

The method is associated with the potential function

$$V(x) = \vartheta \ln \langle c, x \rangle + F(x) : \operatorname{int} K \to \mathbf{R}; \tag{4.3.4}$$

note that, since $F$ is logarithmically homogeneous of the degree $\vartheta$, we have

$$V(tx) = V(x), \quad x \in \operatorname{int} K, \quad t > 0. \tag{4.3.5}$$

Let us show that the accuracy of a strictly feasible (i.e., feasible and belonging to $\operatorname{int} K$) solution $x$ can be evaluated via the quantity $V(x_0)-V(x)$. Indeed, since the feasible set $K_e$ of the problem is bounded and the restriction of $F$ onto $K_e$ is a self-concordant barrier for this set (see Proposition 2.3.1(i)), $F$ is below bounded on the relative interior of $K_e$ (Proposition 2.3.2(ii)). Therefore the quantity

$$R(x_0) = F(x_0) - \min\{F(u) \mid u \text{ is strictly feasible}\}$$

is finite. We have

$$V(x_0) - V(x) = F(x_0) - F(x) + \vartheta \ln \frac{\langle c, x_0 \rangle}{\langle c, x \rangle} \leq R(x_0) + \vartheta \ln \frac{\langle c, x_0 \rangle}{\langle c, x \rangle},$$

so that we have proved the following result.

**Proposition 4.3.1** *Let $x$ be a strictly feasible solution to* (4.3.1). *Then*

$$\langle c, x \rangle \leq \langle c, x_0 \rangle \, R(x_0) \exp\left\{ -\frac{1}{\vartheta}(V(x_0) - V(x)) \right\},$$

$$R(x_0) = F(x_0) - \min\{F(u) \mid u \ \textit{is strictly feasible}\}.$$

Note that the inequality in Proposition 4.3.1 is an accuracy estimate, since the optimal value of the objective is zero.

Proposition 4.3.1 demonstrates that an updating rule that transforms a given strictly feasible solution into a new solution of the same type with by an absolute constant less the value of the potential, being iterated, reduces the errors $\langle c, x_i \rangle$ of the successively updated solutions in the ratio $1 - O(\vartheta^{-1})$. The idea of this rule is as follows. Let $x$ be a strictly feasible solution to be updated. Consider the hyperplane $N$ tangent at $x$ to the corresponding level set of the barrier and let $S$ be the intersection of this plane with $M \bigcap K$. $S$ is a convex set, and the reduction of $F$ onto the relative interior $S'$ of $S$ is a $\vartheta$-self-concordant barrier for $S$ (Proposition 2.3.1(i)). Let us try to find a point $x'$ in $S'$ with the value of the potential being "considerably" less than at $x$. To this end, consider the function

$$\Phi(u) = F(u) + \vartheta \frac{\langle c, u - x \rangle}{\langle c, x \rangle} + \vartheta \ln \langle c, x \rangle, \tag{4.3.6}$$

i.e., the result of linearization of the logarithmic term of the potential at $x$; we regard $\Phi$ as a function defined on $S'$. Since the logarithmic term is concave, we have $F(u) \leq \Phi(u)$, while $F(x) = \Phi(x)$. We see that, to decrease $F$, it suffices to decrease $\Phi$. On the other hand, $\Phi$ is the sum of a self-concordant barrier and a linear function, and therefore it is strongly 1-self-concordant on $S'$. Let us demonstrate that the Newton decrement $\lambda(\Phi, x)$ of $\Phi$ at $x$ *is not small*, so that, to decrease $\Phi$ "substantially," it suffices to perform, starting from $x$, a step of the Newton-type process described in Proposition 2.2.2. Indeed, $F$ has zero derivative at $x$ in each direction from $S - x$ (the origin of $S$), so that

$$(4.3.7) \qquad \lambda(\Phi, x) = \frac{\vartheta}{\langle c, x\rangle} \max\{\langle c, h\rangle \mid h \in \operatorname{Lin}\{S - x\},\ \langle F''(x)h, h\rangle \leq 1\}.$$

To prove that $\lambda(\Phi, x)$ is not small, note that $K$ contains the ray $\{tx^* \mid t > 0\}$, where $x^* \in M \backslash \{0\}$ is such that $\langle c, x^*\rangle = 0$ (indeed, in view of (K.3), we can take as $x^*$ the optimal solution to (4.3.1)). This ray intersects $S$. Indeed, $F$ attains its minimum over the intersection of $N$ and $K$ at $x$ (the origin of $N$) and is a self-concordant barrier for this set (Proposition 2.3.1(i)), so that this intersection is bounded (Proposition 2.3.2(ii)). It means that $N$ intersects with all rays $\{ty \mid t > 0\}$, $y \in K \backslash \{0\}$ and, in particular, with the ray $\{tx^* \mid t > 0\}$. Since the latter ray is also contained in $M$, it intersects $S$.

Since, as we have seen, $x$ is the $F\,|_{S'}$-center of $S$, the latter set is contained in the ellipsoid

$$W' = \{y \in x + \operatorname{Lin}\{S - x\} \mid \langle F''(x)(y - x), (y - x)\rangle \leq (1 + 3\vartheta)^2\}$$

(see (2.3.8)). The linear functional $c_x = \vartheta \langle c, x\rangle^{-1} c$ equals $\vartheta$ at $x$ and equals 0 at certain point $u^* \in S$, namely, at the point where the ray $\{tx^* \mid t > 0\}$ intersects $S$. Thus, on the $(1 + 3\vartheta)$-enlargement, $W' - x$, of the ellipsoid

$$W = \{h \in \operatorname{Lin}\{S - x\} \mid \langle F''(x)h, h\rangle \leq 1\}$$

involved into (4.3.7), $c_x$ varies by at least $2\vartheta$; hence, on $W$ it varies at least by $2\vartheta(1 + 3\vartheta)^{-1}$, and we conclude that

$$(4.3.8) \qquad \lambda(\Phi, x) \geq \tfrac{1}{4}.$$

Now let $x'$ be the iterate of $x$ with respect to the Newton-type process described in Proposition 2.2.2 associated with $a = 1$, so that

$$x' = x + \frac{\lambda(\Phi, x)}{1 + \lambda(\Phi, x)}\zeta,$$

$$(4.3.9) \qquad \zeta = \frac{\vartheta}{\langle c, x\rangle} \operatorname{argmin}\{\langle c, h\rangle \mid h \in W\}.$$

From (4.3.8) and Proposition 2.2.2, it follows that

$$x' \in S',$$

$$V(x) - V(x') \geq \Phi(x) - \Phi(x') \geq \kappa, \qquad \kappa = \tfrac{1}{4} - \ln \tfrac{5}{4} \tag{4.3.10}$$

(the first inequality in (4.3.10) follows from $V(\cdot) \leq \Phi(\cdot)$, $V(x) = \Phi(x)$).

Thus, we have found a point $x' \in S'$ with the value of the potential being "considerably" less than at $x$. Note that $x' \in \operatorname{int} K \bigcap M$, since $S'$ is contained in the latter set. Thus, the only possibility for $x'$ not to be a strictly feasible solution to (4.3.1) is to violate the normalizing constraint $\langle e, \cdot \rangle = 1$. Let us believe for a moment that $\langle e, x' \rangle > 0$ (we prove this later). Let us take $x^+ = \langle c, x' \rangle^{-1} x'$ as the desired updating of $x$. This point belongs to $(\operatorname{int} K) \bigcap M$, since $x'$ belongs to the latter set, $\langle c, x' \rangle > 0$, and evidently satisfies the normalizing constraint. At the same time, in view of (4.3.5), we have $V(x^+) = V(x')$ and, in view of (4.3.10), we obtain

$$x^+ \in \operatorname{rint} K_e,$$

$$V(x) - V(x') \geq \kappa = \tfrac{1}{4} - \ln \tfrac{5}{4}. \tag{4.3.11}$$

It remains to verify that $\langle e, x' \rangle$ is positive. Assume that it is nonpositive; since $x$ is feasible, we have $\langle e, x \rangle = 1$, so that there exists $\theta \in (0, 1]$ such that $\gamma(t) = \langle e, x + t(x' - x) \rangle$ is positive for $0 \leq t < \theta$ and is zero for $t = \theta$. Note that the segment $[x, x']$ is contained in $(\operatorname{int} K) \bigcap M$. It follows that the (clearly unbounded) curve $\gamma^{-1}(t)(x + t(x' - x))$, $0 \leq t < \theta$ is contained in the feasible set of (4.3.1), which is impossible in view of (K.1).

The analytical description of the method is as follows. At $i$th step of the method ($i \geq 1$), the previous approximate solution $x_{i-1} \in \operatorname{rint} K_e$ ($x_0$ is the point mentioned in (K.2)) is transformed into a new strictly feasible solution $x_i$ according to the following rules:

($\mathcal{K}$.1) Compute

$$e_i = F'(x_{i-1}) \in E^*,$$

$$c_i = \vartheta \frac{c}{\langle c, x_i \rangle}$$

and set

$$L_i = \{x \in M \mid \langle e_i, x \rangle = 0\};$$

($\mathcal{K}$.2) Compute the Newton direction of the functional $c_i$ in the subspace $L_i$ with respect to the Euclidean structure defined by $D^2F(x_{i-1})$, namely, set

$$\zeta_i \in \operatorname{Argmin}\{\langle c_i, h \rangle + \tfrac{1}{2} \langle F''(x_{i-1})h, h \rangle \mid h \in L_i\},$$

$$\lambda_i = \langle F''(x_{i-1})\zeta_i, \zeta_i \rangle^{1/2},$$

$$x_i' = x_{i-1} + \frac{\zeta_i}{1 + \lambda_i};$$

($\mathcal{K}$.3) Find $x_i'' \in (x_{i-1} + L_i) \bigcap (\operatorname{int} K)$ such that

$$V(x_i'') \leq V(x_i') \tag{4.3.12}$$

and set

$$x_i = \frac{x_i''}{\langle e, x_i''\rangle}.$$

The $i$th step is completed.

**Remark 4.3.1** (1) *As we have seen, $x_i' \in \operatorname{int} K$, so that an admissible choice of $x_i''$ is $x_i'' = x_i'$; the version based on this choice is called basic.*

(2) *The simplest way to perform "large steps" is to choose $x_i''$ by minimizing the potential in the direction $\zeta_i$:*

$$x_i'' = \operatorname{argmin}\{V(x_{i-1} + t\zeta_i) \mid t \geq 0,\ x_{i-1} + t\zeta_i \in \operatorname{int} K\}.$$

### 4.3.3 Rate of convergence

The convergence properties of the above method can be derived immediately from Proposition 4.3.1, (4.3.11), and (4.3.12) and are as follows.

**Theorem 4.3.1** *Assume that* (K.1)–(K.3) *are satisfied. Then the generalized Karmarkar method associated with a $\vartheta$-self-concordant logarithmically homogeneous barrier for $K$ produces a sequence $\{x_i\}$ of strictly feasible solutions to* (4.3.1), *and the relative accuracy of these solutions can be estimated as*

$$\text{(4.3.13)} \quad \frac{\langle c, x_i\rangle}{\langle c, x_0\rangle} \leq R(x_0)\exp\left\{-\frac{1}{\vartheta}(V(x_i) - V(x_0))\right\} \leq R(x_0)\exp\{-\kappa i\},$$

*where*

$$\kappa = \tfrac{1}{4} - \ln\tfrac{5}{4}, \quad R(x_0) = F(x_0) - \min\{F(x) \mid x \in \operatorname{rint} K_e\}.$$

The factor $R(x_0)$ is responsible for the quality of the initial approximate solution $x_0$. Note that, in the basic version of the Karmarkar method for an LP problem (the latter is precisely the above method associated with $K = \mathbf{R}^n_+$, the $n$-self-concordant barrier (4.3.2)), $e = (1, ..., 1)^T$, and $x_0 = n^{-1}e$), we have $R(x_0) = 1$.

### 4.3.4 Karmarkar method and projective transformations

Let us present another interpretation of the above method (for the LP case, it was found by Bayer and Lagarias; see [BL 91]). For simplicity, let us restrict ourselves to the "large-step" version. From the description of the updating rule, it follows immediately that, in fact, the input and the output of the rule are not strictly feasible solutions, but *strictly feasible rays*, i.e., rays of the type $\{tu \mid t > 0\}$ associated with $u \in (\operatorname{int} K)\bigcap M$. These rays are in a one-to-one correspondence with strictly feasible solutions (such a solution evidently defines a strictly feasible ray, and, as we have seen, the boundedness of the feasible set implies that each strictly feasible ray intersects the "normalizing hyperplane"

$\{x \mid \langle e, x\rangle = 1\}$ and therefore corresponds to a strictly feasible solution). The potential is constant along the rays and therefore can be thought of as the function defined on the space of rays. The updating rule looks as follows: Given a strictly feasible ray $\xi = \{tx \mid t > 0\}$, we consider the two-dimensional angle $U$, which is the intersection of $K$ and the two-dimensional subspace $M_\xi$ passing through $\xi$ and the one-dimensional subspace

$$\eta(\xi) = \operatorname{argmin}\{\langle c, h\rangle + \tfrac{1}{2}\langle F''(u)h, h\rangle \mid h \in M,\ \langle F'(u), h\rangle = 0\} \cdot \mathbf{R},$$

where $u$ is some point of $\xi$ (since $F$ is logarithmically homogeneous, the subspace does not depend on choice of $u \in \xi$). Then we minimize $V$ over the rays from the relative interior of $U$, and this gives us a new strictly feasible ray $\xi^+(\xi)$, which is the result of the updating. In the initial description of the updating rule, we represent $\xi$ by a point $v \in \xi$ and choose as the representative of a ray $\gamma \subset \operatorname{rint} U$ the point at which $\gamma$ intersects the interval $\Delta = \{u \in \operatorname{rint} U \mid \langle F(v), u - v\rangle = 0\}$, so that to minimize $V$ over $U$ is the same as to minimize this function over $\Delta$. Now note that we could choose representatives of rays $\gamma \in \operatorname{rint} U$ in many other ways, e.g., as points at which these rays intersect the level hyperplane

$$C = \{u \mid \langle c, u\rangle = 1\}$$

of the objective. Thus, minimizing $V$ over the rays belonging to $\operatorname{rint} U$ is the same as minimizing this function over the interval $\Delta' = C \bigcap \operatorname{rint} U$. On this interval, however, $V$ coincides with $F$. Let $w$ be the point at which $\xi$ intersects $\Delta'$. It is not difficult to verify that the direction of $\Delta'$ is precisely the Newton direction of $F$ in the affine subspace $M \bigcap C$ (the direction is taken at $w$),

$$\operatorname{Lin}\{\Delta' - w\} = \operatorname{argmin}\{\langle F'(x), h\rangle + \tfrac{1}{2}\langle F''(x)h, h\rangle \mid h \in M, \langle c, h\rangle = 0\} \cdot \mathbf{R}$$

(to verify the latter relation, it suffices to use the fact that the Newton direction of $F$ at $w$ with respect to whole $E$ is collinear to $w$; see (2.3.12)).

Thus, we can describe the updating rule as follows. Let us identify strictly feasible rays and their intersections with the hyperplane $C$; thus, the set of strictly feasible rays corresponds to the relative interior of the closed convex set $T = C \bigcap M \bigcap K$. The restriction of the potential onto $\operatorname{rint} T$ coincides with the restriction of the barrier onto the set and therefore is a self-concordant barrier for $T$. The method of Karmarkar is simply the Newton minimization method as applied to $F\,|_{\operatorname{rint}}$: At each step, the Newton direction at the current point is found, and then $F$ is minimized in this direction (recall that we consider the method with large steps). The fact that $V = F$ decreases "substantially" at each step is an immediate corollary of the fact that $T$ is unbounded (in turn, the latter property is implied by the fact that the optimal value of the objective is 0). Since $T$ is unbounded and $V = F$ is a self-concordant barrier for $T$, this function is below unbounded on $\operatorname{rint} T$ (see Proposition 2.3.2(ii); recall that $T$ is contained in $K$ and therefore does not contain straight lines). Since $V$ is below unbounded, we have $\lambda(V, x) \geq 1$ for all $x \in \operatorname{rint} T$ (see Theorem 2.2.3(i)). The latter inequality, in view of Proposition 2.2.2, implies that each Newton step decreases $V$ by at least $(1 - \ln 2)$.

The summary of the above explanation of the method is the following. In the initial problem, we are given a bounded closed convex feasible domain $G$ and a hyperplane $\Pi$, which is known to be support to $G$ ($\Pi$ is the level hyperplane of the objective corresponding to its optimal value). The problem is to find a close to $\Pi$ point in $G$. To solve this problem, in Karmarkar method, we perform the projective transformation that moves $\Pi$ to infinity; this transformation maps $G$ onto an *unbounded* closed convex domain $G^+$ (in our conic representation, $G$ is $K_e$ and $G^+$ is $T$), and in the transformed space the problem is to find a point of $G^+$ far enough from the origin. To this purpose, we minimize by the Newton method a self-concordant barrier for $G^+$ (in our representation, this barrier is the reduction of $F$ onto $\operatorname{rint} T$). Each step of the method decreases the barrier by an absolute constant, and, since the barrier is below bounded on each bounded subset of $\operatorname{int} G^+$, the corresponding sequence does go to infinity at a rate that ensures the polynomiality of the method.

### 4.3.5 Case of unknown optimal value

From the viewpoint of applications, it might be difficult to provide (K.2) (to point out an initial strictly feasible solution) and especially (K.3) (to point out the optimal value of the objective). It is known how to avoid these difficulties in the case of LP; the same tricks can be used in the general conic case. Let us discuss two possibilities of this type.

*Primal-dual reformulation of the problem.* First, we may simultaneously solve the primal and the dual problems, i.e., to add to the initial (primal) conic problem

$$\text{minimize } \langle c, x\rangle \quad \text{s.t. } x \in K \bigcap (L + b)$$

the dual problem

$$\text{minimize } \langle s, b\rangle \quad \text{s.t. } s \in K^* \bigcap (L^\perp + c).$$

Under the assumption that the feasible set of the primal problem is bounded and intersects $\operatorname{int} K$, this pair of problems in view of the duality theorem (Theorem 4.2.1) is equivalent to the conic problem in the space $E \times E^*$,

$$(\mathsf{PD}): \qquad \text{minimize } \langle c, x\rangle + \langle s, b\rangle \quad \text{s.t. } (x, s) \in (K \times K^*) \bigcap (L \times L^\perp + (b, c))$$

in the sense that optimal solutions to the latter problem are precisely the pairs comprised of optimal solutions to the primal and to the dual problem. It is important for us that the Duality Theorem also states that, under the above assumption, (PD) is solvable and the optimal value of the problem equals $\langle c, b\rangle$. Thus, duality reduces the initial problem to an equivalent conic problem with known optimal value of the objective, so that, for the resulting problem, there are no difficulties with (K.3).

Of course, to solve (PD) with the aid of the method of Karmarkar, we need a logarithmically homogeneous self-concordant barrier for the cone $K \times K^*$. The latter requirement can be easily satisfied if we know a $\vartheta$-logarithmically homogeneous barrier $F$ for $K$ and the Legendre transformation $F^*$ of this barrier. Indeed, by virtue of Theorem 2.4.4, $F^*$ is a $\vartheta$-logarithmically homogeneous self-concordant barrier for the cone $(-K^*)$ anti-dual to $K$, so that $F^+(s) = F^*(-s)$ is the barrier of the same type for $K^*$, and, of course, $F(x) + F^+(s)$ is a $(2\vartheta)$-logarithmically homogeneous self-concordant barrier for $K \times K^*$.

The only difficulty now is to provide (K.2) for (PD). This can be done with the aid of the same Phase 0 as in the LP case. Namely, to find a strictly feasible solution to a conic problem (C) with the feasible set $Q \bigcap (M + d)$ ($Q$ is the cone in certain finite-dimensional space $H$; $M$ is a linear subspace of $H$), we can choose a point $z_0 \in \operatorname{int} Q$ and form the auxiliary conic problem

$$\text{minimize} \quad t \quad \text{s.t. } (t, z) \in (\mathbf{R}_+ \times Q), z - tz_0 \in (M + (1-t)d).$$

If the initial problem is consistent, then clearly the optimal value in the latter problem is 0, while the point $(1, z_0)$ is a strictly feasible solution to it. If we know a $\nu$-logarithmically homogeneous self-concordant barrier $\Phi$ for $Q$, then the function $-\ln t + \Phi(z)$ is a $(\nu + 1)$-logarithmically homogeneous self-concordant barrier for $\mathbf{R}_+ \times Q$ (Proposition 2.3.1(iii) combined with Example 1 of §2.3), and the latter problem can be solved by the basic version of the method. We can terminate the iterations at the first moment when the open ellipsoid $W = \{u \mid \langle \Phi''(z)(u - z), u - z\rangle < 1\}$ associated with $z$-component of current approximate solution intersects $M + d$. Indeed, this ellipsoid is contained in $\operatorname{int} Q$ (Theorem 2.1.1(ii)), and therefore each point of the intersection $(M + d) \bigcap W$ can be taken as the desired strictly feasible solution to (C).

### 4.3.6 Sliding objective approach

Now let us describe the purely primal technique that allows us to avoid (K.3). The advantage of this technique is that we do not require any knowledge of a barrier for the dual cone. As far as (K.2) is concerned, we assume that this condition is satisfied (otherwise, we could apply to (4.3.1) the above Phase 0). Thus, we assume that (K.1) and (K.2) are satisfied, and, instead of (K.3), we introduce a more realistic assumption

(K.3′) We are given an a priori lower bound $\sigma$ for the optimal value of the objective.

Of course, we can also assume that the objective is not constant on the feasible set of (4.3.1).

Under these assumptions, we can modify the method as follows. At the $i$th step of the method, we have *the current objective*

$$c(i-1) = c - t_{i-1}e \qquad (4.3.14)$$

such that the optimal value of this linear functional on the feasible set of (4.3.1) is nonnegative. At the first step, the latter property can be provided by the choice $t_0 = \sigma$ (see (K.3$'$)).

The current objective generates the *current potential*

$$V_{i-1}(x) = F(x) + \vartheta \ln \langle c(i-1), x \rangle .$$

Let $(\mathsf{P}(d))$ denote the problem that is obtained from (4.3.1) by replacing the objective by $d$ and let $x^+ = \mathcal{K}(d; x)$ denote the updating of a strictly feasible solution $x$ into a new strictly feasible solution $x^+$ defined by rules ($\mathcal{K}$.1)–($\mathcal{K}$.3) as applied to problem $(\mathsf{P}(d))$. From the above analysis of a step, it follows that the updating is well defined, provided that the optimal value of the objective is nonnegative.

At the $i$th step, we first try to update $x_{i-1}$ with the aid of $\mathcal{K}(c(i-1), \cdot)$. In particular, we compute the quantity $\lambda_i = \lambda(x_{i-1}, c(i-1))$, which, as we know, is the Newton decrement (at $x_{i-1}$) of the function

$$\Phi_{x_{i-1}}^{c(i-1)}(u) = F(u) + \vartheta \frac{\langle c(i-1), u - x_{i-1} \rangle}{\langle c(i-1), x_{i-1} \rangle}$$

regarded as a function defined on $\operatorname{rint} S_i$, where

$$S_i = K \bigcap \{x \in M \mid \langle F'(x_{i-1}), x - x_{i-1} \rangle = 0\}.$$

It may happen that

$$\lambda(x_{i-1}, c(i-1)) \geq \tfrac{1}{4}; \tag{4.3.15}$$

in this case, we set

$$c(i) = c(i-1), \qquad x_i = \mathcal{K}(c(i), x_{i-1}) \tag{4.3.16}$$

and go to the next step. In the opposite case, we choose the largest $\tau = \tau_i > 0$ satisfying the relations

$$\lambda(x_{i-1}, c(i-1) - \tau e) \leq \tfrac{1}{4}, \qquad \langle c(i-1) - \tau e, x_{i-1} \rangle > 0. \tag{4.3.17}$$

It is easily seen that $\tau_i$ is well defined and

$$\lambda(x_{i-1}, c(i-1) - \tau_i e) = \tfrac{1}{4}. \tag{4.3.18}$$

Indeed, let $\tau_i^*$ be the minimal value of the current objective $c(i-1)$ on the feasible set of (4.3.1). Note that this value is nonnegative (by virtue of our assumption about $c(i-1)$). The optimal value of the objective $c(i-1) - \tau_i^* e$ on the feasible set of (4.3.1) is zero, and this objective is not constant on the latter set, so that $\langle c(i-1) - \tau_i^* e, x_{i-1} \rangle > 0$ and $\lambda(x_{i-1}, c(i-1) - \tau^* e) \geq \frac{1}{4}$ (see (4.3.8)). At the same time, if $\tau > \tau_i^*$ is such that $\langle c(i-1) - \tau e, x_{i-1} \rangle > 0$ and the optimal value of the objective $c(i-1) - \tau e$ on the feasible set of (4.3.1) is negative, then the arguments used in the proof of (4.3.8) demonstrate that $\lambda(x_{i-1}, c(i-1) - \tau e) > \frac{1}{4}$. Thus, the maximal $\tau$ satisfying (4.3.17) does exist, satisfies (4.3.18), and belongs to $[0, \tau_i^*]$.

After $\tau_i$ is chosen, we set

$$c(i) = c(i-1) - \tau_i e, \qquad x_i = \mathcal{K}(c(i), x_{i-1}) \tag{4.3.19}$$

and go to the next step. Note that, in view of $\tau_i \in [0, \tau_i^*]$, the optimal value of $c(i)$ on the feasible set of (4.3.1) still is nonnegative.

Let us verify that, for the modified method, the accuracy estimate similar to (4.3.13) holds, namely,

$$\frac{\langle c, x_i\rangle - c^*}{\langle c, x_0\rangle - \sigma} \leq R(x_0) \exp\left\{-\frac{1}{\vartheta}(V_i(x_i) - V_0(x_0))\right\} \leq R(x_0)\exp\{-\kappa i\}, \tag{4.3.20}$$

where

$$\kappa = \tfrac{1}{4} - \ln \tfrac{5}{4}, \qquad R(x_0) = F(x_0) - \min\{F(x) \mid x \in \operatorname{rint} K_e\},$$

and $c^*$ denotes the optimal value in (4.3.1).

Let us verify first that

$$V_i(x_i) \leq V_{i-1}(x_{i-1}) - \kappa, \qquad \kappa = \tfrac{1}{4} - \ln\tfrac{5}{4}. \tag{4.3.21}$$

Indeed, we have

$$V_i(x_i) \leq V_i(x_{i-1}) - \kappa, \tag{4.3.22}$$

since the only property (additional to the nonnegativity of the optimal value of the objective) used in the proof of (4.3.11) was the relation that in our notation is now $\lambda(x_{i-1}, c(i)) \geq \frac{1}{4}$, and the modified method does maintain this relation. Furthermore, either $V_{i-1}(\cdot) \equiv V_i(\cdot)$, and then (4.3.21) is equivalent to (4.3.22), or $c(i) = c(i-1) - \tau_i e$ with positive $\tau_i$. In the latter case,

$$V_i(x_{i-1}) = F(x_{i-1}) + \vartheta \ln(\langle c(i-1) - \tau_i e, x_{i-1}\rangle)$$

$$< F(x_{i-1}) + \vartheta\ln(\langle c(i-1), x_{i-1}\rangle) = V_{i-1}(x_{i-1})$$

(we considered that $x_{i-1}$ is feasible, so that $\langle e, x_{i-1}\rangle = 1$), and (4.3.21) follows from (4.3.22).

From (4.3.21), it follows (compare with the proof of Proposition 4.3.1) that

$$V_0(x_0) - V_i(x_i) = F(x_0) - F(x_i) + \vartheta \ln\frac{\langle c(0), x_0\rangle}{\langle c(i), x_i\rangle} \geq i\kappa,$$

and this relation, combined with $\langle c(i), x_i\rangle = \langle c, x_i\rangle - t_i$ (since $c(i) = c - t_i e$) and $t_0 = \sigma$, implies that

$$\frac{\langle c, x_i\rangle - t_i}{\langle c, x_0\rangle - \sigma} \leq R(x_0)\exp\left\{-\frac{1}{\vartheta}(V_0(x_0) - V_i(x_i))\right\}. \tag{4.3.23}$$

Since $c(i)$ is nonnegative on the feasible set of (4.3.1) and since for feasible solutions $x$ we have $\langle c(i), x\rangle = \langle c, x\rangle - t_i$, we have $t_i \le c^*$. Thus, (4.3.23) implies the first inequality in (4.3.20). The second inequality in (4.3.20) follows from (4.3.21). Thus, (4.3.20) is proved.

To conclude our description of the sliding objective version of the Karmarkar method, note that it is difficult to expect it to be very good in practice, since the updating rule for $c(\cdot)$ is based on the worst-case analysis and therefore contradicts the spirit of large steps.

## 4.4 Projective method and linear-fractional problems

In this section, we describe another potential reduction interior-point method for (4.3.1), the so called (generalized) projective method. An advantage of this method as compared to the method of Karmarkar is that the new method does not require assumptions like (K.2) and (K.3). Besides this, the projective method can be easily extended from problem (4.3.1) the *generalized linear-fractional problem.*

### 4.4.1 Formulation of the problem. Assumptions

The generalized linear-fractional problem is as follows. Let $P$ and $Q$ be closed convex cones in $\mathbf{R}^k$ and $\mathbf{R}^l$, respectively, and let $A$, $B$, $C$ be affine mappings from $\mathbf{R}^n$ into $\mathbf{R}^k$ ($A$ and $B$) and into $\mathbf{R}^l$ ($C$). These data define the problem

$$\mathcal{P}: \qquad \text{minimize} \quad \lambda$$

subject to

$$\lambda B(v) - A(v) \in P, \tag{4.4.1}$$

$$B(v) \in P, \tag{4.4.2}$$

$$C(v) \in Q. \tag{4.4.3}$$

Let us indicate some interesting particular cases.

(i) When $P = \mathbf{R}_+$ and $B(v) \equiv 1$, then (4.4.1)–(4.4.3) is the problem of minimizing the linear form $A(v)$ over $v \in \mathbf{R}^n$ under the constraint $C(v) \in Q$; i.e., this is a conic problem associated with the cone $Q$; we refer to this case as to *linear.*

(ii) When $P = \mathbf{R}_+$ and $Q = \mathbf{R}^l_+$, then $\mathcal{P}$ is the usual linear-fractional problem with the objective $A(v)/B(v)$ and linear constraints (4.4.3) and $\{B(v) \ge 0\}$.

(iii) When $P = \mathbf{R}^k_+$, $\mathcal{P}$ is an optimization problem with the quasiconvex objective

$$\max_{1\le i\le k} \frac{A_i(v)}{B_i(v)}$$

and convex constraints (4.4.3) and $\{B_i(v) \ge 0,\ i = 1, ..., k\}$.

(iv) When $P$ is the cone of positive semidefinite matrices in the space of symmetric $m \times m$ matrices, $\mathcal{P}$ becomes the problem of minimizing $\lambda$ under the constraints $\{B(v) \geq 0,\ A(v) \leq \lambda B(v)\}$ and (4.4.3) ("$\leq$ " is understood here in the operator sense). This problem is of interest for control theory; some related applications are discussed in §5.4.

*Preliminary remarks and notation.* To proceed, it is convenient to slightly transform the problem under consideration. It is better to deal with linear (i.e., affine and homogeneous) mappings $A$, $B$, $C$ than with affine ones. To this purpose, let us add to the control vector $v$ one extra variable $t$ and represent $A(v)$, $B(v)$, $C(v)$ as $\bar{A}(v,1)$, $\bar{B}(v,1)$, $\bar{C}(c,1)$, where $\bar{A}$, $\bar{B}$, $\bar{C}$ are linear. This transforms $\mathcal{P}$ into a similar problem involving the new control vector $\bar{v} = (v,t)$ and the additional constraint $t = 1$. Since we deal only with the transformed problem, we can forget about our initial $v$, $A$, $B$, $C$ and omit bars in the notation for our new data. Thus, we henceforth formulate our problem as

$$\mathcal{P}: \qquad \text{minimize} \quad \lambda$$

subject to

$$(\lambda B - A)v \in P, \tag{4.4.4}$$

$$Bv \in P, \tag{4.4.5}$$

$$Cv \in Q, \tag{4.4.6}$$

$$\langle e, v \rangle = 1; \tag{4.4.7}$$

herein $v \in \mathbf{R}^n$, $P$ and $Q$ are convex cones in $\mathbf{R}^k$ and $\mathbf{R}^l$, respectively; $A$ and $B$ are linear homogeneous mappings from $\mathbf{R}^n$ into $\mathbf{R}^k$; $C$ is a similar mapping from $\mathbf{R}^n$ into $\mathbf{R}^l$ and $e \in \mathbf{R}^n$. Henceforth $\langle \cdot, \cdot \rangle$ denotes the standard scalar product in the corresponding $\mathbf{R}^{(\cdot)}$.

Let us denote

$$E = \mathbf{R}^k \times \mathbf{R}^k \times \mathbf{R}^l, \qquad K = P \times P \times Q,$$

so that $K$ is a convex cone in $E$, and let

$$K^* = \{s \mid \langle s, x \rangle \geq 0 \ \forall x \in K\}$$

be the cone dual to $K$. Also, let $R(\lambda)$ for every $\lambda \in \mathbf{R}$ be the linear mapping from $\mathbf{R}^n$ into $E$ defined by

$$R(\lambda)v = ((\lambda B - A)v, Bv, Cv);$$

also, let

$$E(\lambda) = \operatorname{Im} R(\lambda), \quad E^+(\lambda) = \{R(\lambda)v \mid v \in \mathbf{R}^n,\ \langle e, v \rangle \geq 0\}.$$

*Assumptions.* In what follows, we assume that the following statements are true.

(P.1) $P$ and $Q$ are closed convex pointed cones with nonempty interiors (note that then $K$ and $K^*$ also are closed convex pointed cones with nonempty interiors).

(P.2) $\mathcal{P}$ is solvable, and, for a certain optimal solution $v^*$ to $\mathcal{P}$, we have $Bv^* >_P 0$ (we write $a >_M b$, $M$ being a cone, if and only if $a - b \in \operatorname{int} M$, and $a \geq_M b$, if and only if $a - b \in M$).

(P.3) There exists $z^* \in \mathbf{R}^n$ such that

$$\langle e, z^* \rangle = 1 \quad \text{and} \quad Cz^* >_Q 0.$$

Our last assumption about $\mathcal{P}$ (it is not so crucial, but simplifies exposition) is the following statement.

(P.4) $\operatorname{Ker} R(\lambda) = \{0\}$ for all $\lambda \in \mathbf{R}$.

Note that, in the case when $\mathcal{P}$ represents problem (4.3.1) (see §4.4.1 item (i)), assumptions (P.1)–(P.3) mean exactly that (4.3.1) is strictly feasible (i.e., the feasible plane intersects the interior of $K$) and solvable.

These were assumptions about $\mathcal{P}$. Now we present the assumptions about our a priori information in the following problem.

(I.1) We are given two bounds $\tilde{\lambda}$ and $\bar{\lambda}$ such that

$$\tilde{\lambda} < \lambda^* < \bar{\lambda},$$

$\lambda^*$ being the optimal value in $\mathcal{P}$.

(I.2) We are given an oracle that, at any input, $x \in E$ reports whether $x \in \operatorname{int} K$.

(I.3) We are given a $\vartheta$-logarithmically homogeneous self-concordant barrier $F^*$ for the cone $K^*$ dual to $K$.

Note that the only description of $P$ and $Q$ used by the method that we develop is the one given by (I.2) and (I.3).

### 4.4.2 Description of the method

The idea of the method is as follows. Let $\varepsilon$ be a small positive real. From (P.1)–(P.4), it is easily seen that there exists $\bar{v} = \bar{v}(\varepsilon)$ such that $\langle e, \bar{v} \rangle > 0$, $B\bar{v} >_P 0$ and $R(\lambda)\bar{v} >_K 0$ for all $\lambda \geq \lambda^* + \varepsilon$. We associate with $\bar{v}$ "the potential"

$$\mathcal{V}_\varepsilon(s, \lambda) = F^*(s) + \vartheta \ln \langle s, R(\lambda)\bar{v} \rangle\,;$$

here $\lambda \geq \lambda^* + \varepsilon$ and $s \in \operatorname{int} K^*$. Since $B\bar{v} >_P 0$, the potential is monotone in $\lambda$; aside from this, from $\vartheta$-logarithmical homogeneity of $F^*$, we can easily derive that the potential is below bounded in $s \in \operatorname{int} K^*$ and $\lambda \geq \lambda^* + \varepsilon$,

$$\mathcal{V}_\varepsilon(s, \lambda) \geq r(\varepsilon).$$

Our method is based on a strategy for updating a given pair $(s, \lambda)$, $s \in \operatorname{int} K^*$, $\lambda > \lambda^*$ into a new pair $(s', \lambda') = \mathcal{U}(s, \lambda)$ with the following properties:

($\alpha$) $s' \in \operatorname{int} K^*$; $\lambda' \leq \lambda$;

($\beta$) If $\lambda' \geq \lambda^* + \varepsilon$, then $\mathcal{V}_\varepsilon(s', \lambda')$ is "significantly" (at least by an absolute constant) less than $\mathcal{V}_\varepsilon(s, \lambda)$;

($\gamma$) The updating is organized in such a way that, if it results in $\lambda' < \lambda$, then a byproduct of the updating is a feasible solution to $\mathcal{P}$ with the value of the objective equal to $\lambda'$.

It should be stressed that, although the properties of the updating are related to the potential, the updating itself does not use neither $\bar{v}(\varepsilon)$ nor $\varepsilon$ and is therefore *independent of the potential.*

Our method is to iterate the updating $\mathcal{U}$, starting with a pair $(\bar{s}, \bar{\lambda})$, where $\bar{s}$ is an arbitrary point from $\operatorname{int} K^*$. From ($\alpha$) and ($\beta$), it follows that, until the current value of $\lambda$ remains greater of equal $\lambda^* + \varepsilon$, the potential associated with $\varepsilon$ decreases at a step at least by an absolute constant. Since the potential is below bounded, we conclude that, in a finite number of steps (not exceeding $O(\mathcal{V}_\varepsilon(\bar{s}, \bar{\lambda}) - r(\varepsilon)))$, the current $\lambda$ becomes less than $\lambda^* + \varepsilon$. On the other hand, ($\gamma$) claims that, if the current value of $\lambda$ is less than $\bar{\lambda}$, then we already have found a solution to $\mathcal{P}$ within accuracy $(\lambda - \lambda^*)$. These observations imply convergence of the method; evaluating $r(\varepsilon)$ and $\mathcal{V}_\varepsilon(\bar{s}, \bar{\lambda})$, we come to the efficiency estimate.

It remains to explain the nature of the updating $\mathcal{U}$. Let

$$\bar{F}(x) = \sup\{\langle s, x\rangle - F^*(s) \mid s \in \operatorname{int} K^*\} \tag{4.4.8}$$

be the Legendre transformation of $F^*$ and let $F(x) = \bar{F}(-x)$. As we know (see Theorem 2.4.4 and Proposition 2.4.1), $F$ is a $\vartheta$-logarithmically homogeneous self-concordant barrier for $K$, and, if $s \in \operatorname{int} K^*$, then the point

$$x = -(F^*)'(s)$$

belongs to $\operatorname{int} K$; besides this, if $r < 1$, then the Dikin's ellipsoid of the barrier $F$

$$W_r(x) = \{y \mid \| y - x \|_{x,F} \leq r\},$$

where

$$\| u \|_{x,F} = \langle F''(x)u, u\rangle^{1/2} = \left\langle [(F^*)''(s)]^{-1}u, u\right\rangle^{1/2},$$

is contained in $\operatorname{int} K$ (Proposition 2.3.2(i.1)). Consider the set

$$E^+(\lambda) = \{R(\lambda)v \mid v \in \mathbf{R}^n,\ \langle e, v\rangle \geq 0\}$$

and let $V(\lambda) = R(\lambda)v(\lambda)$ be the point of this set closest to $x$ in the norm $\| \cdot \|_{x,F}$.

It may happen (case I) that $\| V(\lambda) - x \|_{x,F} \geq 0.99$. It turns out that in this case the direction

$$\zeta = -\Phi(V(\lambda) - x), \qquad \Phi = F''(x) = [(F^*)''(s)]^{-1}$$

is a descent direction for $\mathcal{V}_\varepsilon(\cdot,\lambda)$ at the point $s$, provided that $\lambda \geq \lambda^* + \varepsilon$, and, under the latter assumption, an appropriate step $s' = s + \rho\zeta$ in this direction keeps the point in $\operatorname{int} K^*$ and decreases $\mathcal{V}_\varepsilon$ at least by an absolute constant. Note that, to choose $\rho$, we *should not* know the potential, same as we should not bother about whether $\lambda \geq \lambda^* + \varepsilon$ or not. After $s$ is updated, we set $\lambda' = \lambda$; note that the result of our updating satisfies $(\alpha)$–$(\gamma)$.

Now consider the case opposite to $I$, i.e., when $\| V(\lambda) - x \|_{x,F} < 0.99$. Since the ellipsoids $\Omega_r(s)$ are contained in $\operatorname{int} K$ for $r < 1$, the relation $\| x - V(\lambda') \|_\Phi < 1$ implies that $V(\lambda') \in \operatorname{int} K$, so that we can immediately find a feasible solution $z_{\lambda'}$ to $\mathcal{P}$ with the value of the objective equal to $\lambda'$. Namely, if $\langle e, v(\lambda')\rangle > 0$, then we can take

$$z_{\lambda'} = \frac{v(\lambda')}{\langle e, v(\lambda')\rangle}.$$

If $\langle e, v(\lambda')\rangle = 0$, then it is possible to find $\delta > 0$ such that $R(\lambda')(v(\lambda') + \delta e) \in \operatorname{int} K$, and we can take

$$z_{\lambda'} = \frac{v(\lambda') + \delta e}{\langle e, v(\lambda') + \delta e\rangle}.$$

No other possibilities can occur, since by construction $\langle e, v(\lambda')\rangle \geq 0$.

Our observations demonstrate that the relation $\| V(\lambda') - x \|_{x,F} < 1$ for sure is not satisfied if $\lambda' \leq \lambda^*$. On the other hand, in the relevant case, we have

$$\| V(\lambda) - x \|_{x,F} < 0.99.$$

Thus, bisection in $\lambda' \in [\tilde{\lambda}, \lambda]$ allows us to find $\lambda' < \lambda$ such that

$$0.99 \leq \| V(\lambda') - x \|_{x,F} < 1.$$

This $\lambda'$ is the $\lambda$-component of the updated pair; note that $z_{\lambda'}$ is the feasible solution to $\mathcal{P}$ required by $(\gamma)$. It remains to form $s'$ and to ensure $(\alpha)$ and $(\beta)$. To this end, let us note that the pair $(s, \lambda')$ satisfies the requirements of case $I$, so that the updating $s \to s'$ corresponding to this case ensures the relation

$$\{s' \in \operatorname{int} K^*\} \& \{\lambda' \geq \lambda + \varepsilon \Rightarrow \mathcal{V}_\varepsilon(s', \lambda') \leq \mathcal{V}_\varepsilon(s, \lambda') - \text{const}\},$$

const being a positive absolute constant. Since $\mathcal{V}_\varepsilon(s, \cdot)$ is monotone, the latter relation implies $(\alpha)$ and $(\beta)$.

The method we describe produces sequences $s_i \in E$, $\lambda_i \in \mathbf{R}$, and $z_i \in \mathbf{R}^n \bigcup \{*\}$ such that $\lambda_1 \geq \lambda_2 \geq \dots$, and the following predicate is maintained:

$$(\mathrm{L}_i): \qquad \{s_i \in \operatorname{int} K^*\} \& \{\lambda_i > \lambda^*\},$$

$$\{\{\lambda_i < \bar{\lambda}\} \Rightarrow \{z_i \in \mathbf{R}^n \text{ and } (\lambda_i, z_i) \text{ is feasible for } \mathcal{P}\}\}. \tag{4.4.9}$$

To initialize the method, we choose $s_0$ as an arbitrary point of $\operatorname{int} K^*$ and set

$$\lambda_0 = \bar{\lambda}, \qquad z_0 = *;$$

note that this ensures $(\mathrm{L}_0)$.

At $i$th iteration of the method, we transform $(s_{i-1}, \lambda_{i-1})$ into $(s_i, \lambda_i, z_i)$. This transformation looks as follows (italicized text describes the rules, and the comments are in the standard font).

$1^0$. *Set*

$$x_i = -(F^*)'(s_{i-1}), \quad \Psi_i = (F^*)''(s_{i-1}), \quad \Phi_i = \Psi_i^{-1} \tag{4.4.10}$$

*and go to* $2^0$.

**Remark 4.4.1** *As we know (Theorem* 2.4.4, *Propositions* 2.3.5 *and* 2.3.2(i.1)),

$$x_i \in \operatorname{int} K, \tag{4.4.11}$$

$$\Psi_i \textit{ is nondegenerate} \tag{4.4.12}$$

*(so that* $\Phi_i$ *is well defined), and the inclusions*

$$W_r^i \equiv \{x \mid \langle \Phi_i(x - x_i), x - x_i \rangle \le r^2\} \subseteq \operatorname{int} K, \qquad 0 \le r < 1, \tag{4.4.13}$$

$$\Omega_r^i \equiv \{s \mid \langle s - s_{i-1}, \Psi_i(s - s_{i-1}) \rangle \le r^2\} \subseteq \operatorname{int} K^*, \qquad 0 \le r < 1 \tag{4.4.14}$$

*hold.*

For $\lambda \in \mathbf{R}$, let

$$V_i(\lambda) \equiv R(\lambda) v_i(\lambda)$$

be the closest to $x_i$ point of $E^+(\lambda)$ (see §4.4.1), the distance being measured with respect to the Euclidean structure $\mathcal{E}_i$ defined by the scalar product

$$\langle x, Y \rangle_i = \langle \Phi_i x, Y \rangle .$$

It is easily seen that $V_i(\lambda)$ can be found as follows. First, compute the $\mathcal{E}_i$-closest to $x_i$ point $H_i(\lambda)$ in $E(\lambda)$ as follows:

$$H_i(\lambda) = R(\lambda) h_i(\lambda), \qquad h_i(\lambda) = [\phi_i(\lambda)]^{-1} R^T(\lambda) \Phi_i x_i, \tag{4.4.15}$$

where

$$\phi_i(\lambda) = R^T(\lambda) \Phi_i R(\lambda). \tag{4.4.16}$$

Second, check whether $\langle e, h_i(\lambda) \rangle \ge 0$. If so, set $v_i(\lambda) = h_i(\lambda)$, and consequently $V_i(\lambda) = H_i(\lambda)$. Otherwise, choose $v_i(\lambda)$ as the point of the half-space $\{\langle e, \cdot \rangle \ge 0\}$ closest to $h_i(\lambda)$ in the Euclidean metric defined by the scalar product $\langle \phi_i(\lambda) x, y \rangle$ on $\mathbf{R}^n$,

$$v_i(\lambda) = h_i(\lambda) - \frac{\langle h_i(\lambda), e \rangle}{\langle [\phi_i(\lambda)]^{-1} e, e \rangle} [\phi_i(\lambda)]^{-1} e \tag{4.4.17}$$

and set $V_i(\lambda) = R(\lambda) v_i(\lambda)$.

We are ready to describe the second operation of $i$th iteration.

$2^0$. *Compute* $V_i(\lambda_{i-1})$ *and check whether* $V_i(\lambda_{i-1}) \in \operatorname{int} K$.

*If no (case* $\mathrm{I}_i$*), set* $\lambda_i = \lambda_{i-1}$, $z_i = z_{i-1}$ *and go to* $4^0$.

*If yes (case* $\mathrm{II}_i$*), go to* $3^0$.

$3^0$. (1) *Applying dichotomy in* $\lambda \in [\tilde{\lambda}, \lambda_{i-1}]$, *find* $\lambda_i \in [\tilde{\lambda}, \lambda_{i-1}]$ *such that*

$$V_i(\lambda_i) \in (\operatorname{int} K)\backslash W^i_{0.99}. \tag{4.4.18}$$

$3^0$. (2) *Choose* $v_i \in \mathbf{R}^n$ *such that*

$$\{R(\lambda_i)v_i \in (\operatorname{int} K)\backslash W^i_{0.99}\} \quad and \quad \{\langle e, v_i\rangle > 0\}, \tag{4.4.19}$$

*set*

$$z_i = \frac{v_i}{\langle e, v_i\rangle}, \tag{4.4.20}$$

*and go to* $4^0$.

Rule $3^0$ needs justification. We should prove, first, that $3^0$. (1) leads to the required $\lambda_i$ after a finite number of dichotomy steps and, second, we must explain how to find $v_i$ required in $3^0$. (2).

Let us start with the first of these two issues. In view of $2^0$, we should perform $3^0$ only if

$$V_i(\lambda_{i-1}) \in \operatorname{int} K.$$

Also, if $V_i(\lambda_{i-1}) \notin W^i_{0.99}$, then $\lambda_{i-1}$ can be chosen as the required $\lambda_i$, and there is no problem with $3^0$. (1). It remains to consider the case when

$$V_i(\lambda_{i-1}) \in W^i_{0.99}. \tag{4.4.21}$$

Note that

$$V_i(\tilde{\lambda}) \notin \operatorname{int} K. \tag{4.4.22}$$

Indeed, otherwise, every $v \in \mathbf{R}^n$ close enough to $v_i(\tilde{\lambda})$ would satisfy the inclusion $R(\tilde{\lambda})v \in \operatorname{int} K$, and, since $\left\langle e, v_i(\tilde{\lambda})\right\rangle \geq 0$ (by construction), we could choose the above "close enough to $v_i(\tilde{\lambda})$" $v$ in the half-space $\{\langle e, \cdot\rangle > 0\}$, so that $\langle e, v\rangle > 0$ and $R(\tilde{\lambda})v \in \operatorname{int} K$, which means that the pair $(\tilde{\lambda}, \langle e, v\rangle^{-1} v)$ is feasible for $\mathcal{P}$. The latter is impossible, since $\lambda < \lambda^*$. Thus, (4.4.22) does hold.

Now note that (see Remark 5.1)

$$W^i_{0.99} \subseteq W^i_{0.999} \subseteq \operatorname{int} K. \tag{4.4.23}$$

In view of (P.4), the curve $V_i(\lambda)$, $\lambda \in [\tilde{\lambda}, \lambda_{i-1}]$ is continuous; as $\lambda$ varies from $\lambda_{i-1}$ to $\lambda$, this curve passes from a point belonging to $W^i_{0.99}$ (see (4.4.21)) to a point outside $\operatorname{int} K$ (see (4.4.22)), and therefore (see (4.4.23)), the dichotomy in $\lambda$ after a finite number of steps results in a point $\lambda' \in W^i_{0.999}\backslash W^i_{0.99}$; in view of (4.4.23), this $\lambda'$ can be taken as the desired $\lambda_i$.

Thus, $3^0$. (1) is justified.

$3^0$. (2) is almost evident: Since $V_i(\lambda_i) = R(\lambda_i)v_i(\lambda_i) \in (\operatorname{int} K)\backslash W^i_{0.99}$ and $0 \le \langle e, v_i(\lambda_i)\rangle$, to obtain $v_i$ required in $3^0$. (2), it suffices to set, say,

$$v_i = v_i(\lambda_i) + \varepsilon_i e$$

with small enough $\varepsilon_i > 0$.

**Remark 4.4.2** *Note that we have only proved that our dichotomy is finite, but did not evaluate the required number of steps. Of course, dichotomy is a fast process, so that, in practice, we should not bother much about its laboriousness; nevertheless, theoretically, our scheme should be regarded as presenting a "conceptual" polynomial-time method, rather than a method that is actually polynomial. We do not investigate the problem of how long dichotomy could be in a general linear-fractional case; let us simply indicate that, in the linear case (see* §4.4.1, *item* (i)), *we can avoid necessity in dichotomy at all. Indeed, it is easily seen that, to find the required* $\lambda_i$, *it suffices to minimize a linear-fractional function over the ellipsoid under the constraint that the denominator is nonnegative. In the case in question, we have* $P = \mathbf{R}_+$, $A(v) = a^T v$, $B(v) = e^T v$. *Assume (as it is normally the case) that the barrier* $F^*$ *is of the form* $F^*(t, \tau, \sigma) = -\ln t - \ln \tau + \Xi^*(\sigma)$, *where* $t$, $\tau$, *and* $\sigma$ *are projections of* $s \in K^* = \mathbf{R}^+ \times \mathbf{R}^+ \times Q^*$ *onto the direct factors and* $\Xi^*$ *is a logarithmically homogeneous self-concordant barrier for* $Q^*$. *Then the positive definite quadratic form* $x^T \Phi x$ *is a sum of two other positive definite forms, the first of them depending on the projection of* $x$ *onto* $P \times P \equiv \mathbf{R}^+ \times \mathbf{R}^+$ *and the second,* $y^T \bar{\Phi} y$, *depending only on the projection* $y$ *of* $x$ *onto* $Q$. *It is easily seen that we then can set*

$$\lambda_i = \min\{\lambda \mid \lambda e^T v \ge a^T v,\ e^T v \ge 0 \textit{ for some } v \textit{ such that } v^T C^T \bar{\Phi} C v \le 0.999\}.$$

*The optimal value in the latter problem, if it exists, is given by simple explicit formulae, and it turns out that under assumptions* (P.1)–(P.3), *the solution does exist. Thus, in the linear case our method is actually polynomial.*

**Remark 4.4.3** *Note that in the case* $II_i$ *(i.e., when* $3^0$ *is applied)* $(\lambda_i, z_i)$ *is strictly feasible to* $\mathcal{P}$ *(i.e.,* $z_i \in \mathbf{R}^n$, $\langle e, z_i\rangle = 1$ *and* $R(\lambda_i)z_i \in \operatorname{int} K$), *so that we have*

$$(4.4.24)\qquad \begin{aligned} &(M_i): \quad \{\lambda_i > \lambda^*\}, \textit{ and} \\ &\{\{\lambda_i < \bar{\lambda}\} \Rightarrow \{z_i \in \mathbf{R}^n \textit{ and } (\lambda_i, z_i) \textit{ is feasible for } \mathcal{P}\}\}. \end{aligned}$$

*In the case of* $I_i$, *we have* $(\lambda_i, z_i) = (\lambda_{i-1}, z_{i-1})$, *and* (4.4.24) *follows from* $(M_{i-1})$. *Thus,* $(M_i)$ *always holds.*

Our last rule is as follows.

$4^0$. (1) *Set*

(4.4.25) $$\xi_i = V_i(\lambda_i) - x_i,$$

(4.4.26) $$\zeta_i = \Phi_i \xi_i,$$

(4.4.27) $$\rho_i = \{\langle \zeta_i, \xi_i \rangle\}^{1/2},$$

(4.4.28) $$\tau_i = \frac{1}{1+\rho_i}.$$

$4^0$. (2) *Choose $t_i \geq 0$ such that*

(4.4.29) $$\pi_i(t_i) \leq \pi_i(\tau_i),$$

*where*

$$\pi_i(t) = F^*(s_{i-1} - t\zeta_i) : \Delta_i \to \mathbf{R},$$

(4.4.30) $$\Delta_i = \{t \in \mathbf{R} \mid s_{i-1} - t\zeta_i \in \operatorname{int} K^*\},$$

*and set*

(4.4.31) $$s_i = s_{i-1} - t_i \zeta_i.$$

*ith iteration is completed.*

Rule $4^0$ needs justification, just as rule $3^0$ did. We should prove that $\pi_i$ is well defined at $\tau_i$, i.e., that

(4.4.32) $$\tau_i \in \Delta_i,$$

so that (4.4.29) can be satisfied simply by the choice $t_i = \tau_i$.

**Lemma 4.4.1** *We have*

(4.4.33) $$\langle \zeta_i, \Psi_i \zeta_i \rangle = \langle \Phi_i \xi_i, \xi_i \rangle = \langle \zeta_i, \xi_i \rangle = \rho_i^2 \geq (0.99)^2;$$

(4.4.34) $$\tau_i^2 \langle \zeta_i, \Psi_i \zeta_i \rangle < 1.$$

**Proof.** The two equalities in (4.4.33) immediately follow from (4.4.26); the third equality in (4.4.33) is the definition of $\rho_i$. To prove the inequality in (4.4.33), note that, in the case of $I_i$, we have $V_i(\lambda_i) = V_i(\lambda_{i-1}) \notin \operatorname{int} K$, which, in view of (4.4.13), implies that

$$\langle \xi_i, \xi_i \rangle_i \equiv \langle \Phi_i(V_i(\lambda_i) - x_i), V_i(\lambda_i) - x_i \rangle \geq 1 \geq (0.99)^2,$$

as it is required in (4.4.33). In the case of $II_i$, we have

(4.4.35) $$V_i(\lambda_i) \in (\operatorname{int} K) \backslash W^i_{0.99}$$

(see rule $3^0$. (1)), which again results in $\langle \xi_i, \xi_i \rangle_i \geq (0.99)^2$. Thus, (4.4.33) does hold.

Equation (4.4.34) is an immediate corollary of (4.4.33) and the definition of $\tau_i$. □

Now we are ready to prove that $\tau_i \in \Delta_i$. Indeed, from (4.4.34), it follows that

$$T \equiv s_{i-1} - \tau_i \zeta_i \in \operatorname{int} \Omega_1^i \tag{4.4.36}$$

(see (4.4.14)), so that $T \in \operatorname{int} K^*$ in view of (4.4.14), and therefore $\tau_i \in \Delta_i$ (see (4.4.30)).

Thus, $4^0$ is justified. Since $t_i \in \Delta_i$ by construction, the point $s_i = s_{i-1} - t_i \zeta_i$ belongs to $\operatorname{int} K^*$ (see the definition of $\Delta_i$), which, combined with ($\mathrm{M}_i$), proves ($\mathrm{L}_i$). Thus, our iteration is well defined and maintains ($\mathrm{L}_i$).

### 4.4.3 Rate of convergence

Since $Bx^* \in \operatorname{int} P$, the following two constants are well defined:

$$\gamma_1 = 1 + \min\{\gamma \geq 0 \mid \gamma B x^* + B z^* \geq_P 0\}, \tag{4.4.37}$$

$$\gamma_2 = 1 + \min\{\gamma \geq 0 \mid \gamma B x^* + (\lambda^* B - A) z^* \geq_P 0\}. \tag{4.4.38}$$

Let

$$\gamma(\varepsilon) = \frac{\varepsilon}{\gamma_1 \varepsilon + \gamma_2}; \tag{4.4.39}$$

note that

$$0 < \gamma_1 \gamma(\varepsilon) < 1, \qquad \varepsilon > 0. \tag{4.4.40}$$

Also, let

$$V^* = (Bx^*, Bx^*, Cz^*); \tag{4.4.41}$$

note that, in view of (P.2) and (P.3), we have

$$V^* \in \operatorname{int} K. \tag{4.4.42}$$

Let us define the quantity $L(s_0)$ by the relation

(4.4.43)
$$\ln L(s_0) = \frac{1}{\vartheta}\{F^*(s_0) + F(V^*)\} + 1 - \ln \vartheta$$
$$+ \ln(\max\{\langle s_0, R(\bar{\lambda})(\varepsilon z^* + (1-\varepsilon) x^*)\rangle \mid 0 \leq \varepsilon \leq 1\}).$$

Let us define $z_\varepsilon$, $0 < \varepsilon \leq 1$ by the relation

$$z_\varepsilon = (1 - \gamma(\varepsilon)) x^* + \gamma(\varepsilon) z^*. \tag{4.4.44}$$

Let us also write

$$x \succ_K Y,$$

if $x \geq_K Y$ and $x \neq Y$.

**Lemma 4.4.2** *For every $\varepsilon \in (0,1)$ and for all $\lambda > \lambda^* + \varepsilon$, we have*

$$\langle e, z_\varepsilon \rangle = 1 \tag{4.4.45}$$

*and*

$$R(\lambda) z_\varepsilon \succ_K \gamma(\varepsilon) V^*. \tag{4.4.46}$$

*In particular,*

$$R(\lambda) z_\varepsilon \in E^+(\lambda) \bigcap (\operatorname{int} K). \tag{4.4.47}$$

**Proof**. Let us fix $\varepsilon \in (0,1)$ and $\lambda > \lambda^* + \varepsilon$. (4.4.45) is evident. To prove (4.4.46), note that

$$\begin{aligned} (\lambda B - A) z_\varepsilon - \gamma(\varepsilon) B x^* = (1 - \gamma(\varepsilon))(\lambda^* B - A) x^* \\ + (1 - \gamma(\varepsilon))(\lambda - \lambda^*) B x^* \\ + \gamma(\varepsilon)(\lambda - \lambda^*) B z^* \\ + \gamma(\varepsilon)(\lambda^* B - A) z^* - \gamma(\varepsilon) B x^*. \end{aligned} \tag{4.4.48}$$

Since $x^*$ is feasible for $\mathcal{P}$, we have $(\lambda^* B - A) x^* \geq_P 0$, and therefore (see (4.4.40))

$$(1 - \gamma(\varepsilon))(\lambda^* B - A) x^* \geq_P 0. \tag{4.4.49}$$

Furthermore, in view of (4.4.37) and (4.4.38),

$$\gamma(\varepsilon) B z^* \geq_P -\gamma(\varepsilon)(\gamma_1 - 1) B x^* \tag{4.4.50}$$

and

$$\gamma(\varepsilon)(\lambda^* B - A) z^* \geq_P -\gamma(\varepsilon)(\gamma_2 - 1) B x^*. \tag{4.4.51}$$

Relations (4.4.48)–(4.4.51) imply that

$$\begin{aligned} (\lambda B - A) z_\varepsilon - \gamma(\varepsilon) B x^* \geq_P (\lambda - \lambda^*)\{1 - \gamma_1 \gamma(\varepsilon)\} B x^* \\ -\gamma(\varepsilon)(\gamma_2 - 1) B x^* - \gamma(\varepsilon) B x^* \succ_P \\ \succ_P \{\varepsilon(1 - \gamma_1 \gamma(\varepsilon)) - \gamma_2 \gamma(\varepsilon)\} B x^* = 0 \end{aligned} \tag{4.4.52}$$

(we considered that $\lambda - \lambda^* > \varepsilon$ and the definition of $\gamma(\cdot)$).

Now, by similar reasoning,

$$\begin{aligned} B z_\varepsilon - \gamma(\varepsilon) B x^* = (1 - 2\gamma(\varepsilon)) B x^* + \gamma(\varepsilon) B z^* \geq_P \\ \geq_P (1 - 2\gamma(\varepsilon)) B x^* - (\gamma_1 - 1)\gamma(\varepsilon) B x^* \geq_P 0. \end{aligned} \tag{4.4.53}$$

Since $C x^* \geq_P 0$, we also have

$$C z_\varepsilon - \gamma(\varepsilon) C z^* \geq_Q \gamma(\varepsilon) C z^* - \gamma(\varepsilon) C z^* \geq_Q 0. \tag{4.4.54}$$

Equations (4.4.52)–(4.4.54) mean precisely that (4.4.46) holds. □

The main result on the above method is as follows.

**Theorem 4.4.1** *Let*

$$\delta_i = \pi_i(0) - \pi_i(t_i); \tag{4.4.55}$$

*then, for all $i$, we have*

$$\delta_i \geq \kappa \equiv 0.99 - \ln(1.99) > 0. \tag{4.4.56}$$

*Furthermore, let $i$ be such that*

$$\ln \gamma(1) > \ln L(s_0) - \frac{1}{\vartheta}\sum_{j=1}^{i} \delta_j \tag{4.4.57}$$

*(in view of (4.4.56) the latter relation is satisfied when $i > \rfloor \vartheta\kappa^{-1} \ln(L(s_0)/\gamma(1)) \lfloor$ ) and let $\varepsilon_i \in (0,1)$ be defined by the relation*

$$\ln \gamma(\varepsilon_i) = \ln L(s_0) - \frac{1}{\vartheta}\sum_{j=1}^{i} \delta_j; \tag{4.4.58}$$

*then*

$$\varepsilon_i \leq \frac{L(s_0)}{\gamma(1)} \exp\left\{-\frac{1}{\vartheta}\sum_{j=1}^{i} \delta_j\right\} \leq \frac{L(s_0)}{\gamma(1)} \exp\left\{-\frac{\kappa}{\vartheta} i\right\}, \tag{4.4.59}$$

*and*

$$\lambda_i \leq \lambda^* + \varepsilon_i. \tag{4.4.60}$$

*In particular, for all $i$ satisfying (4.4.57) and the relation*

$$\lambda^* + \varepsilon_i < \bar{\lambda}, \tag{4.4.61}$$

*the point $z_i$ belongs to $\mathbf{R}^n$, and $(\lambda_i, z_i)$ is a feasible solution to $\mathcal{P}$ with the accuracy (in terms of the objective of $\mathcal{P}$) no worse than $\varepsilon_i$.*

*Thus, to solve $\mathcal{P}$ to an accuracy $\varepsilon \in (0,1)$, it suffices to perform no more than*

$$M(\varepsilon) = O(1)\vartheta \ln \frac{L(s_0)}{\gamma(1)\varepsilon}$$

*iterations of the method, $O(1)$ being an absolute constant.*

**Proof.** $1^0$. Let us prove (4.4.56). Since $F^*$ is $\vartheta$-logarithmically homogeneous self-concordant barrier for $K^*$, the function $\pi_i$ is strongly self-concordant on $\Delta_i$ with the parameter of self-concordance equal to 1. Furthermore, in view of $(F^*)'(s_{i-1}) = -x_i$ (the origin of $x_i$), we have

$$\begin{aligned} \pi_i'(0) &= \langle \zeta_i, x_i \rangle = \langle \Phi_i(V_i(\lambda_i) - x_i), x_i \rangle = \langle V_i(\lambda_i) - x_i, x_i \rangle_i \\ &= -\langle V_i(\lambda_i) - x_i, V_i(\lambda_i) - x_i \rangle_i + \langle V_i(\lambda_i) - x_i, V_i(\lambda_i) \rangle_i . \end{aligned} \tag{4.4.62}$$

Since $V_i(\lambda_i)$ is the $\mathcal{E}_i$-closest to $x_i$ point of $E^+(\lambda_i)$ and the latter set is a cone, we have $\langle V_i(\lambda_i) - x_i, V_i(\lambda_i)\rangle_i = 0$, so that (4.4.62) implies that

$$\pi_i'(0) = -\langle V_i(\lambda_i) - x_i, V_i(\lambda_i) - x_i\rangle_i = -\langle \xi_i, \xi_i\rangle_i = -\rho_i^2 \tag{4.4.63}$$

(the latter equality follows from (4.4.33)).

Furthermore,

$$\pi_i''(0) = \langle \zeta_i, \Psi_i \zeta_i\rangle = \rho_i^2 \tag{4.4.64}$$

(see (4.4.33)). In particular, the Newton decrement $\lambda(\pi_i, 0)$ of $\pi_i$ at 0 equals to $\rho_i$. It follows that $\tau_i$ is precisely the iterate of 0 with respect to the Newton-type process described in Proposition 2.2.2 (where we should set $a = 1$), and, in view of the latter proposition, we have

$$\pi_i(\tau_i) \le \pi_i(0) - \{\rho_i - \ln(1+\rho_i)\} \le \pi_i(0) - \kappa, \quad \kappa = 0.99 - \ln(1.99) \tag{4.4.65}$$

(we used (4.4.33)), which, in view of (4.4.29), implies (4.4.56).

$2^0$. Let us fix $i$ satisfying (4.4.57); then $\varepsilon_i$ is well defined and belongs to $(0,1)$. Moreover, we clearly have $\gamma(\varepsilon) \ge \varepsilon\gamma(1)$, $0 \le \varepsilon \le 1$, whence

$$\varepsilon_i\gamma(1) \le \gamma(\varepsilon_i) = L(s_0)\exp\left\{-\frac{1}{\vartheta}\sum_{j=1}^{i}\delta_j\right\},$$

which, combined with (4.4.56), immediately implies (4.4.59). In view of (4.4.9), to complete the proof of the theorem, it remains to establish (4.4.60).

Assume that (4.4.60) does not hold, so that

$$\lambda_i > \lambda^* + \varepsilon_i. \tag{4.4.66}$$

From the description of the method, it follows immediately that $\lambda_j$ do not increase, so that (4.4.66) implies that

$$\lambda_j > \lambda^* + \varepsilon_i, \qquad j = 0, ..., i. \tag{4.4.67}$$

Henceforth, let $j$ take values in $\{0, 1, ..., i\}$.

Let $z = z_{\varepsilon_i}$ and let

$$Z_j = R(\lambda_j)z; \tag{4.4.68}$$

in view of Lemma 4.4.2 and the relation $\lambda_j \ge \lambda_i > \lambda^* + \varepsilon_i$, we have

$$Z_j \succ_K \gamma(\varepsilon_i)V^*, \qquad Z_j \in E^+(\lambda_j) \bigcap \operatorname{int} K. \tag{4.4.69}$$

Now set

$$\nu_j = F^*(s_j) + \vartheta \ln\langle s_j, Z_j\rangle \tag{4.4.70}$$

and let us prove that

$$\nu_{j-1} - \nu_j \ge \delta_j, \qquad j = 1, ..., i. \tag{4.4.71}$$

Indeed, we have

$$\nu_{j-1} - \nu_j = F^*(s_{j-1}) + \vartheta \ln \langle s_{j-1}, Z_{j-1} \rangle - F^*(s_j) - \vartheta \ln \langle s_j, Z_j \rangle . \tag{4.4.72}$$

Furthermore,

$$Z_{j-1} - Z_j = ((\lambda_{j-1} - \lambda_j)Bz, 0, 0),$$

and $Bz \geq_P 0$ (since $Bz$ is the second component of $Z_j$ and $Z_j >_K 0$, see (4.4.69)). Also, since $\lambda_j \leq \lambda_{j-1}$, we conclude that $Z_{j-1} - Z_j \geq_K 0$, whence, in view of $s_{j-1} \in \operatorname{int} K^*$, $\langle s_{j-1}, Z_{j-1} \rangle \geq \langle s_{j-1}, Z_j \rangle$. Therefore (4.4.72) implies that

$$\begin{aligned} \nu_{j-1} - \nu_j &\geq F^*(s_{j-1}) + \vartheta \ln \langle s_{j-1}, Z_j \rangle - F^*(s_j) - \vartheta \ln \langle s_j, Z_j \rangle \\ &= F^*(s_{j-1}) - F^*(s_{j-1} - t_j \zeta_j) \\ &\quad + \vartheta \{ \ln \langle s_{j-1}, Z_j \rangle - \ln \langle s_{j-1} - t_j \zeta_j, Z_j \rangle \}. \end{aligned} \tag{4.4.73}$$

Now we have

$$\langle \zeta_j, Z_j \rangle = \langle \Phi_j \xi_j, Z_j \rangle = \langle \xi_j, Z_j \rangle_j = \langle V_j(\lambda_j) - x_j, Z_j \rangle_j .$$

Since $Z_j \in E^+(\lambda_j)$ (see (4.4.69)), the points $V_j(\lambda_j) + tZ_j$ belong to $E^+(\lambda_j)$, $t \geq 0$, and, since $V_j(\lambda_j)$ is the $\mathcal{E}_j$-closest to $x_j$ point in $E^+(\lambda_j)$, we conclude that $\langle V_j(\lambda_j) - x_j, Z_j \rangle_j \geq 0$. Thus, $\langle \zeta_j, Z_j \rangle \geq 0$, and, since $t_j \geq 0$ (see $4^0$.(2)), we obtain

$$\ln \langle s_{j-1}, Z_j \rangle - \ln \langle s_{j-1} - t_j \zeta_j, Z_j \rangle \geq 0,$$

so that (4.4.73) results in

$$\nu_{j-1} - \nu_j \geq F^*(s_{j-1}) - F^*(s_{j-1} - t_j \zeta_j) = \pi_j(0) - \pi_j(t_j) = \delta_j,$$

as required in (4.4.71).

$3^0$. Relation (4.4.71) implies $\nu_i \leq \nu_0 - \sum_{j=1}^i \delta_j$, or, which is the same,

$$F^*(s_i) + \vartheta \ln \langle s_i, Z_i \rangle \leq F^*(s_0) + \vartheta \ln \langle s_0, Z_0 \rangle - M_i, \qquad M_i \equiv \sum_{j=1}^{i} \delta_j. \tag{4.4.74}$$

In view of the first relation in (4.4.69) combined with $s_i \in \operatorname{int} K^*$, we have

$$\vartheta \ln \langle s_i, Z_i \rangle > \vartheta \ln \langle s_i, V^* \rangle + \vartheta \ln \gamma(\varepsilon_i),$$

so that

$$\begin{aligned} F^*(s_i) + \vartheta \ln \langle s_i, Z_i \rangle &> F^*(s_i) + \vartheta \ln \langle s_i, V^* \rangle + \vartheta \ln \gamma(\varepsilon_i) \\ &\geq -F(V^*) + \vartheta \ln \gamma(\varepsilon_i) + \vartheta \ln \vartheta - \vartheta \end{aligned} \tag{4.4.75}$$

(the last inequality in (4.4.75) follows from Proposition 2.4.1).

Furthermore, since $Z_0 = R(\bar{\lambda})z_{\varepsilon_i}$, we have (see (4.4.43))

$$\vartheta \ln \langle s_0, Z_0\rangle \le \vartheta \ln L(s_0) - F^*(s_0) - F(V^*) + \vartheta \ln \vartheta - \vartheta,$$

which, combined with (4.4.75) and (4.4.74), results in

$$\vartheta \ln \gamma(\varepsilon_i) < \vartheta \ln L(s_0) - M_i.$$

The resulting inequality contradicts the definition of $\varepsilon_i$, and this contradiction demonstrates that (4.4.66) is impossible. Thus, (4.4.60) is proved. □

### 4.4.4 Concluding remarks

*Large-step strategy.* There are two possibilities to include into the method a "large-step" strategy. First, when implementing rule $3^0$. (1), we could take as $\lambda_i$ not an arbitrary $\lambda \in [\tilde{\lambda}, \lambda_{i-1}]$ satisfying (4.4.18), but as "close" to $\min\{\lambda \in [\tilde{\lambda}, \lambda_{i-1}] \mid V_i(\lambda) \notin \operatorname{int} K\}$. Moreover, it is easily seen that all we require of rule $3^0$ is to find $\lambda < \lambda_{i-1}$ and $v$, $\langle e, v\rangle > 0$, such that

$$\text{(4.4.76)} \qquad R(\lambda)v \in \operatorname{int} K \quad \text{and} \quad E^+(\lambda) \bigcap W^i_{0.99} = \emptyset;$$

such a $\lambda$ can be chosen as $\lambda_i$, and the vector $\langle e, v\rangle^{-1} v$ can be chosen as the approximate solution $z_i$ (compare with (4.4.20)). Rule $3^0$ in its basic version provides us with certain $\lambda = \hat{\lambda}$, $v = \hat{v}$ satisfying (4.4.76), and, after these quantities are found, we can try to improve the pair $(\hat{\lambda}, \hat{v})$, i.e., to replace it by a pair $(\lambda_i, v_i)$ satisfying (4.4.76) with $\lambda_i \le \hat{\lambda}$. The less $\lambda_i$ is, the better the approximate solution to $\mathcal{P}$ found at the $i$th iteration, and the larger $\rho_i$ is, i.e., the larger the (lower estimate of the) contribution $\delta_i$ of the iteration into the right-hand side of (4.4.59) is. The second possibility to implement large steps is to find $t_i$ in the $4^0$. (2) via minimization of $\pi_i(t)$ over $t \in \Delta_i$.

*"Inner" dichotomy versus the "outer" one.* An unpleasant feature of the projective method for the linear-fractional problem is the necessity to perform, even in the basic version of the method, the "inner" dichotomy in $\lambda$ (see rule $3^0$. (2)); we can eliminate this necessity only in the linear case (see Remark 4.4.2). As far as a generalized linear-fractional problem is concerned, we should note that, although this problem, generally speaking, is quasiconvex (see §4.4.1 item (iii)), we can immediately reduce it to a "small" series of convex problems. Indeed, checking whether a given $\lambda$ is greater than the optimal value to $\mathcal{P}$ and finding a feasible solution $(v, \lambda)$ to $\mathcal{P}$ is the same as finding whether the usual conic problem with trivial objective

$$\mathcal{P}_\lambda : \qquad \text{find } v \text{ such that } \lambda Bv - Av \in P,\ Bv \in P,\ Cv \in Q$$

is feasible and finding a feasible solution to $\mathcal{P}_\lambda$. To solve $\mathcal{P}$, we could perform "outer" dichotomy in $\lambda$ and use, say, the method of Karmarkar or the primal-dual method (see the next section) to analyse the arising auxiliary problems

$\mathcal{P}_\lambda$. What are the advantages of the inner dichotomy used in the projective method as compared to the strategy based on the outer dichotomy?

To understand the cost of the inner dichotomy used in projective method, we should consider the following. Our point is that the inner dichotomy, in contrast to the outer one, in many important cases is almost costless, since the computational effort required by this dichotomy is dominated by other computations required at an iteration of a potential reduction interior point method (like the projective method we are now studying, the method of Karmarkar from the previous section, and the primal-dual method that is our subject in the next section). Consider, for example, the problem mentioned in §4.4.1 item (iv), i.e., the case of trivial $Q$ and $P$ being the cone of positive semidefinite $m \times m$ matrices, and let $n$ be the dimension of the control vector $v$. In the case in question, $E$ is the space of $(2m) \times (2m)$ block-diagonal symmetric matrices with two $m \times m$ diagonal blocks; $K$ is the cone comprised by positive semidefinite matrices from $E$. We provide $E$ with the standard Euclidean structure $\langle A, B \rangle = \mathrm{Tr}\{AB\}$; then $K^* = K$. As we see in Chapter 5, the cone $K^* = K$ admits $2m$-logarithmically homogeneous self-concordant barrier

$$F^*(s) = -\ln \mathrm{Det}\, s;$$

we perform complexity analysis of an ($i$th) iteration of the projective method associated with this barrier. Let $M$ be the amount of the dichotomy steps at this iteration.

In the case under consideration, we have

$$(F^*)'(s) = -s^{-1}; \qquad (F^*)''(s)y = s^{-1}ys^{-1}. \tag{4.4.77}$$

Thus, our general scheme leads to the following computations (the quantities in angle brackets indicate the arithmetic cost of the corresponding computation).

*Rule* $1^0$. This rule requires computing $x_i = (s_{i-1})^{-1}$ $\langle O(m^3) \rangle$. Note that, in our general scheme, the next instruction was "compute the matrix $(F^*)''(s_{i-1})$ and its inverse"; to implement it literally (i.e., to compute and store $O(m^2) \times O(m^2)$ matrices), it would require at least $O(m^4)$ operations. Fortunately, what, in fact, is meant in $1^0$ is to "provide the possibility to multiply a given $Y \in E$ by $\Psi_i \equiv (F^*)''(s_{i-1})$ and $\Phi_i = [\Psi_i]^{-1}$." In view of (4.4.77), to this purpose, it suffices to know $s_{i-1}$ and $x_i$ and to compute $\Psi_i Y$ and $\Phi_i Y$ for a given $Y$ it takes $O(m^3)$ operations.

*Rules* $2^0$, $3^0$. The effort required to compute $v_i(\lambda)$ and $V_i(\lambda)$ is dominated by the following operations:

- computing the $n \times n$ matrix $\phi_i(\lambda)$;
- given $\phi_i(\lambda)$, computing its inverse $\langle O(n^3) \rangle$;
- a single multiplication of a $Y \in E$ by $R^T(\lambda)$, $\langle O(m^2 n) \rangle$, and a single multiplication of a $y \in \mathbf{R}^n$ by $R(\lambda)$, $\langle O(m^2 n) \rangle$.

It takes additionally $O(m^3)$ operations to check whether $V_i(\lambda) \in \operatorname{int} K$ and whether $V(\lambda) \in W_\cdot^i$.

Now, to compute $\phi_i(\lambda)$ for *a single* $\lambda$, it requires $N = O(m^3 n + m^2 n^2)$ operations (we should multiply each of $n$ columns $R_j$ of $R(\lambda)$ by $\Phi_i$, i.e., compute $n$ matrices $s_{i-1} R_j s_{i-1}$, $j = 1, ..., n$, which takes $O(m^3 n)$ operations, and then compute scalar products of each of the results with each of $R_j$, which takes $O(m^2 n^2)$ operations more). However (now is the crucial point), to compute $\phi_i(\cdot)$ at a series of $k$ values of $\lambda$, it takes less than $kN$ operations, namely, no more than $3N + (k-3)O(n^2)$. Indeed, $\phi_i(\lambda)$ is quadratic in $\lambda$. Therefore, we can compute and store the values of this matrix for three values of $\lambda$, say, $\lambda = 0, 1, 2$, $\langle O(m^3 n + m^2 n^2) \rangle$, and, after that, it takes only $O(n^2)$ operations to compute $\phi_i(\lambda)$ for any other $\lambda$.

*Rule* $4^0$. The arithmetic cost of this rule (at least for the choice $t_i = \tau_i$) is clearly $O(m^3)$.

It follows that the arithmetic cost of an iteration with $M$ dichotomy steps is

$$O(m^3) + O(m^3 n + m^2 n^2) + M \times O(n^2 + n^3 + m^2 n + m^3)$$

$$= O(m^3 n + m^2 n^2) + M \times O(n^3 + m^2 n + m^3).$$

(the first left-hand side term represents all expenses for $1^0$ and $4^0$, the second is the cost of computing $\phi_i(\lambda)$ for $\lambda = 0, 1, 2$, and the factor at $M$ in the third term is the cost of computing $v_i(\lambda)$, $V_i(\lambda)$ and checking whether $V_i(\lambda)$ belongs to $\operatorname{int} K$ and $W_\cdot^i$). We see that, if $n \ll m^2$, then the term $M \times O(n^3 + m^2 n + m^3)$ is dominated by the term $O(m^3 n + m^2 n^2)$ in a large enough range of values of $M$, so that, in this case, the built-in dichotomy is basically costless.

Note that the case of the usual linear-fractional programming (§4.4.1, item (ii)) is even more preferable for the inner dichotomy. Indeed, in this case, $\phi_i(\lambda)$ is a 3-rank correction of, say, $\phi_i(0)$, and it takes only $O(ln)$ operations to compute $\phi_i(\lambda)$ and $[\phi_i(\lambda)]^{-1}$ after $\phi_i(0)$ and $[\phi_i(0)]^{-1}$ are computed, where $l = \dim Q$.

## 4.5 Primal-dual potential reduction method

From the theoretical viewpoint, a disadvantage of the generalized method of Karmarkar and projective method, as well as that one of their initial LP versions, is that the efficiency estimate of the methods is proportional to the parameter of the underlying barrier, while, for the path-following methods, the estimate is proportional to the square root of the parameter. Below, we present the primal-dual potential reduction method, which is free from this shortcoming. The method extends onto the general conic case the algorithm developed by Todd and Ye (see [TY 87], [Ye 88a], [Ye 89]) for LP problems.

### 4.5.1 Formulation of the problem. Assumptions

Let us fix a primal-dual pair of conic problems

$$\text{(P):} \qquad \text{minimize } \langle c, x\rangle \quad \text{s.t. } x \in K\bigcap(L+b),$$

$$\text{(D):} \qquad \text{minimize } \langle s, b\rangle \quad \text{s.t. } s \in K^*\bigcap(L^\perp+c),$$

where, as usual, $K$ is a closed convex pointed cone in $E$ with a nonempty interior, $K^* \subset E^*$ is the dual cone, $L \subset E$ is a linear subspace, and $L^\perp \subset E^*$ is the annulator of $L$. For brevity, let us denote

$$\begin{aligned} M &= L + b \text{ (the primal feasible plane)},\\ P &= M\bigcap K \text{ (the primal feasible set)},\\ P' &= \operatorname{rint} P \quad \text{(the set of primal strictly feasible solutions)}, \end{aligned}$$

and, similarly,

$$M^* = L^\perp + c, \quad D = M^*\bigcap K^*, \quad D' = \operatorname{rint} D.$$

A pair $(x, s)$ comprised of strictly feasible primal and dual solutions (i.e., of $x \in P'$, $s \in D'$) is called a *strictly feasible primal-dual pair.*

We henceforth make the following regularity assumption:

$$\text{(PD):} \qquad ((\mathsf{P}), (\mathsf{D})) \text{ admits strictly feasible primal-dual pairs.}$$

In terms of §4.2, (PD) means that, for our problems, $(\mathrm{I}_p)$ and $(\mathrm{I}_d)$ hold. In view of implications (1) and (5) from §4.2, (PD) implies that both primal and dual problems are solvable and that their optimal values $\mathsf{P}^*$ and $\mathsf{D}^*$ satisfy the duality relation

$$\mathsf{P}^* + \mathsf{D}^* = \langle c, b\rangle\,. \tag{4.5.1}$$

Note also, that to ensure (PD), it suffices to assume that, say, the primal feasible set is bounded and intersects $\operatorname{int} K$ (assumption (K.1) from §4.3). Indeed, the latter assumption, in terms of §4.2, is $(\mathrm{I}_p)$ and $(\mathrm{B}_p)$, while (PD) is precisely $(\mathrm{I}_p)$ and $(\mathrm{I}_d)$; to prove the implication $(\mathrm{I}_p)$ and $(\mathrm{B}_p) \Rightarrow (\mathrm{I}_p)$ and $(\mathrm{I}_d)$, it suffices to note that $\rceil,(\mathrm{I}_d)$ would imply $(\mathrm{R}_p)$; see §4.2, property (2), which, combined with $(\mathrm{F}_p)$, clearly would imply $\rceil,(\mathrm{B}_p)$.

Note that, from (4.5.1), it follows (see (4.2.2)) that, for each pair $(x, s)$ of feasible primal and dual solutions, we have

$$\nu(x, s) \equiv \varepsilon(x) + \varepsilon^*(s) = \langle s, x\rangle\,, \tag{4.5.2}$$

where

$$\varepsilon(x) = \langle c, x\rangle - \mathsf{P}^*, \qquad \varepsilon^*(s) = \langle s, b\rangle - \mathsf{D}^*$$

are the primal and the dual accuracies.

The only representation of the cones $K$, $K^*$ used in the primal-dual method is the knowledge (i.e., the possibility to compute the values and first- and second-order derivatives) of a $\vartheta$-logarithmically homogeneous self-concordant barrier $F$ for the primal cone $K$ together with the Legendre transformation $F^*$ of $F$. As we know (Theorem 2.4.4), $F^*$ is $\vartheta$-logarithmically homogeneous self-concordant barrier for the cone $(-K^*)$ anti-dual to $K$, so that the function $F^+(s) = F^*(-s) : \operatorname{int} K^* \to \mathbf{R}$ is a barrier of the same type for the dual cone $K^*$. We assume that we are given $F$ and $F^+$.

To simplify our considerations, we also assume that we know an initial strictly feasible primal-dual pair $(x^0, s^0)$ (to find such a pair, we can use, say, the preliminary stage of the barrier path-following method, and the projective method, and so forth).

### 4.5.2 Primal-dual potential function

The potential underlying the method depends on a parameter $\gamma > 0$ (which is chosen as an absolute constant later) and is as follows:

$$V_*(x,s) = F(x) + F^+(s) + (\vartheta + \gamma\vartheta^{1/2}) \ln \langle s, x \rangle : \operatorname{int} K \times \operatorname{int} K^* \to \mathbf{R}.$$

Also, let

$$U(x,s) = F(x) + F^+(s) + \vartheta \ln \langle s, x \rangle,$$

so that

$$V_*(x,s) = U(x,s) + \gamma\vartheta^{1/2} \ln \langle s, x \rangle.$$

It turns out that the relative value of the duality gap $\nu(\cdot,\cdot)$ at a strictly feasible primal-dual pair $(x,s)$ (and, consequently, the accuracy of $x$ and $s$ as approximate solutions to the corresponding problems) can be estimated in terms of the potential, as it is shown in the following result.

**Proposition 4.5.1** *Let $(x,s)$ be a strictly feasible primal-dual pair and let $(x^0, s^0)$ be the initial strictly feasible primal-dual pair. Then*

$$\frac{\nu(x,s)}{\nu(x^0,s^0)} \le R(x^0,s^0) \exp\left\{-\frac{V_*(x^0,s^0) - V_*(x,s)}{\gamma\vartheta^{1/2}}\right\}, \tag{4.5.3}$$

*where*

$$R(x^0,s^0) = \exp\left\{\frac{U(x^0,s^0) - \vartheta\ln\vartheta + \vartheta}{\gamma\vartheta^{1/2}}\right\}. \tag{4.5.4}$$

**Proof.** By definition of $V_*(\cdot,\cdot)$ and by virtue of (4.5.2), we have

$$V_*(x,s) - V_*(x^0,s^0) = \gamma\vartheta^{1/2} \ln \frac{\nu(x,s)}{\nu(x^0,s^0)} + U(x,s) - U(x^0,s^0).$$

This inequality implies the statement under consideration, since

$$U(x,s) - U(x^0,s^0) \ge \vartheta\ln\vartheta - \vartheta - U(x^0,s^0)$$

(see Proposition 2.4.1; note that $U(x,s) = V(-s,x)$, $V$ being the function involved into Proposition 2.4.1). □

### 4.5.3 Basic primal-dual procedure

The above statement means that, to approximate the pair of optimal primal and dual solutions, it suffices to decrease at an appropriate rate the value of the potential function. In the method under consideration, this is ensured by a procedure that updates a given strictly feasible primal-dual pair into another pair of the same type with the potential at the updated pair being by an absolute constant less than at the initial pair. This procedure $\mathcal{PD}(\gamma, \delta)$ depends on $\gamma$ (which is responsible for the potential) and on certain positive parameter $\delta$, which is specified (as an absolute constant) later.

The procedure $\mathcal{PD}$ $(\gamma, \delta)$, as applied to a pair of strictly feasible primal and dual solutions $(x, s)$, transforms it into the pair $(x', s') = \mathcal{PD}(\gamma, \delta)[(x, s)]$ of the same type in accordance with the following rules:

$\mathcal{PD}$. (1) Set

$$\tau = (\vartheta + \gamma\vartheta^{1/2})\frac{s}{\langle s, x\rangle},$$

form the function

$$v(u) = F(u) + \langle \tau, u - x\rangle : \operatorname{int} K \to \mathbf{R},$$

and find the Newton direction, $\xi$, of $v\,|_{P'}$ at $x$,

$$\xi = \operatorname{argmin}\{\langle v'(x), h\rangle + \tfrac{1}{2}\langle v''(x)h, h\rangle \mid h \in L\}.$$

Also, let $\lambda$ be the Newton decrement of $v|_{P'}$ at $x$,

$$\lambda = \langle v''(x)\xi, \xi\rangle^{1/2}.$$

$\mathcal{PD}$. (2) If

$$\lambda > \delta, \tag{4.5.5}$$

then set

$$x' = x + \frac{\xi}{1+\lambda}, \qquad s' = s,$$

and terminate; otherwise set

$$x' = x, \qquad s' = \frac{\langle s, x\rangle}{\vartheta + \gamma\vartheta^{1/2}}[-F'(x) - F''(x)\xi]$$

and terminate.

**Proposition 4.5.2** *Under assumption* (PD) *for each* $\gamma > 0$ *and all* $\delta > 0$ *satisfying the relation*

$$(\mathcal{C}): \qquad \delta < 1, \qquad \Omega(\gamma, \delta) \equiv \frac{\gamma(\gamma(1-\delta) - \delta)}{1+\gamma} - \frac{\delta^2}{2(1-\delta)^2} > 0$$

(*this relation is satisfied for all small enough positive* $\delta$) *and for each pair* $(x, s)$ *of strictly feasible solutions to* (P) *and* (D), *the pair* $(x', s') = \mathcal{PD}(\gamma, \delta)[(x, s)]$ *is strictly feasible, and*

$$V_*(x', s') \le V_*(x, s) - \min\{\delta - \ln(1+\delta); \Omega(\gamma, \delta)\}. \tag{4.5.6}$$

**Proof.** Let us fix $\delta$ satisfying $(\mathcal{C})$ and a strictly feasible primal-dual pair $(x, s)$.

$1^0$. Assume first that (4.5.5) is the case. Then $s' = s$ is a strictly feasible dual solution. Furthermore, $\lambda = \lambda(v\mid_{P'}, x)$, and $x'$ is exactly the image of $x$ with respect to the updating described in Proposition 2.2.2 (this updating is applied to $v\mid_{P'}$). By virtue of Proposition 2.2.2, $x' \in P'$, and

$$v(x') \le v(x) - \{\lambda - \ln(1+\lambda)\} \le v(x) - \{\delta - \ln(1+\delta)\}. \tag{4.5.7}$$

Since $\ln(\cdot)$ is concave and $\tau$ is precisely the derivative of $(\vartheta + \gamma\vartheta^{1/2})\ln\langle s, u\rangle$ in $u$ at $x$, we have

$$\begin{aligned} V_*(x', s') - V_*(x, s) &= V_*(x', s) - V_*(x, s) \\ &= (\vartheta + \gamma\vartheta^{1/2})\{\ln\langle s, x'\rangle - \ln\langle s, x\rangle\} + F(x') - F(x) \\ &\le \langle \tau, x' - x\rangle + F(x') - F(x) = v(x') - v(x), \end{aligned}$$

so that (4.5.7) implies that

$$V_*(x', s') - V_*(x, s) \le -\{\delta - \ln(1+\delta)\}. \tag{4.5.8}$$

Thus, in the case of (4.5.5), $(x', s')$ is strictly dual feasible, and (4.5.6) holds.

$2^0$. Now consider the case when

$$\lambda \le \delta. \tag{4.5.9}$$

Then $x' = x$ is strictly primal feasible. To prove that $s'$ is strictly dual feasible, denote $s^* = -F'(x)$; then, under the notation

$$\rho = \gamma\vartheta^{1/2}, \qquad \pi = \frac{\langle s, x\rangle}{\vartheta + \rho},$$

we have

$$s' = \pi s'', \qquad s'' = s^* - F''(x)\xi. \tag{4.5.10}$$

We have $s^* \in \operatorname{int} K^*$ (see Theorem 2.4.4). To prove that $s'$ is strictly dual feasible, let us first prove that

$$\| s'' - s^* \|_{s^*}^2 \equiv \langle s'' - s^*, (F^+)''(s^*)(s'' - s^*)\rangle = \langle \xi, F''(x)\xi\rangle = \lambda^2 < \delta^2. \tag{4.5.11}$$

We have

$$\langle s'' - s^*, (F^+)''(s^*)(s'' - s^*)\rangle = \langle F''(x)\xi, (F^+)''(s^*)F''(x)\xi\rangle = \langle F''(x)\xi, \xi\rangle$$

(we considered that $F^+(-u)$ is the Legendre transformation of $F$, so that $(F^+)''(-F'(w)) = (F''(w))^{-1}$, $w \in \operatorname{int} K$, while $s^* = -F'(x)$). Thus, (4.5.11) follows from the definition of $\lambda$ and (4.5.9).

Since $\delta$ is less than 1, (4.5.11) means that $s''$ belongs to the open ellipsoid

$$W = \{u \in E^* \mid \langle u - s^*, (F^+)''(s^*)(u - s^*)\rangle < 1\},$$

which, in view of Proposition 2.3.2, item (i.1), is contained in $\operatorname{int} K^*$. Thus, $s'' \in \operatorname{int} K^*$, and, since $K^*$ is a cone, $s' \in \operatorname{int} K^*$ as well.

$3^0$. Now let us prove that $s' \in M^*$. By definition, we have $s'' = -F'(x) - F''(x)\xi$; by definition of $\xi$, the linear functional

$$f \equiv F''(x)\xi + F'(x) + \tau = -s'' + \tau$$

vanishes at the linear subspace $L$, so that $f = -s'' + \tau \in L^\perp$. We have $s' - c = \pi s'' - c = (\tau - f)\pi - c = (\pi\tau - c) - \pi f$. We have $\pi\tau = s$ (the definitions of $\pi$ and $\tau$), so that $s' - c = (s - c) - \pi f \in L^\perp$ (recall that $s$ is dual feasible, and $f \in L^\perp$). Thus, $s'$ does belong to $M^*$, and since we already have seen that $s' \in \operatorname{int} K^*$, $s'$ is strictly dual feasible.

Thus, $(x', s')$ is a strictly feasible primal-dual pair.

$4^0$. Now let us evaluate the quantity $\omega = V_*(x', s') - V_*(x, s)$. We have

$$\begin{aligned}\omega = V_*(x', s') - V_*(x, s) &= (\vartheta + \rho)\{\ln\langle s', x'\rangle - \ln\langle s, x\rangle\} + F^+(s') - F^+(s) \\ &= (\vartheta + \rho)\{\ln\langle \pi s'', x\rangle) - \ln\langle s, x\rangle\} + F^+(\pi s'') - F^+(s) \\ &= \vartheta\{\ln\langle s'', x\rangle - \ln\langle s, x\rangle\} + F^+(s'') - F^+(s) + \rho\ln\frac{\langle \pi s'', x\rangle}{\langle s, x\rangle}\end{aligned}$$

(note that $F^+$ is $\vartheta$-logarithmically homogeneous and $s' = \pi s''$).

Furthermore,

$$s'' = -F'(x) - F''(x)\xi, \qquad \pi = \frac{\langle s, x\rangle}{\vartheta + \rho},$$

so that

$$\langle s'', x\rangle = \langle -F'(x) - F''(x)\xi, x\rangle = \vartheta - \eta,$$

where

$$\eta = \langle F''(x)\xi, x\rangle$$

(we considered that $\langle -F'(x), x\rangle = \vartheta$; see (2.3.12)). Thus,

$$\ln\frac{\langle \pi s'', x\rangle}{\langle s, x\rangle} = -\ln\left(1 + \frac{\rho}{\vartheta}\right) + \ln\left(1 - \frac{\eta}{\vartheta}\right),$$

so that

$$\begin{aligned}\omega &= \vartheta\{\ln\vartheta + \ln\left(1 - \frac{\eta}{\vartheta}\right) - \ln\langle s, x\rangle\} \\ &\quad + \rho\ln\left(1 - \frac{\eta}{\vartheta}\right) - \rho\ln\left(1 + \frac{\rho}{\vartheta}\right) + F^+(s'') - F^+(s).\end{aligned} \tag{4.5.12}$$

Furthermore, by virtue of $s'' = s^* - F''(x)\xi$, we have $F^+(s'') = \phi(1)$, where

$$\phi(t) = F^+(s^* - tF''(x)\xi).$$

Note that, due to (4.5.11) and Theorem 2.1.1 (applied to 1-strongly self-concordant function $F^+$), we have

$$(1-t\delta)^2\phi''(0) \le \phi''(t) \le \frac{\phi''(0)}{(1-t\delta)^2}, \qquad 0 \le t \le 1,$$

so that

$$\begin{aligned} F^+(s'') = \phi(1) \le \phi(0) + \phi'(0) + \frac{\phi''(0)}{2(1-\delta)^2} &= F^+(-F'(x)) \\ &+ \langle -F''(x)\xi, (F^+)'(-F'(x))\rangle \\ &+ \frac{1}{2(1-\delta)^2}\langle F''(x)\xi, (F^+)''(-F'(x))F''(x)\xi)\rangle \\ = F^+(-F'(x)) + \langle -F''(x)\xi, (F^+)'(-F'(x))\rangle &+ \frac{1}{2(1-\delta)^2} \parallel s'' - s^* \parallel^2_{s^*}. \end{aligned}$$

From the origin of $F^+$, it follows that $(F^+)'(-F'(x)) = -x$, so that our computation leads to

$$F^+(s'') \le F^+(-F'(x)) + \langle F''(x)\xi, x\rangle + \frac{\delta^2}{2(1-\delta)^2} \tag{4.5.13}$$

(we considered that $\parallel s'' - s^* \parallel^2_{s^*} \le \delta^2$; see (4.5.11)).

From Proposition 2.4.1, it follows that $U(x, -F'(x)) = \vartheta \ln \vartheta - \vartheta$, or, due to the definition of $U(\cdot,\cdot)$, $F^+(-F'(x)) + F(x) + \vartheta \ln \langle -F'(x), x\rangle = \vartheta \ln \vartheta - \vartheta$. At the same time, (2.3.12) means that $\langle -F'(x), x\rangle = \vartheta$, and we obtain

$$F^+(-F'(x)) = -\vartheta - F(x).$$

Thus, (4.5.13) implies

$$F^+(s'') \le -\vartheta - F(x) + \eta + \frac{\delta^2}{2(1-\delta)^2} \tag{4.5.14}$$

(we considered that $\eta = \langle F''(x)\xi, x\rangle$). Now (4.5.12) implies that

$$\begin{aligned} \omega &\le \vartheta\{\ln\vartheta + \ln\left(1 - \frac{\eta}{\vartheta}\right) - \ln\langle s, x\rangle\} + \rho\left\{\ln\left(1-\frac{\eta}{\vartheta}\right) - \ln\left(1+\frac{\rho}{\vartheta}\right)\right\} \\ &\quad -\vartheta - F(x) + \eta + \frac{\delta^2}{2(1-\delta)^2} - F^+(s) \\ &= \{\vartheta\ln\vartheta - \vartheta - F(x) - F^+(s) - \vartheta\ln\langle s, x\rangle\} \\ &\quad +\vartheta\ln\left(1-\frac{\eta}{\vartheta}\right) + \rho\ln\left(1-\frac{\eta}{\vartheta}\right) - \rho\ln\left(1+\frac{\rho}{\vartheta}\right) + \eta + \frac{\delta^2}{2(1-\delta)^2} \\ &\le \vartheta\ln\left(1-\frac{\eta}{\vartheta}\right) + \rho\ln\left(1-\frac{\eta}{\vartheta}\right) - \rho\ln\left(1+\frac{\rho}{\vartheta}\right) + \eta + \frac{\delta^2}{2(1-\delta)^2} \end{aligned}$$

(the latter inequality follows from (2.4.8)). Since $\ln(1-\eta/\vartheta) \le -\eta/\vartheta$, it follows that

$$\omega \le \rho \ln\left(1 - \frac{\eta}{\vartheta}\right) - \rho \ln\left(1 + \frac{\rho}{\vartheta}\right) + \frac{\delta^2}{2(1-\delta)^2}. \tag{4.5.15}$$

$5^0$. To complete the proof, it remains to evaluate $\eta$. We have

$$\eta^2 = \langle F''(x)\xi, x\rangle^2 \le \langle F''(x)\xi, \xi\rangle \langle F''(x)x, x\rangle .$$

We have $\langle F''(x)\xi, \xi\rangle = \lambda^2$ (the definition of $\lambda$) and $\langle F''(x)x, x\rangle = \vartheta$ ((2.3.14)), so that $| \eta | \le \delta\vartheta^{1/2}$. The latter inequality, combined with (4.5.15) and the relations $\rho = \gamma\vartheta^{1/2}$, $\vartheta \ge 1$ implies that

$$\omega \le \gamma\delta - \frac{\gamma^2}{1+\gamma} + \frac{\delta^2}{2(1-\delta)^2}.$$

Thus, in the case of (4.5.9), inequality (4.5.6) also holds. □

### 4.5.4 Rate of convergence

The method in its basic version simply iterates the above updating, namely, forms the sequence of strictly feasible primal-dual pairs

$$(x_i, s_i) = \mathcal{PD}(\gamma, \delta)[(x_{i-1}, s_{i-1})],$$

where $(x_0, s_0) = (x^0, s^0)$ is the initial pair mentioned in §4.5.1, and $\gamma$, $\delta$ are positive absolute constants that should satisfy $(\mathcal{C})$.

At the $i$th iteration of a "large-step" version of the method, we first compute an intermediate strictly feasible primal-dual pair

$$(x_i^+, s_i^+) = \mathcal{PD}(\gamma, \delta)[(x_{i-1}, s_{i-1})],$$

and then, using an appropriate line search, transform this intermediate pair into a strictly feasible primal-dual pair $(x_i, s_i)$ in such a way that

$$V_*(x_i, s_i) \le V_*(x_i^+, s_i^+)$$

(some possibilities for the latter transformation are discussed below).

From Propositions 4.5.1 and 4.5.2, we immediately obtain the following rate-of-convergence statement.

**Theorem 4.5.1** *Let a primal-dual pair* (P), (D) *satisfy* (PD) *and let* $(x_i, s_i)$ *be the trajectory of the primal-dual method as applied to the pair*; *control parameters* $\gamma$, $\delta$ *of the method are assumed to satisfy* $(\mathcal{C})$. *Then the duality gap*

$$\nu(x_i, s_i) = \{\langle c, x_i\rangle - \mathsf{P}^*\} + \{\langle s_i, b\rangle - \mathsf{D}^*\}$$

*of the ith pair can be estimated as follows*:

$$
\begin{aligned}
\frac{\nu(x,s)}{\nu(x^0,s^0)} &\le R(x^0,s^0)\exp\left\{-\frac{V_*(x^0,s^0)-V_*(x,s)}{\gamma\vartheta^{1/2}}\right\} \\
&\le R(x^0,s^0)\exp\left\{-\frac{\Theta(\gamma,\delta)}{\gamma\vartheta^{1/2}}\,i\right\},
\end{aligned}
\tag{4.5.16}
$$

*where*

$$R(x^0,s^0)=\exp\left\{\frac{U(x^0,s^0)-\vartheta\ln\vartheta+\vartheta}{\gamma\vartheta^{1/2}}\right\}$$

*and*

$$\Theta(\gamma,\delta)=\min\{\delta-\ln(1+\delta)\,;\,\Omega(\gamma,\delta)\}.$$

As far as the theoretical estimate (4.5.16) is concerned, rational choice of $\gamma$, $\delta$ (i.e., that one minimizing $\kappa\equiv\gamma^{-1}\Theta(\gamma,\delta)$) is

$$\gamma=1.05,\qquad \delta=0.34,$$

which results in $\kappa=0.045$.

### 4.5.5 Large-step strategy

Note that the basic version $\mathcal{PD}$ of the primal-dual method, in fact, does not require knowledge of $F^+$, so that we can implement it if only $F$ is known. Of course, in the latter case, it is impossible to use the main advantage of a potential reduction method, we mean "large steps." Recall that all we are interested in is updating a current strictly feasible primal-dual pair into another pair of the same type with the value of the potential $V_*$ being at least by an absolute additive constant less than at the initial pair. Since we are interested in decreasing the potential by the maximum possible amount, we can try to achieve a larger decreasing, say, with the aid of minimization of $V_*$ in the direction prescribed by $\mathcal{PD}$; that is what is understood as "large step."

For procedure $\mathcal{PD}$, the most natural scheme of large steps seems to be the following. $\mathcal{PD}$ as applied to $(x,s)$ determines a pair of directions; the first one is the direction $\zeta=(1+\lambda)^{-1}\xi$ in the primal space $E$, and the second one is

$$\eta=\frac{\langle s,x\rangle}{\vartheta+\gamma\vartheta^{1/2}}[-F'(x)-F''(x)\xi]-s$$

in the dual space $E^*$. These directions belong to $L$, $L^\perp$, respectively, and $\mathcal{PD}$ sometimes prescribes to update $(x,s)$ as $x'=x+\zeta$, $s'=s$ and sometimes as $x'=x$, $s'=s+\eta$, which ensures at least an absolute constant decreasing of the potential. By virtue of the primal-dual symmetry, the latter is also ensured by the updating $\mathcal{PD}^*$, which is "symmetric," in the natural sense, to $\mathcal{PD}$; let us denote the primal and dual directions defined by $\mathcal{PD}^*$ by $\eta^*$

and $\zeta^*$, respectively. Thus, given a strictly feasible primal-dual pair $(x, s)$, we can compute four directions in $E \times E^*$, namely, $(\zeta, 0)$, $(\eta^*, 0)$, $(0, \eta)$, $(0, \zeta^*)$ with the following properties: The directions belong to $L \times L^{\perp}$, and there are at least two possibilities (the first is given by $\mathcal{PD}$; the second is given by $\mathcal{PD}^*$) to decrease "significantly" the potential by translating $(x, s)$ in an appropriately chosen direction from the above collection. This observation, being interpreted in spirit of "large steps," makes it reasonable to update a given pair by minimizing the potential over the intersection of $(\operatorname{int} K) \times (\operatorname{int} K^*)$ and the four-dimensional affine plane passing through the pair parallel to the above four directions.

# Chapter 5

# How to construct self-concordant barriers

In the previous chapters, we presented a number of polynomial-time interior-point methods for standard and conic settings of convex programming problems. To apply such a method to certain convex program, we should act as follows.

(A) First, we should reformulate it in the standard form

$$(f): \qquad \text{minimize } \langle c, x\rangle \quad \text{s.t. } x \in G$$

or in the conic form

$$(\mathsf{P}): \qquad \text{minimize } \langle c, x\rangle \quad \text{s.t. } x \in K\bigcap(L+b)$$

associated with a closed convex domain $G \subset E$, respectively, a closed convex pointed cone $K \subset E$ with a nonempty interior.

Note that, at least from the theoretical viewpoint, there are no difficulties with (A).

(B) Second, we should find a self-concordant barrier $F$ for $G$ (the case of problem $(f)$) or a logarithmically homogeneous self-concordant barrier $F$ for $K$ (the case of problem $(\mathsf{P})$); in the latter case we sometimes are interested not only in $F$, but also in its Legendre transformation $F^*$. Of course, we should do our best to choose the barrier with the smallest possible value of the parameter.

As far as (B) is concerned, there is theoretically no problem: The theorem on universal barrier (Theorem 2.5.1) states that an $n$-dimensional closed convex domain (cone) always admits a $O(n)$-self-concordant ($O(n)$-logarithmically homogeneous self-concordant) barrier. The universal barrier is, however, useless from the computational viewpoint. Indeed, to implement an interior-point method associated with a barrier, we need "an explicit" representation of the barrier, which allows us to compute at a reasonable cost the values of the barrier and its first- and second-order derivatives. The universal barrier, generally speaking, does not fit this requirement, excluding a very restricted number of simple domains for which it can be presented in a "computable" form. Recall that as of yet we have mentioned only two examples of "computable" barriers, i.e., the standard logarithmic barrier for a convex polytope (§2.3, Example 2) and the barrier for the Lebesgue set of a convex quadratic form (Lemma 3.3.1). Of course, even these two examples allow us to cover linear programming problems and (convex) quadratically constrained quadratic problems; nevertheless, it would be highly desirable to extend the set of "computable" self-concordant barriers. This is our subject in the present chapter.

Below, we develop a kind of *barrier calculus.* This calculus includes the following two parts:

(i) A list of ad hoc computable barriers for a spectrum of concrete convex domains and cones, which are of especial interest for convex programming (§§5.3, 5.4). This spectrum includes epigraphs of the standard univariate functions (the power function, the exponent, powers of the Euclidean norm, fractional-quadratic functions, and so on), the second-order cone, and the cone of positive semidefinite matrices, and so forth;

(ii) A set of *composition rules* that allow us to form a computable self-concordant barrier for a domain that is the result of certain standard operation with convex sets (like taking inverse image, intersection, projection, and so forth), provided that we are given self-concordant barriers for the operands, i.e., the sets involved into the operation (§§5.1, 5.2, 5.5).

Combining barriers mentioned in (i) with rules from (ii), we obtain a very wide spectrum of computable self-concordant barriers and, consequently, provide the possibility to solve in polynomial time various convex programs of appropriate analytical structure (see Chapter 6).

## 5.1 Operations with convex sets and barriers

It turns out that each standard operation $\Pi$ with convex domains (cones) that preserves convexity induces "an explicit" operation $\Pi^*$ with self-concordant barriers for the operands; the result of the latter operation is a self-concordant barrier for the result of $\Pi$. Since $\Pi^*$ is "explicit," the resulting barrier is "computable," provided that the initial barriers possess this property.

Below, we list the standard operations $\Pi$ and present the corresponding $\Pi^*$.

### 5.1.1 Simple combination rules

*Inverse images under affine mappings.* We have (see Propositions 2.3.1(i) and 2.3.3(i)) the following result.

**Proposition 5.1.1** *Let $G$ be a closed convex domain in $E$, let $F$ be a $\vartheta$-self-concordant barrier for $G$, and let $\mathcal{A} : E^+ \to E$ be an affine mapping such that $\mathcal{A}(E^+) \bigcap \operatorname{int} G \neq \emptyset$. Then the function*

$$F^+(y) = F(\mathcal{A}(y)) : \operatorname{int} \mathcal{A}^{-1}(G) \to \mathbf{R}$$

*is a $\vartheta$-self-concordant barrier for the closed convex domain $\mathcal{A}^{-1}(G) \subset E^+$.*

*If, in addition, $G$ is a cone, $F$ is $\vartheta$-logarithmically homogeneous, and $\mathcal{A}$ is homogeneous, then $F^+$ is a $\vartheta$-logarithmically homogeneous self-concordant barrier for the closed convex cone $\mathcal{A}^{-1}(G)$.*

*Direct products.* We have (see Propositions 2.3.1(iii) and 2.3.3(iii)) the following result.

**Proposition 5.1.2** *Let $G_i$ be closed convex domains in $E_i$ and let $F_i(x_i)$ be $\vartheta_i$-self-concordant barriers for $G_i$, $1 \le i \le m$. Then the function*

$$F(x_1, \dots, x_m) = F_1(x_1) + \cdots + F_m(x_m) : \operatorname{int}(G_1 \times \cdots \times G_m) \to \mathbf{R}$$

*is a $(\sum_{i=1}^m \vartheta_i)$-self-concordant barrier for the domain $G_1 \times \cdots \times G_m \subset E_1 \times \cdots \times E_m$.*

*If, in addition, $G_i$ are cones and $F_i$ are $\vartheta_i$-logarithmically homogeneous, $1 \le i \le m$, then $F$ is a $(\sum_{i=1}^m \vartheta_i)$-logarithmically homogeneous self-concordant barrier for the cone $G_1 \times \cdots \times G_m \subset E_1 \times \cdots \times E_m$.*

*Intersections.* We have (see Propositions 2.3.1(ii) and 2.3.3(ii)) the following result.

**Proposition 5.1.3** *Let $G_i$ be closed convex domains in $E$ and let $F_i(x_i)$ be $\vartheta_i$-self-concordant barriers for $G_i$, $1 \le i \le m$. Assume that the set $G = \bigcap_{i=1}^m G_i$ has a nonempty interior. Then the function*

$$F(x) = F_1(x) + \cdots + F_m(x) : \operatorname{int} G \to \mathbf{R}$$

*is a $(\sum_{i=1}^m \vartheta_i)$-self-concordant barrier for the domain $G \subset E$.*

*If, in addition, $G_i$ are cones and $F_i$ are $\vartheta_i$-logarithmically homogeneous, $1 \le i \le m$, then $F$ is a $(\sum_{i=1}^m \vartheta_i)$-logarithmically homogeneous self-concordant barrier for the cone $G \subset E$.*

*Conic hulls and projective transformations.* Let $G$ be a closed convex domain in $E$. As we remember (see §4.1), to reformulate a standard problem associated with $G$ in a conic form, it suffices to identify $E$ with the hyperplane

$$\Pi = \{(x,t) \in E^+ = E \times \mathbf{R} \mid t = 1\}$$

in a $(\dim E + 1)$-dimensional space and then introduce as the cone involved into (P) the *conic hull*

$$K(G) = \operatorname{cl}\left\{(x,t) \in E^+ \mid t > 0,\ \frac{x}{t} \in G\right\}$$

of $G$. It is easily seen that

$$\operatorname{int} K(G) = K' \equiv \left\{(x,t) \in E^+ \mid t > 0,\ \frac{x}{t} \in \operatorname{int} G\right\}, \tag{5.1.1}$$

while the image of $G$ under the above embedding $E \to \Pi$ is the intersection of $P$ and $K(G)$, as follows:

$$\Pi \bigcap K(G) = \{(x,1) \mid x \in G\}. \tag{5.1.2}$$

Now assume that we know a $\vartheta$-self-concordant barrier for $G$. How can we transform it into a logarithmically homogeneous self-concordant barrier for $K(G)$? The answer is given by the following result.

**Proposition 5.1.4** *Let $G$ be a closed convex domain in $E$ and let $F$ be a $\mu$-self-concordant barrier for $G$. Then, for an appropriate* absolute constant $\theta$ *(we can take $\theta = 20$), the function*

$$F^+(x,t) = \theta^2\left(F\left(\frac{x}{t}\right) - 2\mu \ln t\right) : \operatorname{int} K(G) \to \mathbf{R}$$

*is a $(2\theta^2\mu)$–logarithmically homogeneous self-concordant barrier for the conic hull $K(G)$ of $G$.*

**Proof.** $F^+$ clearly is $C^3$-smooth on $K'$ and satisfies (2.3.10) with $\vartheta = 2\theta^2\mu$. It remains to prove that $F^+$ is strongly 1-self-concordant on $K'$.

$1^0$. Let us prove that $F^+$ is 1-self-concordant. To this purpose, let us fix $z = (x,t) \in K'$ and let $w = (h, -s) \in E^+$. Let us compute the derivatives up to the order 3 of $F^+$ at the point $z$ in the direction $w$. For $\tau \in \mathbf{R}$, we have, under the notation $\sigma = t^{-1}s$, $\xi = t^{-1}x \in \operatorname{int} G$, $\eta = t^{-1}h$, $\omega = \sigma\xi + \eta$,

$$\frac{x + \tau h}{t - \tau s} = \xi + \rho(\tau)\omega,$$

where

$$\rho(\tau) = \frac{\tau}{1 - \tau\sigma}.$$

Let

$$\phi(r) = F(\xi + r\omega);$$

then

$$F^+(z + \tau w) = \theta^2(\phi(\rho(\tau)) - 2\mu \ln(1 - \sigma\tau) - 2\mu \ln t).$$

Therefore

$$\begin{aligned} d_1 &\equiv DF^+(z)[w] = \theta^2(2\mu\sigma + \pi_1), \\ d_2 &\equiv D^2F^+(z)[w,w] = \theta^2(2\mu\sigma^2 + 2\sigma\pi_1 + \pi_2^2), \\ d_3 &\equiv D^3F^+(z)[w,w,w] = \theta^2(4\mu\sigma^3 + 6\sigma^2\pi_1 + 6\sigma\pi_2^2 + \pi_3), \end{aligned} \tag{5.1.3}$$

where

$$\pi_1 = \phi'(0), \qquad \pi_2 = (\phi''(0))^{1/2} = (D^2F(\xi)[\omega,\omega])^{1/2},$$
$$\pi_3 = \phi'''(0) = D^3F(\xi)[\omega,\omega,\omega].$$

Since $F$ is a $\mu$-self-concordant barrier for $G$, we have

$$\pi_1^2 \le \mu\pi_2^2; \qquad |\pi_3| \le 2\pi_2^3. \tag{5.1.4}$$

Thus, if $\nu = \mu^{1/2}$, $p = |\nu^{-1}\pi_1|$, then $p \le \pi_2$ and

$$d_2 \ge \theta^2(2\nu^2\sigma^2 - 2\sigma\nu\pi_2 + \pi_2^2) \ge \theta^2 \max\{(\nu\sigma)^2, \pi_2^2/2\} \ge 0. \tag{5.1.5}$$

We see that $F$ is convex. Now (5.1.3)–(5.1.5) lead to the relation

$$
\begin{aligned}
| d_3 | &\leq 2\theta^2 \left\{ \frac{1}{\nu}\{2 \mid \nu\sigma \mid^3 + 3(\nu\sigma)^2 \pi_2 + 3 \mid \nu\sigma \mid \pi_2^2\} + \pi_2^3 \right\} \\
&\leq 2\theta^2 \left\{ 2\left(\frac{d_2}{\theta^2}\right)^{3/2} + 3\left(\frac{d_2}{\theta^2}\right)\left(\frac{2d_2}{\theta^2}\right)^{1/2} + 3\left(\frac{d_2}{\theta^2}\right)^{1/2}\left(\frac{2d_2}{\theta^2}\right) \right. \\
&\qquad \left. + + \left(\frac{2d_2}{\theta^2}\right)^{3/2} \right\} \leq 2\frac{20}{\theta} d_2^{3/2},
\end{aligned}
$$

so that

$$
| d_3 | \leq 2 d_2^{3/2} \tag{5.1.6}
$$

for $\theta \geq 20$ (of course, the appropriate value of $\theta$ can be lowered by more accurate evaluations). Thus, $F^+$ is self-concordant on $K'$.

$2^0$. It remains to verify that $F^+(x_i, t_i) \to \infty$ for each sequence $\{(x_i, t_i) \in K'\}$ converging to a boundary point $(x, t)$ of $K(G)$. Of course, we can assume that the sequence $\{F^+(x_i, t_i)\}$ converges to a point in the extended real axis. If $t > 0$, then the points $u_i = t_i^{-1} x_i$ converge to a boundary point of $G$, so that $F(t_i^{-1} x_i)$ tends to $\infty$ (since $F \in S_1^+(\operatorname{int} G, E)$), and therefore $F^+(x_i, t_i) \to \infty$. Now let $t = 0$. Let $w$ be an interior point of $G$. Then, by virtue of (2.3.3), we have $F(u_i) \geq F(w) + \mu \ln(1 - \pi_{u_i}(w))$, where $\pi_{u_i}$ is the Minkowsky function of $G$ with the pole at $u_i$. Let $\| \cdot \|$ be a norm on $E$. Clearly,

$$
1 - \pi_{u_i}(w) \geq \frac{c(w)}{c(w) + \| u_i - w \|}
$$

for certain $c(w) > 0$; since $x_i$ converge, we have $\| u_i - w \| \leq C t_i^{-1}$ for certain constant $C$. Thus,

$$
\begin{aligned}
F^+(x_i, t_i) &= \theta^2 \{F(u_i) - 2\mu \ln t_i\} \\
&\geq \theta^2 \{F(w) + \mu \ln \frac{c(w) t_i}{c(w) t_i + C} - 2\mu \ln t_i\} \to \infty, \qquad i \to \infty,
\end{aligned}
$$

since $t_i$ tend to 0. □

**Corollary 5.1.1** *Let $G$ be a closed convex domain in $E$, let $F$ be a $\vartheta$-self-concordant barrier for $G$, and let $\mathcal{A}(x) = (\langle \alpha, x \rangle + \beta)^{-1}(Ax + b)$ be a projective transformation of $E$ such that $\langle \alpha, x \rangle + \beta$ is positive on* $\operatorname{int} G$. *Let*

$$
G^+ = \operatorname{cl} \mathcal{A}(\operatorname{int} G)
$$

*be the image of $G$ under this transformation. Then the function*

$$
F^+(x) = \theta^2 (F(\mathcal{A}^{-1}(x)) + 2\vartheta \ln(\left\langle \alpha, \mathcal{A}^{-1}(x) \right\rangle + \beta))
$$

*is a $(2\theta^2\vartheta)$-self-concordant barrier for $G^+$, $\theta$ being the same absolute constant as in Proposition* 5.1.4.

**Proof.** Let $\mathcal{B}(x) = (\langle\sigma, x\rangle + \delta)^{-1}(Cx + d)$ be the transformation inverse to $\mathcal{A}$; without loss of generality, we can assume that $\langle\sigma, x\rangle + \delta$ is positive on $\operatorname{int} G^+$. Let $K(G) \subseteq E \times \mathbf{R}$ be the conic hull of $G$. Consider the affine mapping

$$\mathcal{G}: \ E \to E \times \mathbf{R}: \ \mathcal{G}(x) = (Cx + d, \langle\sigma, x\rangle + \delta).$$

Let us verify that

$$\operatorname{int} G^+ = \mathcal{G}^{-1}(\operatorname{int} K(G)).$$

Indeed,

$$\begin{aligned}\{x \in \operatorname{int} G^+\} &\Leftrightarrow \{\{\mathcal{B}(x) \in \operatorname{int} G\}\&\{\langle\sigma, x\rangle + \delta > 0\}\};\\ &\Leftrightarrow \{\langle\sigma, x\rangle + \delta > 0, (\langle\sigma, x\rangle + \delta)^{-1}(Cx + d) \in \operatorname{int} G\}\\ &\Leftrightarrow \{(Cx + d, \langle\sigma, x\rangle + \delta) \in \operatorname{int} K(G)\} \ \Leftrightarrow \ \{\mathcal{G}(x) \in \operatorname{int} K(G)\}.\end{aligned}$$

Now, in view of Proposition 5.1.4, the function $\Phi(x, t) = \theta^2(F(t^{-1}x) - 2\vartheta \ln t)$ is a $2\theta^2\vartheta$-self-concordant barrier for $K(G)$, so that its superposition with $\mathcal{G}$, i.e., the function

$$\theta^2(F(\mathcal{B}(x)) - 2\vartheta \ln(\langle\sigma, x\rangle + \delta)),$$

is a $2\theta^2\vartheta$-self-concordant barrier for $\operatorname{cl} \mathcal{G}^{-1}(\operatorname{int} K(G)) = G^+$ (see Proposition 5.1.1). The latter function, up to an additive constant, is $F^+$. □

The next pair of operations with convex sets—projection and summation—imply computationally not so simple transformations of barriers and are mainly of theoretical interest.

*Images under affine mappings.* We have the following result.

**Proposition 5.1.5** *Let $G$ be a closed convex domain in $E$, let $F$ be a $\vartheta$-self-concordant barrier for $G$, and let $\mathcal{A} : E \to E^+$ be an affine transformation such that*

$$(5.1.7) \qquad \mathcal{A}(E) = E^+, \ \mathcal{A}^{-1}(\mathcal{A}(x)) \bigcap G \ \textit{ is bounded for each } x \in G.$$

*Then $G^+ = \mathcal{A}(G)$ is a closed convex domain in $E^+$, and the function*

$$F^+(u) = \min\{F(x) \mid x \in \mathcal{A}^{-1}(u) \textstyle\bigcap \operatorname{int} G\} : \operatorname{int} G^+ \to \mathbf{R}$$

*is well defined and is a $\vartheta$-self-concordant barrier for $G^+$.*

*If, in addition to the above assumptions, $G$ is a cone, $\mathcal{A}$ is homogeneous, and $F$ is $\vartheta$-logarithmically homogeneous, then $G^+$ also is a cone, and $F^+$ is $\vartheta$-logarithmically homogeneous.*

**Proof.** $1^0$. In view of $\mathcal{A}(E) = E^+$, the set $G' = \mathcal{A}(\operatorname{int} G)$ is open (and, of course, convex), and, clearly, $G' \subseteq G^+ \subseteq \operatorname{cl} G'$. Let us verify that $G^+ = \operatorname{cl} G'$ (so that $G^+$ is a closed convex domain). From boundedness of the sets

$$S(u) = \mathcal{A}^{-1}(u) \bigcap G$$

for $u \in G^+$, in view of the standard convexity arguments, it follows immediately that the multivalued mapping $S(\cdot)$ is locally bounded: For any bounded $Q \subseteq G^+$, the set $\bigcup_{u \in Q} S(u)$ is also bounded. The latter statement, in view of the standard compactness reasons, implies that $G^+$ is closed.

$2^0$. For $u \in G'$ $(= \operatorname{int} G^+)$, the set $S(u)$ is a bounded closed convex set that intersects $\operatorname{int} G$; thus, by Proposition 5.1.1, the restriction of $F$ onto the relative interior of $S(u)$ is a $\vartheta$-self-concordant barrier for the set $S(u)$ (regarded as a closed convex domain in its affine hull). In view of Proposition 2.3.2 (ii), $F$ attains its minimum over the relative interior of $S(u)$ at a certain point $x(u)$ (recall that $S(u)$ is bounded). Thus, $F^+$ is well defined.

$3^0$. Let us verify that $F^+$ is a barrier, i.e., that $F^+(u_i) \to \infty$ along every sequence $\{u_i \in G'\}$ converging to a boundary point of $G^+$. Since $S$ is locally bounded, for such a sequence $\{u_i\}$, the sequence $\{x(u_i)\}$ also is bounded; all limit points of the latter sequence belong to the boundary of $G$ (since otherwise $\{u_i\}$ would converge to an interior point of $G^+$). Since $F$ is a barrier for $G$, we have $F^+(u_i) = F(x(u_i)) \to \infty$.

$4^0$. It remains to prove that $F^+$ is strongly 1-self-concordant and that $\lambda(F^+, u) \leq \vartheta^{1/2}$, $u \in G'$.

Let us first show that we can reduce the situation to the case when $G$ does not contain straight lines. Without loss of generality, assume that $\mathcal{A}(x) = Ax$ is homogeneous. Assume that $G$ contains straight lines and let $E_0$ be the subspace $\{h \mid G + \mathbf{R}h = G\}$. Note that (5.1.7) implies that $E_0 \bigcap \operatorname{Ker} A = \{0\}$, so that we can find a subspace $E_1 \supset \operatorname{Ker} A$ in $E$ such that $E = E_0 \times E_1$. Let $G_1 = G \bigcap E_1$; then $G_1$ is a closed convex domain in $E_1$. Since $\operatorname{Ker} A \subset E_1$, we have $E^+ = (AE_0) \times (AE_1)$, and the above representation of $G$ implies that $G^+ = (AE_0) \times G_1^+$, where $G_1^+ = A(G_1)$. Note that $F$ is constant along the affine planes $x + E_0$, $x \in \operatorname{rint} G_1$ (we used Proposition 2.3.2(ii)); therefore $F^+$ is constant along the planes $u + AE_0$, $u \in \operatorname{rint} G_1^+$, and, to prove that $F^+$ is a $\vartheta$-self-concordant barrier for $G^+$, it suffices to prove the latter statement for the restriction of $F^+$ onto $\operatorname{rint} G_1^+$. The pair $(G_1^+, F^+ \mid_{\operatorname{rint} G_1^+})$ is obtained from $(G_1, F \mid_{\operatorname{rint} G_1^+})$ with the aid of the same construction as the one transforming $(G, F)$ into $(G^+, F^+)$, and, when proving the statement, we can from the very beginning replace $E$ with $E_1, G$ with $G_1$, and $F$ with its restriction onto the relative interior of $G_1$. Thus, we can suppose that $G$ does not contain straight lines.

Let us fix $u_0 \in G'$ and let $x_0 = x(u_0)$. After appropriate translations of the origins in $E$ and $E^+$, we can assume that $u_0 = 0 \in E^+$ and $x_0 = 0 \in E$; since $u = \mathcal{A}(x(u))$, it means that $\mathcal{A}$ is homogeneous: $\mathcal{A}(x) = Ax$. Let us provide $E$ with the Euclidean structure $\mathcal{E}$ associated with the inner product $(p, q) = D^2F(0)[p, q]$ (this is a nondegenerate scalar product in view of Proposition 2.3.2(ii); recall that $G$ does not contain straight lines) and let $E_0 = \operatorname{Ker} A$, $E_1$ be the $\mathcal{E}$-orthogonal complement to $E_0$. Also, let $\pi$ be the $\mathcal{E}$-orthoprojector of

$E$ onto $E_1$; note that there exists a linear isomorphism $\tau : E_1 \to E^+$ such that

$$\mathcal{A} = \tau \circ \pi. \tag{5.1.8}$$

Let us denote a point of $E$ as $x = (v, w)$, where $v \in E_1$, $w \in E_0$ and let $f'_v$, $f'_w$, denote the partial derivatives of a mapping $f$ with respect to $v$ and $w$; we also write $f''_{vw}$ instead of $(f'_w)'_v$, and so on. Under this notation, we have

$$x(u) = (\tau^{-1}u, w(\tau^{-1}u)), \tag{5.1.9}$$

where the function $w(v)$ defined in a neighborhood of 0 in $E_1$ and taking values in $E_0$ is given by the equation

$$F'_w(v, w) = 0; \tag{5.1.10}$$

note that this equation is satisfied by the pair $(v, w) = (0, 0)$ (since we have assumed that $u_0 = 0, x_0 = 0$), and its left-hand side $\Phi(v, w)$ is $C^2$-smooth and satisfies the relation $\Phi'_w(0, 0) = \mathrm{Id}_{E_0}$, so that the equation does define (locally) a $C^2$-smooth function $w(v)$, $w(0) = 0$.

Note that the identity $F'_w(v, w(v)) = 0$ implies

$$Dw(v)[h] = -[F''_{ww}(v, w(v))]^{-1} F''_{vw}(v, w(v))h, \;\; h \in E_1, \tag{5.1.11}$$

Let

$$H(v) = F(v, w(v)),$$

so that in view of (5.1.9) one has (locally)

$$F^+(u) = H(\tau^{-1}u). \tag{5.1.12}$$

Let us compute the derivatives of $H$ at $v = 0$ in a direction $h \in E_1$,

$$DH(v)[h] = (F'_v(v, w(v)), h) + (F'_w(v, w(v)), w'(v)h) \equiv (F'_v(v, w(v)), h)$$

(we have taken into account (5.1.10)). Since $w(\cdot)$ is $C^2$-smooth and $F$ is $C^3$-smooth, $DH(v)$ is $C^2$-smooth, so that $H$ is $C^3$-smooth.

Furthermore,

$$\begin{aligned} D^2H(v)[h, h] &= (F''_{vv}(v, w(v))h, h) + (F''_{wv}(v, w(v))(w'(v)h), h) \\ &= (F''_{vv}(v, w(v))h, h) \\ &\quad -(F''_{wv}(v, w(v))[F''_{ww}(v, w(v))]^{-1} F''_{vw}(v, w(v))h, h) \end{aligned}$$

(we have used (5.1.11)).

Last,

(5.1.13)
$$\begin{aligned} D^3H(v)[h, h, h] &= ((F'''_{vvv}(v, w(v))h)h, h) + ((F'''_{wvv}(v, w(v))(w'(v)h))h, h) \\ &\quad -(\{(F''_{wv}(v, w(v))[F''_{ww}(v, w(v))]^{-1} F''_{vw}(v, w(v))h, h)\}'_v, h). \end{aligned}$$

Now set $v = 0$ and recall that $w(0) = 0$ and $F''_{vw}(0,0) = 0$ (the latter relation holds, since $E_0$ and $E_1$ are orthogonal to each other with respect to the Euclidean structure defined by $D^2F(0,0)$). We obtain

$$DH(0)[h] = (F'_v(0,w(0)),h) = DF(0,0)[(h,0)],$$
$$D^2H(0)[h,h] = (F''_{vv}(0,w(0))h,h) = D^2F(0,0)[(h,0),(h,0)],$$
$$D^3H(0)[h,h,h] = ((F'''_{vvv}(0,w(0))h)h,h) = D^3F(0,0)[(h,0),(h,0),(h,0)]$$

(we considered that $F''_{vw}(0,0) = 0$ implies that $w'(0) = 0$ (see (5.1.11)) and that the last term in the right-hand side of (5.1.13) vanishes at $v = 0$).

Since $F$ is a $\vartheta$-self-concordant barrier, it follows that

$$| DH(0)[h] | \le \vartheta^{1/2}\{D^2H(0)[h,h]\}^{1/2},$$
$$| D^3H(0)[h,h,h] | \le 2\{D^2H(0)[h,h]\}^{3/2}.$$

Since locally $F^+(u) = H(\tau^{-1}u)$ (see (5.1.12)) and $\tau$ is a linear isomorphism of $E_1$ and $E^+$, the latter inequalities imply that

$$| DF^+(u_0)[p] | \le \vartheta^{1/2}\{D^2F^+(u_0)[p,p]\}^{1/2},$$
$$| D^3F^+(u_0)[p,p,p] | \le 2\{D^2F^+(u_0)[p,p]\}^{3/2}$$

for all $p \in E^+$. Since $u_0$ is an arbitrary point of $G' = \operatorname{int} G^+$, we conclude that $F^+$ is strongly 1-self-concordant on $G'$ and that $\lambda(F^+,u) \le \vartheta^{1/2}$, $u \in G'$.

The concluding statement concerning the "conic" case is evident. □

*Arithmetic summation.* We have the following result.

**Proposition 5.1.6** *Let $G_i$ be closed bounded convex domains in $E$ and let $F_i$ be $\vartheta_i$-self-concordant barriers for $G_i$, $1 \le i \le m$. Denote by $F_i^*$ the Legendre transformations of $F_i$ and let $F^+$ be the Legendre transformation of the function $\sum_{i=1}^m F_i^*$. Then $F^+$ is a $(\sum_{i=1}^m \vartheta_i)$-self-concordant barrier for the set*

$$G^+ = G_1 + \cdots + G_m = \{x \in E \mid \exists x_i \in G_i : x = x_1 + \cdots + x_m\}.$$

Proof of the statement follows immediately from Theorems 2.4.2 and 2.4.3. Note that we can also prove this fact using Proposition 5.1.5 and the following representation of the function $F^+$:

$$F^+(x) \equiv \min\{F_1(x_1) + \cdots + F_m(x_m) \mid x_i \in \operatorname{int} G_i,\ x_1 + \cdots + x_m = x\}.$$

### 5.1.2 Inverse image under nonlinear mappings

Let $G$ be a closed convex domain in $E$, $Q$ be an open convex subset in a linear finite-dimensional space $E^+$ and let $\mathcal{A} : H \to E$ be a continuous not necessarily linear mapping. We would like to transform a self-concordant barrier for $G$ into a similar barrier for the inverse image $G^+ = \mathcal{A}^{-1}(G)$ of $G$ under the mapping $\mathcal{A}$. Of course, it is necessary to make some assumptions about $\mathcal{A}$, since generally $G^+$ may be nonconvex. We start with describing mappings which for sure provide convexity of $G^+$.

### A. Mappings concave with respect to a cone

Let $E$, $E^+$ be finite-dimensional vector spaces, $K$ be a closed convex cone in $E$, and $H$ be a convex set in $E^+$. A mapping $\mathcal{A} : H \to K$ is called *concave with respect to* $K$ ($K$-concave for short) if it is continuous on $H$ and for all $x, y \in H$ and all $t \in [0, 1]$ the vector

$$h = \mathcal{A}(tx + (1-t)y) - t\mathcal{A}(x) - (1-t)\mathcal{A}(y)$$

belongs to $K$.

The cone $K$ defines a partial ordering on $E$: $a \geq_K b$ if and only if $a - b \in K$. The definition of $K$-concavity of $\mathcal{A}$ means simply that

$$t\mathcal{A}(x) + (1-t)\mathcal{A}(y) \leq_K \mathcal{A}(tx + (1-t)y), \qquad x, y \in H, \quad t \in [0,1].$$

Of course, the usual concave real-valued functions are precisely the $\mathbf{R}_+$-concave mappings.

Our first statement claims that convexity of the inverse image $\mathcal{A}^{-1}(G)$ of a convex set $G$ is ensured by concavity of $\mathcal{A}$ with respect to the recessive cone

$$\mathcal{R}(G) = \{h \in E \mid z + th \in G \;\; \forall z \in G \, \forall t \geq 0\}$$

of $G$.

**Lemma 5.1.1** *Let $H$ be an open convex set in $E^+$, let $G$ be a closed convex domain in $E$, and let a mapping $\mathcal{A} : H \to E$ be concave with respect to the recessive cone $\mathcal{R}(G)$ of $G$. Assume that $\mathcal{A}(H)$ intersects* $\operatorname{int} G$. *Then the set* $G^+ = \operatorname{cl} \mathcal{A}^{-1}(\operatorname{int} G)$ *is a closed convex domain in $E^+$, and* $\operatorname{int} G^+ = \mathcal{A}^{-1}(\operatorname{int} G)$.

**Proof.** Since $\mathcal{A}$ is continuous and its image intersects $\operatorname{int} G$, the set $H^+ \equiv \mathcal{A}^{-1}(\operatorname{int} G)$ is nonempty and open in $E^+$. Let us verify that this set is convex. Let $x, y \in H^+$, $t \in [0,1]$. We have

$$\mathcal{A}(tx+(1-t)y) = \{t\mathcal{A}(x)+(1-t)\mathcal{A}(y)\}_1+\{\mathcal{A}(tx+(1-t)y)-t\mathcal{A}(x)-(1-t)\mathcal{A}(y)\}_2.$$

The vector $\{\}_1$ belongs to $\operatorname{int} G$, since $\operatorname{int} G$ is convex and $x, y \in \mathcal{A}^{-1}(\operatorname{int} G)$, and the vector $\{\}_2$ belongs to $\mathcal{R}(G)$, since $\mathcal{A}$ is concave with respect to the latter cone. Therefore $\{\}_1 + \{\}_2 \in \operatorname{int} G$ (an immediate corollary of the definition of $\mathcal{R}(G)$), and $tx + (1-t)y \in H^+$.

Since $H^+$ is an open and nonempty convex set in $E^+$, its closure $G^+$ is a closed convex domain in $E^+$ and $\operatorname{int} G^+ = H^+$. $\square$

The following statement is quite standard.

**Lemma 5.1.2** *Let $H$ be an open convex set in $E^+$ and let $K$ be a convex cone in $E$. A $C^2$-smooth mapping $\mathcal{A} : H \to E$ is $K$-concave if and only if $(-D^2\mathcal{A}(u)[h,h]) \in K$ for all $u \in H$, $h \in E^+$, and, if $\mathcal{A}$ is $K$-concave, then $\mathcal{A}(x+h) \leq_K \mathcal{A}(x) + D\mathcal{A}(x)[h]$, $x, x+h \in H$.*

**Proof.** For a $C^2$-smooth $\mathcal{A}$ and $x, y \in H$, we have, under the notation $z = tx + (1-t)y$,

$$\mathcal{A}(z) - t\mathcal{A}(x) - (1-t)\mathcal{A}(y) = -t\int_0^1 (1-\tau)D^2\mathcal{A}(z+\tau(x-z))[x-z, x-z]d\tau$$
$$-(1-t)\int_0^1 (1-\tau)D^2\mathcal{A}(z+\tau(y-z)))$$
$$[y-z, y-z]d\tau.$$

If $(-D^2\mathcal{A}(u)[h,h])$ always belongs to $K$, then the right-hand side of the latter equality clearly belongs to $K$, which proves the if-part. To prove the remaining part of the statement, it suffices to set in the above equality $t = \frac{1}{2}$, $x = u + \varepsilon h$, $y = u - \varepsilon h$ and to look at the principal term as $\varepsilon \to +0$.

Now let $\mathcal{A}$ be $K$-concave and let $x, x+h \in H$. We have

$$\mathcal{A}(x+h) = \mathcal{A}(x) + D\mathcal{A}(x)[h] - r, \qquad r = -\int_0^1 D^2\mathcal{A}(x+th)[h,h](1-t)dt;$$

since, as we already know, $-D^2\mathcal{A}(x+th)[h,h] \in K$, we have $r \in K$, whence

$$\mathcal{A}(x+h) \leq_K \mathcal{A}(x) + D\mathcal{A}(x)[h],$$

as claimed. □

**Examples of concave mappings.**

**Example 1.** As we just have indicated, the usual concave real-valued function is the same as an $\mathbf{R}_+$-concave mapping taking values in $\mathbf{R}$. Convexity of a real-valued function is the same as its $\mathbf{R}_-$-concavity.

**Example 2.** An affine mapping $\mathcal{A}$ clearly is concave with respect to *any* cone $K$ in the image space.

**Example 3.** Let $S_m$ be the space of symmetric $m \times m$ matrices, let $S_m^+$ be the cone of positive semidefinite matrices from $S_m$, and let $L_{p,q}$ be the space of $p \times q$ matrices.

**Example 3.1.** The mappings

$$\mathcal{A}(x) = -xx^T : L_{m,n} \to S_m$$

and

$$\mathcal{B}(x) = -x^T x : L_{n,m} \to S_m$$

are $S_m^+$-concave (an immediate consequence of Lemma 5.1.2).

**Example 3.2.** The mapping

$$\mathcal{A}(x) = -x^{-1} : \operatorname{int} S_m^+ \to S_m$$

is $S_m^+$-concave.

Indeed, we have

$$DA(x)[h] = x^{-1}hx^{-1},$$
$$-D^2\mathcal{A}(x)[h,h] = 2x^{-1}hx^{-1}hx^{-1} = 2x^{-1/2}(x^{-1/2}hx^{-1/2})^2x^{-1/2} \in S_m^+.$$

**Example 3.3.** The mapping

$$\mathcal{A}(x) = x^{1/2} : S_m^+ \to S_m$$

is $S_m^+$-concave.

**Proof.** $\mathcal{A}$ clearly is continuous on $S_m^+$, so that, to prove $S_m^+$-concavity of $\mathcal{A}$ on $S_m^+$, it suffices to verify that $\mathcal{A}$ is $S_m^+$-concave on $\operatorname{int} S_m^+$. Furthermore, $\mathcal{A}$ is $C^\infty$-smooth on $\operatorname{int} S_m^+$ since locally $\mathcal{A}$ can be represented by the Cauchy integral

$$\mathcal{A}(x) = \frac{1}{2\pi i}\oint_\gamma z^{1/2}(zI - x)^{-1}dz,$$

$\gamma$ being an appropriate curve in the half-plane $\operatorname{Re} z > 0$. To establish concavity, let us use Lemma 5.1.2. We have

$$\mathcal{A}^2(x) \equiv x,$$

whence

$$\mathcal{A}(x)\cdot D\mathcal{A}(x)[h] + D\mathcal{A}(x)[h]\cdot\mathcal{A}(x) = h,$$

which, in turn, leads to

$$2(D\mathcal{A}(x)[h])^2 + \mathcal{A}(x)\cdot D^2\mathcal{A}(x)[h,h] + D^2\mathcal{A}(x)[h,h]\cdot\mathcal{A}(x) = 0.$$

The latter relation is the Lyapunov equation with respect to $D^2\mathcal{A}(x)[h,h]$; since $\mathcal{A}(x)$ is positive definite and symmetric, the unique solution to this equation is

$$D^2\mathcal{A}(x)[h,h] = -2\int_0^\infty \exp\{-\mathcal{A}(x)t\}(D\mathcal{A}(x)[h])^2\exp\{-\mathcal{A}(x)t\}dt,$$

and we see that $-D^2\mathcal{A}(x)[h,h] \in S_m^+$, as claimed. □

## B. Concave mappings compatible with a convex domain. The main result

Our aim is to associate with a mapping $\mathcal{A}$ a rule for updating self-concordant barriers $F$ for $G$ into similar barriers for $G^+ = \mathcal{A}^{-1}(G)$. It turns out that the simplest rule $F \to F\circ\mathcal{A}$, which is good for affine $\mathcal{A}$, can be naturally extended onto certain class of *nonlinear* $\mathcal{A}$. Let us start with necessary definitions.

**Definition 5.1.1** *Let $\beta$ be a nonnegative real; let $E$, $E^+$ be finite-dimensional linear spaces; let $K$ be a closed convex cone in $E$; let $\Gamma$ be a closed convex domain in $E^+$. A mapping $\mathcal{A} : \operatorname{int}\Gamma \to E$ is called $(K,\beta)$-compatible with the domain $\Gamma$, if*

(i) $\mathcal{A}$ *is* $C^3$*-smooth on* $\operatorname{int}\Gamma$*;*
(ii) $\mathcal{A}$ *is concave with respect to* $K$*;*
(iii) *For all* $z \in \operatorname{int}\Gamma$, $h \in E^+$, *such that* $z \pm h \in \Gamma$, *we have*

$$D^3\mathcal{A}(z)[h,h,h] \leq_K -3\beta D^2\mathcal{A}(z)[h,h].$$

Note that (iii) is equivalent to

$$\{z \in \operatorname{int}\Gamma, z \pm h \in \Gamma\} \Rightarrow$$

$$3\beta D^2\mathcal{A}(z)[h,h] \leq_K D^3\mathcal{A}(z)[h,h,h] \leq_K -3\beta D^2\mathcal{A}(z)[h,h]$$

(indeed, if the premise in (iii) is satisfied by $(z,h)$, it is also satisfied by $(z,-h)$).

**Definition 5.1.2** *Let* $\beta$ *be a nonnegative real; let* $E$, $E^+$ *be finite-dimensional linear spaces; let* $K$ *be a closed convex cone in* $E$*; let* $\Gamma$ *be a closed convex domain in* $E^+$*; and let* $\Pi$ *be a self-concordant barrier for* $\Gamma$*. A mapping* $\mathcal{A} : \operatorname{int}\Gamma \to E$ *is called* $(K,\beta)$*-compatible with the barrier* $\Pi$, *if*
(i) $\mathcal{A}$ *is* $C^3$*-smooth on* $\operatorname{int}\Gamma$*;*
(ii) $\mathcal{A}$ *is concave with respect to* $K$*;*
(iii) *For all* $z \in \operatorname{int}\Gamma$, $h \in E^+$, *we have*

$$D^3\mathcal{A}(z)[h,h,h] \leq_K -3\beta D^2\mathcal{A}(z)[h,h]\{D^2\Pi(z)[h,h]\}^{1/2}.$$

Similarly to above, (iii) is equivalent to

$$\{z \in \operatorname{int}\Gamma\} \Rightarrow$$

$$3\beta D^2\mathcal{A}(z)[h,h]\{D^2\Pi(z)[h,h]\}^{1/2} \leq_K D^3\mathcal{A}(z)[h,h,h] \leq_K$$
$$\leq_K -3\beta D^2\mathcal{A}(z)[h,h]\{D^2\Pi(z)[h,h]\}^{1/2}.$$

**Remark 5.1.1** *We already have dealt with the particular case of real-valued mappings compatible with self-concordant barriers, which were called "functions compatible with barriers" (Definition* 3.2.1*). A function* $f : \operatorname{int} G \to \mathbf{R}$, *which is* $\beta$*-compatible with a barrier* $F$ *for* $G$ *in the sense of Definition* 3.2.1, *is a mapping into* $\mathbf{R}$, *which is* $(\mathbf{R}_-, 3^{1/2}\beta)$*-compatible with* $F$ *in the sense of Definition* 5.1.2.

**Remark 5.1.2** *The above properties—compatibility with a domain and compatibility with a barrier for the domain—are closely related: A mapping* $\mathcal{A} : \operatorname{int}\Gamma \to E$, *which is* $(K,\beta)$*-compatible with* $\Gamma$ *is also* $(K,\beta)$*-compatible with any self-concordant barrier for* $\Gamma$*. Conversely, if* $\mathcal{A}$ *is* $(K,\beta)$*-compatible with a* $\theta$*-self-concordant barrier* $\Pi$ *for* $\Gamma$, *then* $\mathcal{A}$ *is* $(K,(3\theta+1)\beta)$*-compatible with* $\Gamma$.

First implication: Let $\mathcal{A}$ be $(K,\beta)$-compatible with $\Gamma$ and $\Pi$ be a self-concordant barrier for $\Gamma$ and let $z \in \operatorname{int}\Gamma$ and $h \in E^+$ be such that

$$D^2\Pi(z)[h,h] \leq 1.$$

Then the points $z \pm h$ belong to the unit Dikin ellipsoid of $\Pi$ centered at $z$ and therefore belong to $\Gamma$ (Proposition 2.3.2(i)). In view of Definition 5.1.1, item (iii), it follows that the inequality required in Definition 5.1.2, item (iii) is satisfied for all $h$ such that $D^2\Pi(z)[h,h] \le 1$ and hence for all $h$ (since both sides of the inequality are of the same homogeneity degree with respect to $h$). Thus, $\mathcal{A}$ satisfies Definition 5.1.2.

Second implication: Let $\mathcal{A}$ be $(K,\beta)$-compatible with a $\theta$-self-concordant barrier $\Pi$ for $G$ and let $z \in \operatorname{int}\Gamma$ and $h \in E^+$ be such that $z \pm h \in \Gamma$. From Proposition 2.3.2(iii), it follows that $\{D^2\Pi(z)[h,h]\}^{1/2} \le 3\theta + 1$, so that Definition 5.1.2, item (iii) implies the required inequality $D^3\mathcal{A}(z)[h,h,h] \le_K -3\beta(3\theta+1)D^2\mathcal{A}(z)[h,h]$. □

**Examples of mappings compatible with their domains.**

**Example 1.** An affine mapping is $(K,0)$-compatible with any closed convex domain in the space of argument for any cone $K$ in the image space.

**Example 2.** Let $\mathcal{A}(z) : E^+ \to E$ be a *quadratic* mapping, i.e., a mapping with the coordinates of the image represented as

$$y_j = \langle Q_j z, z\rangle + \langle b_j, z\rangle + c_j, \qquad j = 1, \dots, l \equiv \dim E$$

($\langle\cdot,\cdot\rangle$ is an inner product on $E^+$, $Q_j$ are linear operators on $E^+$, $b_j \in E^+$ and $c_j \in \mathbf{R}$). Assume that $\mathcal{A}$ is concave with respect to $K$. Then $\mathcal{A}$ evidently is $(K,0)$-compatible with $\Gamma = E^+$.

**Example 3.** A $(\mathbf{R}_+,\beta)$-compatible with $\mathbf{R}_+$ function is a usual concave and $C^3$-smooth real-valued function $f : (0,\infty) \to \mathbf{R}$ such that

$$| f'''(t) | \le -\frac{3\beta}{t} f''(t), \; t > 0$$

(evident). In particular, we have the following examples.

**Example 3.1.** $t^p$ is $(\mathbf{R}_+, (2-p)/3)$-compatible with $\mathbf{R}_+$, $0 < p \le 1$;

**Example 3.2.** $\ln t$ is $(\mathbf{R}_+, 2/3)$-compatible with $\mathbf{R}_+$.

**Example 4.** The mapping $\mathcal{A}(x) = -x^{-1} : \operatorname{int} S_m^+ \to S_m$ (see §5.1.2.A, Example 3.2) is $(S_m^+, 1)$-compatible with $S_m^+$.

Indeed, we already know that $\mathcal{A}(x) = -x^{-1}$ is $S_m^+$-concave. Furthermore, we have

$$-D^2\mathcal{A}(x)[h,h] = 2x^{-1}hx^{-1}hx^{-1} = 2x^{-1/2}(x^{-1/2}hx^{-1/2})^2x^{-1/2},$$

$$D^3\mathcal{A}(x)[h,h,h] = -6x^{-1}hx^{-1}hx^{-1}hx^{-1} = -6x^{-1/2}(x^{-1/2}hx^{-1/2})^3x^{-1/2},$$

so that

$$\Delta \equiv -3D^2\mathcal{A}(x)[h,h] - D^3\mathcal{A}(x)[h,h,h] = 6x^{-1/2}\{Q^2 - Q^3\}x^{-1/2},$$

where $Q = x^{-1/2}hx^{-1/2}$. If $x \pm h \in S_m^+$, then clearly $-I \le_{S_m^+} Q \le_{S_m^+} I$, $I$ being the unit matrix, whence $Q^3 \le_{S_m^+} Q^2$, and therefore $\Delta$ is positive-semidefinite.

The following simple lemma allows us to extend the list of examples.

**Lemma 5.1.3** (i) *If $\mathcal{A} : \operatorname{int}\Gamma \to E$ is $(K,\beta)$-compatible with $\Gamma$ and $K' \supset K$ is a closed convex cone in $E$, then $\mathcal{A}$ is $(K',\beta)$-compatible with $\Gamma$.*

(ii) Stability with respect to summation, multiplication by positive reals and to restriction. *If $\mathcal{A}_i : \operatorname{int}\Gamma_i \to E$ are $(K,\beta_i)$-compatible with closed convex domains $\Gamma_i \subseteq E^+$, $i = 1,2$ and if $\Gamma = \Gamma_1 \bigcap \Gamma_2$ possesses a nonempty interior, then a combination*

$$\mathcal{A}(x) = p\mathcal{A}_1(x) + q\mathcal{A}_2(x) : \operatorname{int}\Gamma \to \mathbf{R}$$

*of $\mathcal{A}_i$ with nonnegative coefficients is $(K, \max\{\beta_1,\beta_2\})$-compatible with $\Gamma$.*

(iii) Stability with respect to affine substitutions of argument. *If $\mathcal{A} : \operatorname{int}\Gamma \to E$ is $(K,\beta)$-compatible with a closed convex domain $\Gamma \subseteq E^+$ and $\mathcal{B}$ is an affine mapping into $E^+$ with the image intersecting $\operatorname{int}\Gamma$, then the mapping $\mathcal{A}\circ\mathcal{B}$ is $(K,\beta)$-compatible with the closed convex domain $\mathcal{B}^{-1}(\Gamma)$.*

(iv) Stability with respect to direct product. *If $\mathcal{A}_i : \operatorname{int}\Gamma_i \to E_i$ are $(K_i,\beta_i)$-compatible with $\Gamma_i$, $i = 1,2$, then the mapping $(x,y) \to (\mathcal{A}_1(x), \mathcal{A}_2(y)) : \operatorname{int}(\Gamma_1 \times \Gamma_2) \to E_1 \times E_2$ is $(K_1 \times K_2, \max\{\beta_1,\beta_2\})$-compatible with $\Gamma_1 \times \Gamma_2$.*

**Proof.** The proof is quite straightforward.

In fact, statement (iii) of the lemma can be significantly strengthened; it is done in §5.1.2.C.

Of course, statements similar to those of Lemma 5.1.3 hold for mappings compatible with barriers.

**Lemma 5.1.4** (i) *If $\mathcal{A} : \operatorname{int}\Gamma \to E$ is $(K,\beta)$-compatible with a self-concordant barrier $\Pi$ for $\Gamma$ and $K' \supset K$ is a closed convex cone in $E$, then $\mathcal{A}$ is $(K',\beta)$-compatible with $\Pi$.*

(ii) Stability with respect to summation, multiplication by positive reals and to restriction. *If $\mathcal{A}_i : \operatorname{int}\Gamma_i \to E$ are $(K,\beta_i)$-compatible with self-concordant barriers $\Pi_i$ for closed convex domains $\Gamma_i \subseteq E^+$, $i = 1,2$ and if $\Gamma = \Gamma_1 \bigcap \Gamma_2$ possesses a nonempty interior, then a combination*

$$\mathcal{A}(x) = p\mathcal{A}_1(x) + q\mathcal{A}_2(x) : \operatorname{int}\Gamma \to \mathbf{R}$$

*of $\mathcal{A}_i$ with nonnegative coefficients is $(K, \max\{\beta_1,\beta_2\})$-compatible with the self-concordant barrier $\Pi_1 + \Pi_2$ for $\Gamma$.*

(iii) Stability with respect to affine substitutions of argument. *If $\mathcal{A} : \operatorname{int}\Gamma \to E$ is $(K,\beta)$-compatible with a self-concordant barrier $\Pi$ for a closed convex domain $\Gamma \subseteq E^+$ and $\mathcal{B}$ is an affine mapping into $E^+$ with the image intersecting $\operatorname{int}\Gamma$, then the mapping $\mathcal{A}\circ\mathcal{B}$ is $(K,\beta)$-compatible with the self-concordant barrier $\Pi\circ\mathcal{B}$ for the closed convex domain $\mathcal{B}^{-1}(\Gamma)$.*

(iv) Stability with respect to direct product. *If $\mathcal{A}_i : \operatorname{int}\Gamma_i \to E_i$ are $(K_i,\beta_i)$-compatible with self-cocnordant barriers $\Pi_i$ for $\Gamma_i$, $i = 1,2$, then the mapping*

$$(x,y) \to (\mathcal{A}_1(x), \mathcal{A}_2(y)) : \operatorname{int}(\Gamma_1 \times \Gamma_2) \to E_1 \times E_2$$

*is $(K_1 \times K_2, \max\{\beta_1,\beta_2\})$-compatible with the barrier $\Pi_1(x) + \Pi_2(y)$ for $\Gamma_1 \times \Gamma_2$.*

**Proof.** The proof again is quite straightforward.

Now let us formulate and prove the central result on mappings compatible with (barriers for) their domains; as we see in §§5.3, 5.4, this statement plays the key role in constructing explicit self-concordant barriers.

**Proposition 5.1.7** *Let $G$ be a closed convex domain in $E$, let $F$ be a $\vartheta$-self-concordant barrier for $G$, let $\Gamma$ be a closed convex domain in $E^+$, let $\Pi$ be a $\nu$-self-concordant barrier for $\Gamma$ (in contrast to our usual practice, now we regard a function $\Pi \equiv \mathrm{const}$ as a 0-self-concordant barrier for the whole space, i.e., for $\Gamma = E^+$), and let $\mathcal{A} : \operatorname{int}\Gamma \to E$ be a mapping. Assume that $\mathcal{A}$ is $(\mathcal{R}(G), \beta)$-compatible with $\Pi$ or with $\Gamma$, where $\mathcal{R}(G)$ is the recessive cone of $G$, and that $\mathcal{A}(\operatorname{int}\Gamma)$ intersects $\operatorname{int} G$.*

*Then the set*

$$G^+ = \operatorname{cl}\{x \in \operatorname{int}\Gamma \mid \mathcal{A}(x) \in \operatorname{int} G\}$$

*is a closed convex domain in $E^+$, and the function*

$$\Psi(x) = \max^2\{\beta, 1\}\{F(\mathcal{A}(x)) + \Pi(x)\}$$

*is a $(\max^2\{\beta, 1\}(\vartheta + \nu))$-self-concordant barrier for $G^+$.*

*Comment.* We see that the simplest rule $F \to F \circ \mathcal{A}$ for updating barriers when passing from $G$ to the inverse image of $G$ under *affine* mappings $\mathcal{A}$ can be easily extended onto *nonlinear* mappings, which are $(\mathcal{R}(G), \beta)$-compatible with their domains. For the latter case, the updating rule is

$$F \mapsto \max^2\{1, \beta\}(F \circ \mathcal{A} + \Pi),$$

where $\Pi$ is a self-concordant barrier for the domain of $\mathcal{A}$. In the case of affine (or $\mathcal{R}(G)$-concave quadratic) $\mathcal{A}$, we can set $\beta = 0$, $\Pi \equiv 0$, which results in the initial rule $F \to F \circ \mathcal{A}$.

**Proof of proposition.** In view of Remark 5.1.2, it suffices to consider the case when $\mathcal{A}$ is $(\mathcal{R}(G), \beta)$-compatible with $\Pi$.

$1^0$. Since $\mathcal{A}$ is concave with respect to $\mathcal{R}(G)$ (see Definition 5.1.1, item (ii)), the set $G^+$ is a closed convex domain in view of Lemma 5.1.1.

To prove that $\Psi$ is a self-concordant barrier with the announced value of the parameter, let us compute the derivatives of $\Psi$ at a point $x \in \operatorname{int}\Gamma$ such that $y \equiv \mathcal{A}(x) \in \operatorname{int} G$ in a direction $h \in E^+$. Under the notation

$$\gamma = \max\{\beta, 1\},$$

$$u = D\mathcal{A}(x)[h], \quad v = D^2\mathcal{A}(x)[h,h], \quad w = D^3\mathcal{A}(x)[h,h,h],$$
$$r = \{D^2F(y)[u,u]\}^{1/2}, \qquad \rho = \{D^2\Pi(x)[h,h]\}^{1/2},$$

we have

(5.1.14) $$D\Psi(x)[h] = \gamma^2\{DF(y)[u] + D\Pi(x)[h]\};$$

(5.1.15) $D^2\Psi(x)[h,h] = \gamma^2\{DF(y)[v] + D^2F(y)[u,u] + D^2\Pi(x)[h,h]\}$;

(5.1.16)
$$D^3\Psi(x)[h,h,h] = \gamma^2\{DF(y)[w] + 3D^2F(y)[u,v] + D^3F(y)[u,u,u] + D^3\Pi(x)[h,h,h]\}.$$

Now, we have

$$-DF(y)[\Delta] \geq \{D^2F(y)[\Delta,\Delta]\}^{1/2}, \qquad \Delta \geq_{\mathcal{R}(G)} 0 \tag{5.1.17}$$

(Corollary 2.3.1), and

$$-v \geq_{\mathcal{R}(G)} 0 \tag{5.1.18}$$

(Lemma 5.1.2 combined with the concavity of $\mathcal{A}$ with respect to $\mathcal{R}(G)$). Furthermore, in view of Definition 5.1.2, item (iii), we have

$$3\beta\rho v \leq_{\mathcal{R}(G)} w \leq_{\mathcal{R}(G)} -3\beta\rho v. \tag{5.1.19}$$

Besides this, since $F$ and $\Pi$ are self-concordant barriers, we have

$$|\, DF(y)[u] \,| \leq \vartheta^{1/2} r, \qquad |\, D\Pi(x)[h] \,| \leq \nu^{1/2}\rho, \tag{5.1.20}$$

$$|\, D^3F(y)[u,u,u] \,| \leq 2r^3, \qquad |\, D^3\Pi(x)[h,h,h] \,| \leq 2\rho^3. \tag{5.1.21}$$

$2^0$. From (5.1.17), (5.1.18), it follows that

$$DF(y)[v] \geq \{D^2F(y)[v,v]\}^{1/2}, \tag{5.1.22}$$

so that the quantity

$$s = \{DF(y)[v]\}^{1/2}$$

is well defined. In view of Cauchy's inequality and (5.1.22), we have

$$|\, D^2F(y)[u,v] \,| \leq \{D^2F(y)[u,u]\}^{1/2}\{D^2F(y)[v,v]\}^{1/2} \leq r\, s^2. \tag{5.1.23}$$

Furthermore, the linear functional $DF(y)[\cdot]$ is nonpositive on the cone $\mathcal{R}(G)$ (see (5.1.17)), so that (5.1.19) implies that

$$|\, DF(y)[w] \,| \leq 3\beta\rho s^2. \tag{5.1.24}$$

In view of (5.1.20)–(5.1.24) we can conclude from (5.1.14)–(5.1.16) that

$$|\, D\Psi(x)[h] \,| \leq \gamma^2\{\vartheta^{1/2} r + \nu^{1/2}\rho\}, \tag{5.1.25}$$

$$D^2\Psi(x)[h,h] = \gamma^2\{s^2 + r^2 + \rho^2\}, \tag{5.1.26}$$

$$\begin{aligned} |\, D^3\Psi(x)[h,h,h] \,| &\leq 3\beta\gamma^2\rho s^2 + 3\gamma^2 r\, s^2 + 2\gamma^2 r^3 + 2\gamma^2\rho^3 \\ &\leq 2\gamma^3\{\tfrac{3}{2}\rho\, s^2 + r^3 + \rho^3 + \tfrac{3}{2} r\, s^2\} \end{aligned} \tag{5.1.27}$$

(the concluding inequality in (5.1.27) follows from $\gamma = \max\{\beta, 1\}$).

From (5.1.25) and (5.1.26), it follows that

$$
\begin{aligned}
| D\Psi(x)[h] | &\le \gamma^2\{\vartheta^{1/2}r + \nu^{1/2}\rho\} \le \gamma^2\{\vartheta+\nu\}^{1/2}\{r^2+\rho^2\}^{1/2} \\
&\le \gamma\{\vartheta+\nu\}^{1/2}\{D^2\Psi(x)[h,h]\}^{1/2}.
\end{aligned} \tag{5.1.28}
$$

Now let us verify that

$$
| D^3\Psi(x)[h,h,h] | \le 2\{D^2\Psi(h,h]\}^{3/2}. \tag{5.1.29}
$$

In view of (5.1.26), (5.1.27), to prove (5.1.29), it suffices to establish that

$$
\tfrac{3}{2}\rho\, s^2 + r^3 + \rho^3 + \tfrac{3}{2} r\, s^2 \le \{r^2+s^2+\rho^2\}^{3/2}
$$

for all nonnegative $r$, $s$, $\rho$, or, which is the same, to prove that

$$
\tfrac{3}{2}\rho\, s^2 + r^3 + \rho^3 + \tfrac{3}{2} r\, s^2 \le 1 \tag{5.1.30}
$$

for all nonnegative $r$, $s$, $\rho$ with $r^2+s^2+\rho^2 = 1$.

We have

$$
\begin{aligned}
\tfrac{3}{2}\rho s^2 + r^3 + \rho^3 + \tfrac{3}{2} r s^2 &= (r+\rho)(r^2+\rho^2 - r\rho + \tfrac{3}{2}s^2) \\
&= (r+\rho)(\tfrac{3}{2}(r^2+s^2+\rho^2) - \tfrac{1}{2}(r+\rho)^2) \\
&= \tfrac{1}{2}(r+\rho)(3-(r+\rho))^2 \le 1
\end{aligned}
$$

(we considered that the unconstrained maximum of $t(3-t^2)$ equals to 2), so that (5.1.30) and therefore (5.1.29) are proved.

$4^0$. Thus, $\Psi$ satisfies (5.1.28), (5.1.29) for all $x \in \operatorname{int} G^+$ and all $h \in E^+$; to complete the proof, it suffices to verify that, if $x_i \in \operatorname{int} G^+$ converge to a boundary point $x$ of $G^+$, then $\Psi(x_i) \to \infty$.

Let $y_i = \mathcal{A}(x_i)$. Assume first that $x \in \operatorname{int}\Gamma$. Then $y_i$ converge to certain $y \in G$ (since $\mathcal{A}$ is continuous on $\operatorname{int}\Gamma$), and $y \notin \operatorname{int} G$ (since otherwise $x$ would belong to $\operatorname{int} G^+$). Since $F$ is a barrier for $G$, we have $F(y_i) \to \infty, \Pi(x_i) \to \Pi(x)$, so that $\Psi(x_i) \to \infty$, as required.

Now let $x \in \partial\Gamma$. It suffices to lead to a contradiction the assumption that $\{\Psi(x_i)\}$ is bounded. Let it be the case. Since $x \in \partial\Gamma$, we have $\Pi(x_i) \to \infty$, so that the only possibility for $\{\Psi(x_i)\}$ to be bounded is $F(y_i) \to -\infty$. Let $E_F$ be the recessive subspace of $F$ and let $\pi$ be the canonical epimorphism of $E$ onto $P \equiv E/E_F$; as we know (see Proposition 2.3.2(iii)), $\bar G \equiv \pi(G)$ is a closed convex domain in $P$, $G = \pi^{-1}(\bar G)$ and $F(\cdot) = \bar F(\pi(\cdot))$ for certain $\vartheta$-self-concordant barrier $\bar F$ for $\bar G$. Let $z_i = \pi(y_i)$; from the correspondence between $(G,F)$ and $(\bar G, \bar F)$ and the fact that $F(y_i) \to -\infty$, it follows that $z_i \in \operatorname{int}\bar G$ and $\bar F(z_i) \to -\infty$. Since $\bar F$ is convex on $\operatorname{int}\bar G$, it is below bounded on any bounded subset of $\operatorname{int}\bar G$, so that the relation $\bar F(z_i) \to -\infty$ implies that $\| z_i \| \to \infty$. Without loss of generality, we can assume that the vectors $h_i = z_i / \| z_i \|$

converge to certain vector $h$; of course, $h \neq 0$ and $h \in \mathcal{R}(\bar{G})$. On the other hand, since $\mathcal{A}$ is $\mathcal{R}(G)$-concave, we have (see Lemma 5.1.2)

$$y_i = y_1 + D\mathcal{A}(x_1)[x_i - x_1] - r_i, \qquad r_i \in \mathcal{R}(G).$$

We conclude that

$$z_i = \pi(y_1) + \pi(D\mathcal{A}(x_1)[x_i - x_1]) - \pi(r_i), \ \pi(r_i) \in \mathcal{R}(\bar{G}).$$

Now, since $x_i$ converge to $x$, the sequence $\{\pi(y_1) + \pi(D\mathcal{A}(x_1)[x_i - x_1])\}$ is bounded, whence $h = \lim_{i\to\infty}(-\pi(r_i)/\parallel z_i \parallel)$, so that $-h \in \mathcal{R}(\bar{G})$.

Thus, $\pm h \in \mathcal{R}(\bar{G})$ and $h \neq 0$, so that $\bar{G}$ contains a line. It follows that $h \in E_{\bar{F}}$ (Corollary 2.3.1), and therefore $E_{\bar{F}} \neq \{0\}$, which is the desired contradiction (see the construction of $\bar{F}$). □

It is time now to enrich the list of examples of mappings compatible with their domains by one more example, which, combined with Proposition 5.1.7, leads to a number of important barriers (for applications, see §5.4).

**Examples of mappings compatible with their domains (continued).**

**Example 5.** Let $E$ be a linear space with a closed convex cone $K$, let $E'$ be a Euclidean space with the inner product $\langle\cdot,\cdot\rangle$, and let $T$ be a closed convex domain in certain linear space $E''$. Let $A(t)$, $t \in E''$ be a symmetric linear operator on $E'$ affinely depending on $t$,

$$A(t) = A_0 + A_1 t_1 + \cdots + A_k t_k,$$

where $t_i, i = 1, \dots, k \equiv \dim E''$ are the coordinates of $t$ in certain basis and $A_i$ are symmetric operators on $E'$.

Also, let $\mathcal{Q}(x', x'')$ be a bilinear symmetric mapping from $E'$ into $E$, i.e., a mapping with the coordinates of the image represented as $y_j = \langle Q_j x', x''\rangle$, $j = 1, \dots, l \equiv \dim E$, where the *coordinate operators* $Q_j$ are linear symmetric operators on $E'$.

Assume that

(A.i) $\mathcal{Q}$ is a convex with respect to $K$ mapping, i.e., $\mathcal{Q}(x, x) \in K$, $x \in E'$;

(A.ii) $A(t)$ is positive definite for $t \in \operatorname{int} T$;

(A.iii) for any $t \in E''$ the bilinear vector-valued function $\mathcal{Q}(A(t)x', x'')$ of $x', x''$ is symmetric in $x', x''$.

Note that (A.iii) is equivalent to the assumption that, for all $t$, the operator $A(t)$ commutates with all coordinate operators $Q_j$.

The above data generate the following quadratic-fractional mapping:

$$\mathcal{A}(y, x, t) = y - \mathcal{Q}(A^{-1}(t)x, x) : H \equiv E \times E' \times (\operatorname{int} T) \to E.$$

**Proposition 5.1.8** (i) *In the above situation, under assumptions* (A.i)–(A.iii) *the mapping* $\mathcal{A}$ *is* $(K, 1)$*-compatible with* $\Gamma \equiv E \times E' \times T$.

(ii) *In particular, if $G \subseteq E$ is a closed convex domain with $\mathcal{R}(G) \supset K$, $F$ is a $\vartheta$-self-concordant barrier for $G$, and $\Phi$ is a $\nu$-self-concordant barrier for $T$, then the function*

$$\Psi(y,x,t) = F(y - \mathcal{Q}(A^{-1}(t)x, x)) + \Phi(t)$$

*is a $(\vartheta + \nu)$-self-concordant barrier for the closed convex domain*

(5.1.31)
$$G^+ = \mathrm{cl}\,\{(y,x,t) \in E^+ \equiv E \times E' \times E'' \mid t \in \mathrm{int}\, T, y - \mathcal{Q}(A^{-1}(t)x, x) \in \mathrm{int}\, G\}.$$

**Proof.** $1^0$. (i): Of course, $\mathcal{A}$ satisfies Definition 5.1.1, item (i). Let us compute the derivatives of $\mathcal{A}$ at a point $X = (y,x,t) \in \mathrm{int}\,\Gamma$ in a direction $\Xi = (\eta, \xi, \tau) \in E^+$. In what follows, we fix coordinates in $E$, and subscript $i$ marks $i$th coordinate of a vector from $E$; $A'(\tau)$ denotes the derivative of the affine mapping $A(\cdot)$ in a direction $\tau$.

We have

$$D\mathcal{A}_i(X)[\Xi] = \eta_i - \left\langle Q_i A^{-1}(t)\{\xi - A'(\tau)A^{-1}(t)x\}, x\right\rangle - \left\langle Q_i A^{-1}(t)x, \xi\right\rangle;$$

(5.1.32)
$$D^2\mathcal{A}_i(X)[\Xi,\Xi] = 2\left\langle Q_i A^{-1}(t)A'(\tau)A^{-1}(t)\{\xi - A'(\tau)A^{-1}(t)x\}, x\right\rangle$$
$$-2\left\langle Q_i A^{-1}(t)\{\xi - A'(\tau)A^{-1}(t)x\}, \xi\right\rangle.$$

As it was already indicated (see the comment to (A.iii)), all values of $A(\cdot)$ commutate with $Q_i$, whence $A'(\tau)$ also commutates with $Q_i$. It follows that

$$\langle Q_i A^{-1}(t)A'(\tau)A^{-1}(t)\{A'(\tau)A^{-1}(t)x - \xi\}, x\rangle$$
$$= \langle Q_i A^{-1}(t)\{A'(\tau)A^{-1}(t)x - \xi\}, A'(\tau)A^{-1}(t)x\rangle,$$

so that (5.1.32) can be rewritten as

(5.1.33)
$$D^2\mathcal{A}_i(X)[\Xi,\Xi]$$
$$= -2\langle Q_i A^{-1}(t)\{\xi - A'(\tau)A^{-1}(t)x\}, \{\xi - A'(\tau)A^{-1}(t)x\}\rangle.$$

From the latter relation, it follows that

$$D^3\mathcal{A}_i(X)[\Xi,\Xi,\Xi]$$
$$= 2\langle Q_i A^{-1}(t)A'(\tau)A^{-1}(t)\{\xi - A'(\tau)A^{-1}(t)x\}, \{\xi - A'(\tau)A^{-1}(t)x\}\rangle$$
$$- 2\langle Q_i A^{-1}(t)A'(\tau)A^{-1}(t)\{A'(\tau)A^{-1}(t)x - \xi\}, \{\xi - A'(\tau)A^{-1}(t)x\}\rangle$$
$$- 2\langle Q_i A^{-1}(t)\{\xi - A'(\tau)A^{-1}(t)x\}, A'(\tau)A^{-1}(t)\{A'(\tau)A^{-1}(t)x - \xi\}\rangle,$$

and, again considering that $A(t)$ and $A'(\tau)$ commutate with $Q_i$, we obtain

(5.1.34)
$$\begin{aligned}&D^3\mathcal{A}_i(X)[\Xi,\Xi,\Xi]\\&= 6\left\langle Q_i A^{-1}(t)A'(\tau)A^{-1}(t)\{\xi - A'(\tau)A^{-1}(t)x\}, \{\xi - A'(\tau)A^{-1}(t)x\}\right\rangle.\end{aligned}$$

Thus, under notation

$$\zeta = \xi - A'(\tau)A^{-1}(t)x,$$

we have

$$D^2\mathcal{A}(X)[\Xi,\Xi] = -2\mathcal{Q}(A^{-1}(t)\zeta,\zeta); \tag{5.1.35}$$

$$D^3\mathcal{A}(X)[\Xi,\Xi,\Xi] = 6\mathcal{Q}(A^{-1}(t)A'(\tau)A^{-1}(t)\zeta,\zeta). \tag{5.1.36}$$

$2^0$. **Mini-Lemma 1.** *Let $B$ be a positive-semidefinite symmetric operator on $E'$ commutating with all coordinate operators $Q_i$ of $\mathcal{Q}$. Then $\mathcal{Q}(Bs,s) \in K$ for all $s \in E'$.*

**Proof.** Since $B$ commutates with all $Q_i$, the operator $B^{1/2}$ also possesses this property, whence

$$\langle Q_i Bs, s\rangle = \left\langle Q_i B^{1/2}s, B^{1/2}s\right\rangle,$$

so that

$$\mathcal{Q}(Bs,s) = \mathcal{Q}(B^{1/2}s, B^{1/2}s),$$

and the latter vector belongs to $K$ in view of assumption (A.i). □

Relation (5.1.35), in view of Mini-Lemma 1 and Lemma 5.1.2, implies that $\mathcal{A}$ satisfies Definition 5.1.1, item (ii), i.e., is concave with respect to $K$ on $\operatorname{int}\Gamma$.

$3^0$. It remains to verify item (iii) of Definition 5.1.1. To this end, assume that $X \pm \Xi \in \Gamma$.

$3^0$.1. Let us verify that the operators

$$B^+ = A^{-1}(t) - A^{-1}(t)A'(\tau)A^{-1}(t), \qquad B^- = A^{-1}(t) + A^{-1}(t)A'(\tau)A^{-1}(t)$$

(which clearly are symmetric) are positive-semidefinite.

Indeed, the points $X \pm \Xi$ belong to $\Gamma$, so that the points $t \pm \tau$ belong to $T$. The affine mapping $A(\cdot)$ takes positive-semidefinite values on the latter set; it follows that $A(t) \pm A'(\tau)$ is positive-semidefinite, so that the operators $A^{-1}(t)\{A(t) \pm A'(\tau)\}A^{-1}(t)$ also are positive-semidefinite, as claimed.

$3^0$.2. Since $B^+$ and $B^-$ are positive-semidefinite and commutate with $Q_i$ (because $A(t)$, $A'(\tau)$ do), Mini-Lemma 1 implies that

$$\mathcal{Q}(A^{-1}(t)\zeta,\zeta) - \mathcal{Q}(A^{-1}(t)A'(\tau)A^{-1}(t)\zeta,\zeta) \in K,$$

$$\mathcal{Q}(A^{-1}(t)\zeta,\zeta) + \mathcal{Q}(A^{-1}(t)A'(\tau)A^{-1}(t)\zeta,\zeta) \in K,$$

or (see (5.1.35), (5.1.36))

$$-\tfrac{1}{2}D^2\mathcal{A}(X)[\Xi,\Xi] \pm \tfrac{1}{6}D^3\mathcal{A}^3(X)[\Xi,\Xi,\Xi] \in K,$$

so that

$$X \pm \Xi \in \Gamma \;\Rightarrow\; 3D^2\mathcal{A}(X)[\Xi,\Xi] \le_K D^3\mathcal{A}(X)[\Xi,\Xi,\Xi] \le_K -3D^2\mathcal{A}(X)[\Xi,\Xi],$$

and $\mathcal{A}$ satisfies Definition 5.1.1, item (iii) with $\beta = 1$. Thus, $\mathcal{A}$ is $(K,1)$-compatible with $\Gamma$, and (i) is proved.

Part (ii) is an immediate corollary of (i) and Proposition 5.1.7 (we should apply Proposition 5.1.7 to $\Pi(y,x,t) \equiv \Phi(t)$ and consider that $\mathcal{A}(\Gamma)$ is the whole $E$ and therefore intersects $\operatorname{int} G$). □

### C. Compatibility of superpositions

A superposition of the usual concave real-valued functions on the axis, generally speaking, is not concave; to ensure the latter property, we should impose the monotonicity of the outer function. The resulting statement can be immediately extended onto vector-valued mappings: If $f : E^+ \to E$ is $K$-concave and $g : E \to E^-$ is $Q$-concave ($K$ and $Q$ are convex cones in $E$, $E^-$, respectively), and, in addition, $g$ is $(K,Q)$-monotone,

$$x' \ge_K x'' \;\Rightarrow\; g(x') \ge_Q g(x''),$$

then $g \circ f$ is $Q$-concave. Surprisingly, it turns out that the latter assumptions that ensure concavity of the superposition also ensure compatibility of $g\circ f$ with the domain of the superposition, provided that $f$ and $g$ possess this property.

**Proposition 5.1.9** *Let $\Gamma$ be a closed convex domain in $E^+$, let $G$ be a closed convex domain in $E$, and let $K$ be a closed convex cone in $E$. Also, let $Q$ be a closed convex cone in $E^-$. Let $\mathcal{F} : \operatorname{int}\Gamma \to E$ and $\mathcal{G} : \operatorname{int} G \to E^-$ be mappings. Assume that*

(i) *$\mathcal{F}$ is $(K,\alpha)$-compatible with $\Gamma$;*
(ii) *$\mathcal{G}$ is $(Q,\gamma)$-compatible with $G$;*
(iii) *$\mathcal{G}$ is $(K,Q)$-monotone on $\operatorname{int} G$:*

$$\{x', x'' \in \operatorname{int} G,\ x' \ge_K x''\} \;\Rightarrow\; \mathcal{G}(x') \ge_Q \mathcal{G}(x'');$$

(iv) *$K \subseteq \mathcal{R}(G)$.*

*Assume also that the set*

$$H = \{x \in \operatorname{int}\Gamma \mid \mathcal{F}(x) \in \operatorname{int} G\}$$

*is nonempty. Then the set $H$ is an open convex domain in $E^+$, the set*

$$G^+ = \mathrm{cl}\ H$$

*is a closed convex domain in $E^+$, and the superposition $\mathcal{S}(x) = \mathcal{G}(\mathcal{F}(x)) : \operatorname{int} G^+ \to E^-$ is $(Q,\beta)$-compatible with $G^+$, where*

$$\beta = \begin{cases} \max\{\alpha,\gamma\}, & \gamma \le \frac{1}{3}, \\ [\alpha+\gamma+\sqrt{(\alpha-\gamma)^2+(3\alpha+2)(3\gamma-1)}]/2, & \textit{otherwise.} \end{cases}$$

**Proof.** $1^0$. In view of Lemma 5.1.1 and (iv), the set $H$ is an open convex nonempty subset of $E^+$, so that $G^+$ is a closed convex domain in $E^+$ and $H = \operatorname{int} G^+$.

$2^0$. The mapping $\mathcal{S}$ clearly is $C^3$-smooth on $H = \operatorname{int} G^+$, so that $(\mathcal{S}, G^+)$ satisfies item (i) of Definition 5.1.1.

$3^0$. Let $x \in H$ and $h \in E^+$. Let us denote

$$f = \mathcal{F}(x), \quad f' = D\mathcal{F}(x)[h], \quad f'' = D^2\mathcal{F}(x)[h,h], \quad f''' = D^3\mathcal{F}(x)[h,h,h].$$

We have

$$D\mathcal{S}(x)[h] = D\mathcal{G}(f)[f'],$$

$$D^2\mathcal{S}(x)[h,h] = D\mathcal{G}(f)[f''] + D^2\mathcal{G}(f)[f',f'], \tag{5.1.37}$$

$$D^3\mathcal{S}(x)[h,h,h] = D\mathcal{G}(f)[f'''] + 3D^2\mathcal{G}(f)[f',f''] + D^3\mathcal{G}(x)[f',f',f']. \tag{5.1.38}$$

Since $\mathcal{F}$ is $K$-concave, we have

$$-f'' \in K \subseteq \mathcal{R}(G) \tag{5.1.39}$$

(Lemma 5.1.2 and (iv)). Furthermore, from $(K,Q)$-monotonicity of $\mathcal{G}$ it immediately follows that

$$D\mathcal{G}(u)[\Delta] \geq_Q 0 \quad \text{whenever } u \in \operatorname{int} G \quad \text{and } \Delta \geq_K 0. \tag{5.1.40}$$

Besides this, $\mathcal{G}$ is $Q$-concave, so that

$$D^2\mathcal{G}(u)[\Delta,\Delta] \leq_Q 0 \quad \text{whenever } u \in \operatorname{int} G \text{ and } \Delta \in E. \tag{5.1.41}$$

From (5.1.39) and (5.1.40), it follows that

$$D\mathcal{G}(f)[f''] \leq_Q 0, \tag{5.1.42}$$

while from (5.1.41) it follows that

$$D^2\mathcal{G}(f)[f'',f''] \leq_Q 0. \tag{5.1.43}$$

Relations (5.1.37), (5.1.42), and (5.1.43) imply that

$$D^2\mathcal{S}(x)[h,h] \leq_Q 0, \qquad x \in \operatorname{int} G^+, \quad h \in E^+,$$

so that $\mathcal{S}$ is $Q$-concave on $\operatorname{int} G^+$ (Lemma 5.1.2), i.e., $\mathcal{S}$ satisfies item (ii) of Definition 5.1.1.

$4^0$. It remains to verify that $\mathcal{S}$ satisfies item (iii) of Definition 5.1.1, i.e., that

$$D^3\mathcal{S}(x)[h,h,h] \leq_Q -3\beta D^2\mathcal{S}(x)[h,h], \tag{5.1.44}$$

provided that $x \pm h \in G^+$. Thus, we henceforth assume that the pair $(x, h)$ under consideration possesses the latter property. Note then that $x \pm h \in \Gamma$, whence, in view of (i),

$$f''' \leq_K -3\alpha f''. \tag{5.1.45}$$

This relation, combined with (5.1.40), implies that

$$D\mathcal{G}(f)[f'''] \leq_Q -3\alpha D\mathcal{G}(f)[f'']. \tag{5.1.46}$$

Thus, the first of three terms comprising $D^3\mathcal{S}(x)[h,h,h]$ is bounded from above (in the sense of $\leq_Q$) by constant times the vector $-D^2\mathcal{S}(x)[h,h]$ (see (5.1.38), (5.1.42), (5.1.43)).

To find a similar bound for the term $D^2\mathcal{G}(f)[f', f'']$, let us use "Cauchy's inequality": For any $\lambda > 0$, we have (see (5.1.41))

$$D^2\mathcal{G}(f)\left[\lambda f' + \frac{1}{\lambda}f'', \lambda f' + \frac{1}{\lambda}f''\right] \leq_Q 0,$$

whence

$$D^2\mathcal{G}(f)[f', f''] \leq_Q -\frac{1}{2}\left\{\lambda^2 D^2\mathcal{G}(f)[f', f'] + \frac{1}{\lambda^2}D^2\mathcal{G}(f)[f'', f'']\right\}. \tag{5.1.47}$$

The first of two terms of the right-hand side of (5.1.47) can be immediately bounded from above by $-D^2\mathcal{S}(x)[h,h]$, and our goal is to bound via the latter vector the vector $D^2\mathcal{G}(f)[f'', f'']$.

$4^0$.1. Let us verify that

$$f \pm f' \in G. \tag{5.1.48}$$

Indeed, $f$ is $K$-concave and $x \pm th \in \operatorname{int} G^+$, $0 \leq t < 1$, so that, in view of Lemma 5.1.2,

$$\operatorname{int} G \ni f(x+th) \leq_K f + tf', \qquad \operatorname{int} G \ni f(x-th) \leq_K f - tf', \quad 0 \leq t < 1;$$

since $K \subseteq \mathcal{R}(G)$, we conclude that $f \pm tf' \in \operatorname{int} G$, $0 \leq t < 1$, and (5.1.48) follows.

In view of (5.1.48) and (ii), we have

$$D^3\mathcal{G}(f)[f', f', f'] \leq_Q -3\gamma D^2\mathcal{G}(f)[f', f']. \tag{5.1.49}$$

$4^0$.2. Let us prove that

$$f + f' + \frac{f''}{3\alpha+2} \in G, \qquad f - f' + \frac{f''}{3\alpha+2} \in G. \tag{5.1.50}$$

$4^0$.3. Let

$$f(t) = \mathcal{F}(x+th),\ 0 \leq t < 1;$$

we claim that

$$f(t) \leq_K f + tf' + \frac{1}{3\alpha+1}\left\{t + \frac{(1-t)^{3\alpha+2}-1}{3\alpha+2}\right\} f'', \tag{5.1.51}$$

$$0 \leq t < 1;$$

note that (5.1.51), combined with the fact that $f(t) \in G$, $0 \leq t < 1$ (the latter is ensured by $x \in \operatorname{int} G^+$, $x+h \in G^+$) and (iv), implies the first inclusion in (5.1.50) (we should pass to limit as $t \to 1-0$).

To prove (5.1.51) is the same as proving that every function $\phi(t) = \langle \theta, f(t) \rangle$, $\theta$ being a linear functional on $E^+$ that is nonnegative on $K$, satisfies the inequality

$$\phi(t) \leq \phi(0) + t\phi'(0) + \frac{1}{3\alpha+1}\left\{t + \frac{(1-t)^{3\alpha+2}-1}{3\alpha+2}\right\}\phi''(0). \tag{5.1.52}$$

Note that, if $u = x + th$, $e = (1-t)h$, then $u \in \operatorname{int}\Gamma$ and $u \pm e \in \Gamma$, so that (i) implies $D^3\mathcal{F}(u)[e,e,e] \leq_K -3\alpha D^2\mathcal{F}(u)[e,e]$, whence

$$D^3\mathcal{F}(u)[h,h,h] \leq_K -\frac{3\alpha}{1-t}D^2\mathcal{F}(u)[h,h]$$

or

$$f'''(t) \leq_K -\frac{3\alpha}{1-t}f''(t)$$

or

$$\phi'''(t) \leq -\frac{3\alpha}{1-t}\phi''(t), \qquad 0 \leq t < 1. \tag{5.1.53}$$

Besides this, $K$-concavity of $\mathcal{F}$ implies that $f''(t) \leq_K 0$, whence

$$\phi''(t) \leq 0, \qquad 0 \leq t < 1. \tag{5.1.54}$$

Let us derive from (5.1.53), (5.1.54) that

$$\phi''(t) \leq (1-t)^{3\alpha}\phi''(0), \qquad 0 \leq t < 1. \tag{5.1.55}$$

Indeed, in the case of $\phi''(0) = 0$, relation (5.1.55) follows from (5.1.54). Now let $\phi''(0) \neq 0$, so that, in view of (5.1.54), $\phi''(0) < 0$. Let $\Delta$ be the largest of those half-intervals $[0,T)$, $T \leq 1$, on which $\phi'' < 0$. For $t \in \Delta$, we have (see (5.1.53))

$$-\frac{d}{dt}\ln(-\phi'''(t)) \leq \frac{3\alpha}{1-t},$$

whence

$$\frac{\phi''(0)}{\phi''(t)} \leq \frac{1}{(1-t)^{3\alpha}}, \qquad t \in \Delta,$$

or, taking into account (5.1.54), $\phi''(t) \leq (1-t)^{3\alpha}\phi''(0)$, $t \in \Delta$. From the latter inequality and the definition of $\Delta$, it follows that $\Delta = [0,1)$, and (5.1.55) is proved.

Two sequential integrations of (5.1.55) imply that

$$\phi(t) - \phi(0) - \phi'(0)t \leq \frac{\phi''(0)}{3\alpha+1}\left\{t + \frac{(1-t)^{3\alpha+2}-1}{3\alpha+2}\right\},$$

as required in (5.1.52). Thus, we have proved (5.1.51) and, consequently, the first inclusion in (5.1.50). In other words, we have proved that

$$\mathcal{F}(x) + D\mathcal{F}(x)[h] + \frac{1}{3\alpha+2}D^2\mathcal{F}(x)[h,h] \in G$$

whenever $x \in \operatorname{int} G^+$ and $h \in E^+$ are such that $x \pm h \in G^+$; if a pair $(x,h)$ satisfies the latter premise, than the pair $(x,-h)$ also does, so that

$$\mathcal{F}(x) - D\mathcal{F}(X)[h] + \frac{1}{3\alpha+2}D^2\mathcal{F}(x)[h,h] \in G,$$

and, in fact, we have proved the entire (5.1.50).

$4^0.4$. From (5.1.50), it follows that the middle of the segment with the endpoints involved into (5.1.50), i.e., the point $f + (3\alpha+2)^{-1}f''$ belongs to $G$,

(5.1.56) $$f - e \in G, \quad \text{where } e = -\frac{1}{3\alpha+2}f''.$$

On the other hand, from (5.1.39), it follows that

(5.1.57) $$f + te \in \operatorname{int} G, \qquad t \geq 0.$$

From (5.1.56) and (5.1.57), we conclude that, if

$$y(t) = f + te, \qquad e(t) = (1+t)e,$$

then $y(t) \in \operatorname{int} G$ and $y(t) \pm e(t) \in G$. In view of (ii), we have

$$D^3\mathcal{G}(y(t))[e(t),e(t),e(t)] \leq_Q -3\gamma D^2\mathcal{G}(y(t))[e(t),e(t)],$$

or, which is the same, that

$$D^3\mathcal{G}(y(t))[e,e,e] \leq_Q -\frac{3\gamma}{1+t}D^2\mathcal{G}(y(t))[e,e], \qquad t \geq 0.$$

In other words, if

$$g(t) = \mathcal{G}(y(t)),$$

then

(5.1.58) $$g'''(t) \leq_Q -3\gamma(1+t)^{-1}g''(t), \qquad t \geq 0;$$

besides this, from $Q$-concavity of $\mathcal{G}$, it follows that

(5.1.59) $$g''(t) \leq_Q 0, \qquad t \geq 0.$$

Following the line of argument from $4^0.3$, we can conclude from (5.1.58), (5.1.59) that

$$g''(t) \leq_Q \frac{1}{(1+t)^{3\gamma}} g''(0), \qquad t \geq 0. \tag{5.1.60}$$

Relation (5.1.60) implies that

$$g'(t) - g'(0) \leq_Q \int_0^t \frac{d\tau}{(1+\tau)^{3\gamma}} g''(0). \tag{5.1.61}$$

Since $g'(t) = D\mathcal{G}(f+te)[e]$, $e \in K$ (see (5.1.56) and (5.1.39)) and $\mathcal{G}$ is $(K,Q)$-monotone, we have $g'(t) \geq_Q 0$, so that (5.1.61) and (5.1.56) imply that

$$\begin{aligned} \frac{1}{3\alpha+2} D\mathcal{G}(f)[f''] = -g'(0) &\leq_Q \int_0^t \frac{d\tau}{(1+\tau)^{3\gamma}} g''(0) \\ &= \frac{1}{3\alpha+2} \int_0^t \frac{d\tau}{(1+\tau)^{3\gamma}} D^2\mathcal{G}(f)[f'',f''], \qquad t \geq 0, \end{aligned}$$

whence

$$-D^2\mathcal{G}(f)[f'',f''] \leq_Q -\frac{3\alpha+2}{\int_0^t \frac{d\tau}{(1+\tau)^{3\gamma}}} D\mathcal{G}(f)[f''], \qquad t > 0. \tag{5.1.62}$$

Note also that

$$-D\mathcal{G}(f)[f''] \geq_Q 0 \tag{5.1.63}$$

(see (5.1.39) and (5.1.40)).

$5^0$. From (5.1.62) and (5.1.47), it follows that, for any $\lambda > 0$ and any $t > 0$, we have

$$\begin{aligned} D^2\mathcal{G}(f)[f',f''] \leq_Q & -\tfrac{1}{2}\lambda^2 D^2\mathcal{G}(f)[f',f'] \\ & - \frac{3\alpha+2}{2\lambda^2 \int_0^t \frac{d\tau}{(1+\tau)^{3\gamma}}} D\mathcal{G}(f)[f'']. \end{aligned} \tag{5.1.64}$$

From (5.1.64), (5.1.46), (5.1.49), and (5.1.38), it follows that for any $\lambda$, $t > 0$, we have

$$\begin{aligned} D^3\mathcal{S}(x)[h,h,h] \leq_Q & \left(\frac{3}{2}\lambda^2 + 3\gamma\right)\left(-D^2\mathcal{G}(f)[f',f']\right) \\ & + \left(3\alpha + \frac{3(3\alpha+2)}{2\lambda^2 \int_0^t \frac{d\tau}{(1+\tau)^{3\gamma}}}\right)(-D\mathcal{G}(f)[f'']), \end{aligned} \tag{5.1.65}$$

whence, in view of (5.1.42), (5.1.43), and (5.1.37), for any $\lambda$, $t > 0$, we have

(5.1.66)
$$D^3 S(x)[h,h,h] \leq_Q \max\left\{ \frac{3\lambda^2}{2} + 3\gamma, 3\alpha + \frac{3(3\alpha+2)}{2\lambda^2 \int\limits_0^t \frac{d\tau}{(1+\tau)^{3\gamma}}} \right\} (-D^2 S(x)[h,h]).$$

Inequality (5.1.66) implies that $S$ satisfies item (iii) of Definition 5.1.1 with

$$\beta = \inf_{\lambda, t>0} \max\left\{ \frac{\lambda^2}{2} + \gamma, \alpha + \frac{3\alpha+2}{2\lambda^2 \int\limits_0^t \frac{d\tau}{(1+\tau)^{3\gamma}}} \right\}.$$

A straightforward computation of $\beta$ leads to

$$\beta = \begin{cases} \max\{\alpha, \gamma\}, & \gamma \leq \frac{1}{3}, \\ [\alpha + \gamma + \sqrt{(\alpha-\gamma)^2 + (3\alpha+2)(3\gamma-1)}]/2, & \text{otherwise}, \end{cases}$$

as claimed. □

## 5.2 Barrier calculus

The technique developed so far for constructing barriers can be essentially strengthened by the following simple observation. Assume that we wish to solve a standard problem

$$(f): \qquad f(x) \equiv \langle c, x \rangle \to \min \mid x \in G,$$

and, for this purpose, we are looking for a "computable" barrier for the corresponding closed convex domain $G \subset E$. It may happen that we cannot find such a barrier, but we can point out another closed convex domain $G^+ \subset E^+$, an affine mapping $\pi : E^+ \to E$ such that $\pi(G^+) = G$, and a computable barrier $F^+$ for $G^+$. In this case, it suffices to replace $(f)$ with the equivalent problem

$$(f^+): \qquad f^+(y) = \langle c, \pi(y) \rangle \to \min \mid y \in G^+$$

and to solve the latter problem by an interior-point method associated with $F^+$.

Of course, similar idea can be used to transform $(f)$ into a conic problem.

Although the above trick is very simple and traditional (it is common to simplify a problem by introducing appropriate additional variables), it turns out to be very important, and it is reasonable to study this in detail, as is done below.

### 5.2.1 Coverings and conic representations

The above considerations motivate the following definition.

**Definition 5.2.1** *Let $G$ be a closed convex domain in $E$, let $l \geq 0$ be an integer, and let $\vartheta > 1$.*

(i) *A $(\vartheta, l)$-covering for $G$ is, by definition, a collection*

$$\Gamma = (E^+, G^+, F^+, \pi)$$

*comprised of a linear space $E^+$, $\dim E^+ = \dim E + l$; a closed convex domain $G^+ \subset E^+$; a $\vartheta$-self-concordant barrier $F^+$ for $G^+$; an affine mapping $\pi : E^+ \to E$ such that $\pi(G^+) = G$. The quantities $l$ and $\vartheta$ are called the codimension and the parameter of $\Gamma$, respectively.*

(ii) *A $\vartheta$-conic representation of $G$ is, by definition, a collection*

$$\mathbf{K} = (E^+, K^+, F^+; H, \pi, \gamma),$$

*where $K^+$ is a closed convex cone with a nonempty interior in a finite-dimensional linear space $E^+$, $\pi$ is an affine mapping from finite-dimensional linear space $H$ into $E^+$ with $\pi(H) \bigcap \operatorname{int} K^+ \neq \emptyset$, $\gamma$ is an affine epimorphism of $H$ onto $E$ such that $G = \gamma(\pi^{-1}(K^+))$, and $F^+$ is a $\vartheta$-logarithmically homogeneous self-concordant barrier for $K^+$.*

Thus, regarding (i), finding a covering for $G$ is the same as representing $G$ as a projection of a closed convex domain $G^+$ in a "larger" space and pointing out a self-concordant barrier for the latter domain.

Regarding (ii), a conic representation of $G$ is a covering of a special type: The associated $G^+$ should be represented as an inverse image of a cone $K^+$ under an affine mapping, and we should know a logarithmically self-concordant barrier for $K^+$.

Usually, since convex domains occurring in convex optimization are defined by functional inequalities, we are especially interested in coverings for epigraphs of convex functions.

**Definition 5.2.2** *Let $G$ be a closed convex domain in $E$ and let $f$ be a lower semicontinuous convex function defined on $G$ and taking values in the extended real axis $\mathbf{R} \bigcup \{+\infty\}$ such that $f$ is finite on the interior of $G$. The pair $(G, f)$ is called a functional element (f.e.) on $E$. The epigraph of an f.e. $(G, f)$ is defined as the set*

$$\mathsf{G}(G, f) = \{(t, x) \in \mathbf{R} \times E \equiv E_+ \mid x \in G, \;\; t \geq f(x)\}$$

*(note that it is a closed convex domain in $E_+$). A $(\vartheta, l)$-covering (a $\vartheta$-conic representation) of the epigraph is called a $(\vartheta, l)$-covering (respectively, a $\vartheta$-conic representation) of the f.e. $(G, f)$.*

*An f.e. $(G, f)$ is called $(\vartheta, l)$-regular if it admits a $(\vartheta, l)$-covering.*

In what follows, we do not distinguish between a continuous convex function $\phi : E \to \mathbf{R}$ and the corresponding f.e. $(E, \phi)$, so we can speak about $(\vartheta, l)$-regular functions.

### 5.2.2 Calculus for coverings

The calculus that we develop is a straightforward extension of the results stated in the previous section.

**Proposition 5.2.1 (images)** *Let $\Gamma = (E^+, G^+, F^+, \pi)$ be a $(\vartheta, l)$-covering for a closed convex domain $G \subset E$ and let $\sigma : E \to E_1$ be an affine transformation such that the set $G_1 = \sigma(G)$ also is a closed convex domain. Then*

$$\Gamma_1 = (E^+, G^+, F^+, \sigma \circ \pi)$$

*is a $(\vartheta, l + (\dim E - \dim E_1))$-covering for $G_1$.*

*Under the same assumptions on $\sigma$, if $\mathbf{K} = (E^+, K^+, F^+; H, \pi, \gamma)$ is a $\vartheta$-conic representation of $G$, then*

$$\mathbf{K}^+ = (E^+, K^+, F^+; H, \pi, \sigma \circ \gamma)$$

*is a $\vartheta$-conic representation for $G_1$.*

**Proof.** This is an immediate consequence of the definitions. □

In what follows, Ker $\mathcal{A}$ for an *affine* mapping $\mathcal{A}$ denotes the kernel of the homogeneous part of the mapping.

**Proposition 5.2.2 (inverse images)** *Let $\Gamma = (E^+, G^+, F^+, \pi)$ be a $(\vartheta, l)$-covering for a closed convex domain $G \subset E$ and let $\sigma : E_1 \to E$ be an affine transformation such that $\sigma(E_1) \bigcap \operatorname{int} G \neq \emptyset$. Let $G_1 = \sigma^{-1}(G)$. Then $G_1$ is a closed convex domain in $E_1$, and $\Gamma$ induces a $(\vartheta, l)$-covering $\Gamma_1 = (E^\#, G^\#, F^\#, \pi^\#)$ for $G_1$. This covering is as follows:*

$$E^\# = \{(y, s) \mid y \in E^+,\ s \in E_1,\ \pi(y) = \sigma(s)\};$$

$$G^\# = \{(y, s) \in E^\# \mid y \in G^+\};$$

$$F^\#(y, s) = F(y), \qquad (y, s) \in \operatorname{int} G^\#;$$

$$\pi^\#(y, s) = s, \qquad (y, s) \in E^\#.$$

*Under the same assumptions on $\sigma$, a $\vartheta$-conic representation $\mathbf{K} = (E^+, K^+, F^+; H, \pi, \gamma)$ of $G$ induces a $\vartheta$-conic representation $\mathbf{K}^\# = (E^+, K^+, F^+; H^\#, \pi^\#, \gamma^\#)$ of $G_1$. This representation is as follows:*

$$H^\# = \{\gamma^{-1}(\mathrm{Im}\ \sigma)\} \times \mathrm{Ker}\ \sigma;$$

$$\pi^\#(u, s) = \pi(u), \qquad u \in \gamma^{-1}(\mathrm{Im}\ \sigma), \quad s \in \mathrm{Ker}\ \sigma;$$

$$\gamma^\#(u, s) = \tau(\gamma(u)) + s, \qquad u \in \gamma^{-1}(\mathrm{Im}\ \sigma), \quad s \in \mathrm{Ker}\ \sigma,$$

*where $\tau$ is a "partial inverse" to $\sigma$, i.e., an affine mapping from* Im $\sigma$ *into $E_1$ such that $\sigma(\tau(v)) \equiv v$.*

**Remark 5.2.1** *Rigorously speaking, the relation for, say, $E^{\#}$ defines not a linear, but an affine space, and, to obtain a linear space, we should somehow choose the origin. A similar situation occurs in some of the below statements.*

**Proof of Proposition 5.2.2.** $G^{\#}$ evidently is closed and convex. Since $G^+$ is a closed convex domain in $E^+$ and $\mathrm{Im}\,\sigma$ intersects the set $\mathrm{int}\,\pi(G^+) = \pi(\mathrm{int}\,G^+)$, there exist $y_0 \in \mathrm{int}\,G^+$ and $s_0 \in E_1$ such that $\pi(y_0) = \sigma(s_0)$. The point $(y_0, s_0)$ evidently is an interior point of $G^{\#}$, so that $G^{\#}$ is a closed convex domain in $E^{\#}$.

Let us verify that $F^{\#}$ is a $\vartheta$-self-concordant barrier for $G^{\#}$. $G^{\#}$ is the inverse image of $G^+$ under the mapping $\tau(y,s) = y : E^{\#} \to E^+$, and $\tau(E^{\#}) \ni y_0$, where $y_0$ is the above point from $\mathrm{int}\,G^+$. Thus, $\tau(E^{\#})$ intersects $\mathrm{int}\,G^+$, so that $F^+$ possesses the desired property in view of Proposition 5.1.1.

Now let us prove that $\pi^{\#}(G^{\#}) = G_1$. If $s \in G_1$, then $\sigma(s) \in G$ (the definition of $G_1$), and therefore $\sigma(s) = \pi(y)$ for some $y \in G^+$ (since $\Gamma$ is a covering for $G$). From the definitions of $G^{\#}$ and $\pi^{\#}$, it follows that $(y,s) \in G^{\#}$ and $\pi^{\#}(y,s) = s$. Thus, $\pi^{\#}(G^{\#}) \supset G_1$. Conversely, if $(y,s) \in G^{\#}$, then, by virtue of definitions of $\pi^{\#}$ and $G^{\#}$, we have $s = \pi^{\#}(y,s)$, $\sigma(s) = \pi(y) \in \pi(G^+) = G$, so that $s \in G_1 = \sigma^{-1}(G)$. Thus, $\pi^{\#}(G^{\#}) \subset G_1$, so that $\pi^{\#}(G^{\#}) = G_1$.

Now, for a point $(y,s) \in E^{\#}$ the relation $\pi^{\#}(y,s) = 0 \in E_1$ holds if and only if $s = 0$ and $\pi(y) = \sigma(0)$, so that the dimension of the inverse image $(\pi^{\#})^{-1}(s)$ of a point $s \in E_1$ coincides with the dimension of the inverse image $\pi^{-1}(x)$ of a point $x \in E$, whence the codimension of $\Gamma^{\#}$ equals to $l$.

The statement concerning conic representations can be proved by similar reasoning. $\square$

**Corollary 5.2.1** *Let $(G,f)$ be a functional element on $E$ and let $\mathcal{A} : E' \to E$ be an affine mapping such that $\mathcal{A}(E')$ intersects $\mathrm{int}\,G$. Then a $(\vartheta,l)$-covering for $(G,f)$ induces a $(\vartheta,l)$-covering for the functional element $(\mathcal{A}^{-1}(G), f\circ\mathcal{A})$, and a $\vartheta$-conic representation of $(G,f)$ induces a $\vartheta$-conic representation of the f.e. $(\mathcal{A}^{-1}(G), f\circ\mathcal{A})$.*

Indeed, $\mathsf{G}(\mathcal{A}^{-1}(G), f\circ\mathcal{A})$ is the inverse image of $\mathsf{G}(G,f)$ under the affine mapping $(t,u) \to (t,\mathcal{A}(u))$, and the image of this mapping intersects $\mathrm{int}\,\mathsf{G}(G,f)$.

**Proposition 5.2.3 (direct products)** *Let $\Gamma_i = (E_i^+, G_i^+, F_i^+, \pi_i)$ be $(\vartheta_i, l_i)$-coverings for closed convex domains $G_i$ belonging to the spaces $E_i$, $1 \le i \le m$ and let $G = G_1 \times \cdots \times G_m \subset E = E_1 \times \cdots \times E_m$. Then the coverings $\Gamma_i$ induce a $(\sum_{i=1}^m \vartheta_i, \sum_{i=1}^m l_i)$-covering $\Gamma = (E^+, G^+, F^+, \pi)$ for $G$. This covering is as follows:*

$$E^+ = E_1^+ \times \cdots \times E_m^+,$$

$$G^+ = G_1^+ \times \cdots \times G_m^+,$$

$$F^+(x_1,\ldots,x_m) = F_1^+(x_1) + \cdots + F_m^+(x_m), \qquad x_i \in \mathrm{int}\,G_i^+,$$

$$\pi(x_1, \ldots, x_m) = (\pi_1(x_1), \ldots, \pi_m(x_m)).$$

*$\vartheta_i$-conic representations $\mathbf{K}_i = (E_i^+, K_i^+, F_i^+; H_i, \pi_i, \gamma_i)$ of the above $G_i$ induce a $(\sum_{i=1}^m \vartheta_i)$-conic representation $\mathbf{K} = (E^+, K^+, F^+; H, \pi, \gamma)$ of $G$, which is as follows:*

$$E^+ = E_1^+ \times \cdots \times E_m^+,$$

$$K^+ = K_1^+ \times \cdots \times K_m^+,$$

$$F^+(x_1, \ldots, x_m) = F_1^+(x_1) + \cdots + F_m^+(x_m), \qquad x_i \in \operatorname{int} K_i^+,$$

$$H^+ = H_1 \times \cdots \times H_m,$$

$$\pi(x_1, \ldots, x_m) = (\pi_1(x_1), \ldots, \pi_m(x_m)),$$

$$\gamma(x_1, \ldots, x_m) = (\gamma_1(x_1), \ldots, \gamma_m(x_m)).$$

**Proof.** This is an immediate consequence of definitions and Proposition 5.1.2.
□

**Proposition 5.2.4 (intersections)** *Let $\Gamma_i = (E_i^+, G_i^+, F_i^+, \pi_i)$ be $(\vartheta_i, l_i)$-coverings for closed convex domains $G_i \subset E$, $1 \le i \le m$ and let $G \equiv \bigcap_{i=1}^m G_i$ possess a nonempty interior. Then the coverings $\Gamma_i$ induce a $(\sum_{i=1}^m \vartheta_i, \sum_{i=1}^m l_i)$-covering $\Gamma = (G^+, E^+, F^+, \pi)$ for $G$. The covering is as follows:*

$$E^+ = \{(x_1, \ldots, x_m) \in E^\# \equiv E_1 \times \cdots \times E_m \mid \pi_1(x_1) = \pi_2(x_2) = \cdots = \pi_m(x_m)\},$$

$$G^+ = \{(x_1, \ldots, x_m) \in E^+ \mid x_i \in G_i^+,\ 1 \le i \le m\},$$

$$F^+(x_1, \ldots, x_m) = F_1^+(x_1) + \cdots + F_m^+(x_m), \qquad (x_1, \ldots, x_m) \in G^+,$$

$$\pi(x_1, \ldots, x_m) = \pi_1(x_1)\ (= \pi_2(x_2) = \cdots = \pi_m(x_m)), \qquad (x_1, \ldots, x_m) \in E^+.$$

*Under the same assumptions about $\{G_i\}$, $\vartheta_i$-conic representations*

$$\mathbf{K}_i = (E_i^+, K_i^+, F_i^+; H_i, \pi_i, \gamma_i)$$

*of $G_i$ induce a $(\sum_{i=1}^m \vartheta_i)$-conic representation*

$$\mathbf{K} = (E^+, K^+, F^+; H, \pi, \gamma)$$

*of $G$, which is as follows:*

$$E^+ = E_1 \times \cdots \times E_m,$$

$$K^+ = K_1^+ \times \cdots \times K_m^+,$$

$$F^+(x_1, \ldots, x_m) = F_1^+(x_1) + \cdots + F_m^+(x_m), \qquad x_i \in \operatorname{int} K_i^+,$$

$$H = \{(u_1, \ldots, u_m) \in H_1 \times \cdots \times H_m \mid \gamma_1(u_1) = \cdots = \gamma_m(u_m)\},$$

$$\pi(u_1, \ldots, u_m) = (\pi_1(u_1), \ldots, \pi_m(u_m)),$$

$$\gamma(u_1, \ldots, u_m) = \gamma_1(u_1)\ (= \gamma_2(u_2) = \cdots = \gamma_m(u_m)), \qquad (u_1, \ldots, u_m) \in H.$$

**Proof.** $G^+$ clearly is a closed convex domain in $E^+$. Indeed, each point $x \in \operatorname{int} G$ can be represented as $\pi_i(x_i)$ for some $x_i \in \operatorname{int} G_i^+$, and $(x_1, \ldots, x_m)$ evidently is an interior point of $G^+$, so that this set has a nonempty interior; closedness and convexity of $G^+$ are evident. The same reasoning implies that, if $\tau$ is the natural embedding of $E^+$ into $E^\#$, then $G^+$ is the inverse $\tau$-image of $G_1^+ \times \cdots \times G_m^+$ and that $\tau(E^+)$ intersects the interior of the latter set; in view of Propositions 5.1.2 and 5.1.1, these observations imply that $F^+$ is a $(\sum_{i=1}^m \vartheta_i)$-self-concordant barrier for $G^+$.

Relations $\pi(G^+) = G$ and $\sum_{i=1}^m l_i = \dim E^+ - \dim E$ are evident. The statement about conic representation of $G$ is quite straightforward. □

**Corollary 5.2.2** *Let $(G_i, f_i)$ be functional elements on $E$, $1 \le i \le m$, such that*

$$\operatorname{int}\left(\bigcap_{i=1}^m G_i\right) \neq \emptyset.$$

*Then a collection of $(\vartheta_i, l_i)$-coverings ($\vartheta_i$-conic representations) of functional elements $(G_i, f_i)$ induces a $(\sum_{i=1}^m \vartheta_i, \sum_{i=1}^m l_i)$-covering (a $(\sum_{i=1}^m \vartheta_i)$-conic representation, respectively) of the functional element*

$$\left(\bigcap_{i=1}^m G_i, \max_{1\le i\le m} f_i\right).$$

**Proof.** Indeed, $\mathsf{G}(\bigcap_{i=1}^m G_i, \max_{1\le i\le m} f_i) = \bigcap_{i=1}^m \mathsf{G}(G_i, f_i)$. □

The following simple statement is evident.

**Proposition 5.2.5** *Let $(G, f)$ be a functional element on $E$ and let $\Gamma = (E^+, G^+, F^+, \pi)$ be a $(\vartheta, l)$-covering for this f.e. Let $\alpha > 0$ and let $\phi(\cdot)$ be an affine function on $E$. Then $\Gamma$ induces a $(\vartheta, l)$-covering $\Gamma^\# = (E^+, G^+, F^+, \pi')$ for the f.e. $(G, f^+(\cdot) = \alpha f(\cdot) + \phi(\cdot))$. The corresponding mapping $\pi'$ is defined as follows: Let $\pi(x) = (\tau(x), \sigma(x))$, $\tau(x) \in \mathbf{R}$, $\sigma(x) \in E$. Then*

$$\pi'(x) = (\alpha\tau(x) + \phi(\sigma(x)), \sigma(x)).$$

*A $\vartheta$-conic representation $\mathbf{K} = (E^+, K^+, F^+; H, \pi, \gamma)$ of $(G, f)$ induces a $\vartheta$-conic representation $\mathbf{K}^\# = (E^+, K^+, F^+; H, \pi, \gamma')$ of the f.e. $(G, f^+)$, where $\gamma'$ is defined as $\sigma \circ \tau$, $\sigma(t, x) = (\alpha t + \phi(x), x) : \mathbf{R} \times E \to \mathbf{R} \times E$.*

The following simple statement is rather useful.

**Proposition 5.2.6** *Let $(G, f)$ be a functional element on $E$ and let*

$$\Gamma = (E^+, G^+, F^+, \pi)$$

*be a $(\vartheta, l)$-covering for this f.e. Assume that the set $G_0' = \{u \in G \mid f(u) < 0\}$ is nonempty and let $G_0 = \{u \in G \mid f(u) \le 0\}$. Then $\Gamma$ induces a $(\vartheta + 1, l + 1)$-covering $\Gamma^\# = (E^+, G^\#, F^\#, \pi')$ for the set $G_0$. This covering is as follows. Let $\pi(x) = (\tau(x), \sigma(x))$, $\tau(x) \in \mathbf{R}$, $\sigma(x) \in E$. Then*

$$G^\# = \{x \in G^+ \mid \tau(x) \le 0\},$$

$$F^{\#}(x) = F^{+}(x) - \ln(-\tau(x)), \qquad \pi'(x) = \sigma(x).$$

*Under the same assumptions about $G$ and $f$, a $\vartheta$-conic representation*

$$\mathbf{K} = (E^{+}, K^{+}, F^{+}; H, \pi, \gamma)$$

*of the f.e. $(G, f)$ induces a $(\vartheta + 1)$-conic representation*

$$\mathbf{K}^{\#} = (E^{\#}, K^{\#}, F^{\#}; H^{\#}, \pi^{\#}, \gamma^{\#})$$

*of the set $G_0$, defined as follows. Since $\gamma : H \to \mathbf{R} \times E$, we can write*

$$\gamma(u) = (\tau(u), \sigma(u)), \qquad \tau(u) \in \mathbf{R}, \quad \sigma(u) \in E.$$

*We set*

$$E^{\#} = \mathbf{R} \times E^{+},$$

$$K^{\#} = \{t \geq 0\} \times K^{+},$$

$$F^{\#}(t, x) = F(x) - \ln t, \qquad x \in \operatorname{int} K^{+}, \quad t > 0,$$

$$H^{\#} = H,$$

$$\pi^{\#}(u) = (-\tau(u), \pi(u)), \qquad u \in H,$$

$$\gamma^{\#} = \sigma.$$

**Proof.** Since $f$ is lower semicontinuous and convex and $G_0' \neq \emptyset$, $G_0$ is a closed convex domain in $E$, and $G_0'$ intersects $\operatorname{int} G$. If $u \in G_0'$, then we can choose $t \in (f(u), 0)$, so that the point $(t, u)$ belongs to $\operatorname{int} \mathsf{G}(G, f)$, and therefore there exists $x \in \operatorname{int} G^{+}$ such that $\pi(x) = (t, u)$. Since $t < 0$, $x$ is an interior point of $G^{\#}$, so that $G^{\#}$ is a closed convex domain in $E^{+}$. Furthermore, $G^{\#}$ is the intersection of $G^{+}$ and the half-space $\Pi = \{x \in E^{+} \mid \tau(x) \leq 0\}$. The function $-\ln(-\tau(x))$ is a 1-self-concordant barrier for $\Pi$ (see §2.3, Example 2), and, since $\operatorname{int} G^{\#} \neq \emptyset$, $F^{\#}$ is a $(\vartheta{+}1)$-self-concordant barrier for $G^{\#}$ (see Proposition 5.1.3). The relation $\pi'(G^{\#}) = G_0$ and the fact that the codimension of $\Gamma^{\#}$ is $(l + 1)$ are evident.

The statement about conic representation is quite straightforward. □

Note that the above transformations of barriers are "simple" in the sense that they are straightforward and involve "rational" linear algebra techniques only. The progress with respect to the results of §5.1 is in Proposition 5.2.1: A "computable" covering for a domain immediately implies similar coverings for its images under affine transformations, which is not the case for barriers (see Proposition 5.1.5).

### 5.2.3 Legendre transformation

Let $(G, f)$ be a functional element on $E$ and let $E^*$ be the space conjugate to $E$. Consider the Legendre transformation of $f$, i.e., the function

$$f^*(\xi) \equiv \sup_{x \in \mathrm{Dom}\{f\}} \{\langle \xi, x\rangle - f(x)\} : E^* \to \mathbf{R} \bigcup \{+\infty\},$$

where $\mathrm{Dom}\{f\}$ denotes the subset of $G$ formed by the points at which $f$ is finite. Clearly, $f^*$ is a convex lower semicontinuous function taking values in the extended axis. The domain $\mathrm{Dom}\{f^*\}$ of this function (the set of all $\xi$ such that $f^*(\xi) < \infty\}$ is convex and nonempty (it clearly contains the set of all subgradients of $f$ at the interior points of $G$), and the set $G^* = \mathrm{cl}\ \mathrm{Dom}\{f^*\}$ is closed, convex, and nonempty. Thus, we have obtained the functional element $(G^*, f^*)$ on the affine span $E^{\#}$ of $\mathrm{Dom}\{f^*\}$; this element is called the Legendre transformation of $(G, f)$.

Our aim now is to demonstrate how a $\vartheta$-self-concordant barrier for the epigraph of $(G, f)$ can be transformed into a self-concordant barrier for the epigraph of $(G^*, f^*)$.

To this purpose, consider the conic hull of the epigraph $\mathsf{G}(G, f)$ of the element $(G, f)$, i.e., the cone

$$K = \mathrm{cl} \left\{ (s, t, x) \in E^+ \equiv \mathbf{R} \times \mathbf{R} \times E \mid s > 0,\ \frac{x}{s} \in \mathrm{Dom}\{f\},\ t \geq sf\left(\frac{x}{s}\right) \right\}.$$

This is a closed convex cone in $E^+$ with a nonempty interior, and the intersection of the cone with the hyperplane $L = \{s = 1\}$ is exactly $\mathsf{G}(G, f)$.

Now let us look at the dual cone $K^*$. A functional $(\sigma, \tau, \xi) \in (E^+)^* = \mathbf{R} \times \mathbf{R} \times E^*$ $(\langle (\sigma, \tau, \xi), (s, t, x)\rangle = \sigma s + \tau t + \langle \xi, x\rangle)$ is nonnegative on $K$ if and only if the implication

$$\left\{ s > 0,\ \frac{x}{s} \in \mathrm{Dom}\{f\},\ t \geq sf\left(\frac{x}{s}\right) \right\} \Rightarrow \{\sigma s + \tau t + \langle \xi, x\rangle \geq 0\}$$

holds or, similarly, if and only if the implication

$$\{x \in \mathrm{Dom}\{f\},\ t \geq f(x)\} \Rightarrow \{\sigma + \tau t + \langle \xi, x\rangle \geq 0\}$$

is true. Thus,

$$K^* = \{(\sigma, \tau, \xi) \mid \sigma + \tau t + \langle \xi, x\rangle \geq 0\ \ \forall (t, x) \in \mathsf{G}(G, f)\}.$$

Consider the intersection $H$ of $K^*$ and the hyperplane $\{\tau = 1\}$. This intersection is, in fact, the set $\{(\sigma, \xi) \in \mathbf{R} \times E^* \mid \sigma \geq \sup\{-t - \langle \xi, x\rangle \mid (t, x) \in \mathsf{G}(G, f)\}\}$, and, since

$$\begin{aligned} &\sup\{-t - \langle \xi, x\rangle \mid (t, x) \in \mathsf{G}(G, f)\} \\ &= \sup\{-f(x) - \langle \xi, x\rangle \mid x \in \mathrm{Dom}\{f\}\} = f^*(-\xi), \end{aligned}$$

we conclude that $H = \{(\sigma,\xi) \mid \sigma \geq f^*(-\xi)\}$, so that $H$ is the image of $\mathsf{G}(G^*, f^*)$ under a linear authomorphism.

Thus, the epigraph of the Legendre transformation of $(G, f)$ is, up to a simple linear authomorphism, the intersection of the cone dual to $K$ and an appropriate affine hyperplane not containing the origin; such a hyperplane, of course, intersects the relative interior of $K^*$. Therefore, to point out a self-concordant barrier for $\mathsf{G}(G^*, f^*)$, it suffices to find a logarithmically homogeneous self-concordant barrier for $K^*$ or a similar barrier for the cone $(-K^*)$ anti-dual to $K$. In turn, we can take as the latter barrier the Legendre transformation of a logarithmically homogeneous self-concordant barrier for $K$ (Theorem 2.4.4). We know how to update a $\vartheta$-self-concordant barrier for the epigraph of $(G, f)$ into an $O(1)\vartheta$-logarithmically homogeneous self-concordant barrier for the conic hull $K$ of this epigraph (Proposition 5.1.4). Thus, to update a $\vartheta$-self-concordant barrier $F$ for the epigraph of a given functional element $(G, f)$ into an $O(1)\vartheta$-self-concordant barrier for the epigraph of$(G^*, f^*)$, we are required to compute the Legendre transformation of the (slightly modified) function $F$.

### 5.2.4 Superposition theorem

The theorem that we establish is the main result of our barrier calculus.

Let $(G_i, \phi_i)$ be functional elements on $E$, $1 \leq i \leq m$ and let $(G, \phi)$ be a functional element on $\mathbf{R}^m$.

Consider the following pair of sets of assumptions. The first set is shown below:

(I.i) $\phi_i$ are finite and continuous on $G_i$, $1 \leq i \leq m$ and $\phi$ is finite and continuous on $G$;

(I.ii) The set $U = \bigcap_{i=1}^m G_i$ has a nonempty interior $U'$, and the image $Q$ of $U$ under the mapping $x \to (\phi_1(x), \ldots, \phi_m(x))$ is contained in $G$.

(I.iii) $G \supset Q + \mathbf{R}^n_+$, and $\phi$ is monotone on the latter set,

$$(\forall u \in Q,\ y \in \mathbf{R}^n_+) : \phi(u + y) \geq \phi(u).$$

The second set of assumptions is the following:

(II.i) If a sequence $\{t_i \in G\}$ is such that $\| t_i \|_2 \to \infty$ and $t_i \geq t$ for some $t \in \mathbf{R}^m$ and all $i$, then $\phi(t_i) \to \infty$;

(II.ii) the set $U = \bigcap_{i=1}^m G_i$ has a nonempty interior $U'$, and, for every $x \in U$ such that $\phi_i(x) < \infty$, $i = 1, \ldots, m$, the vector $(\phi_1(x), \ldots, \phi_m(x))$ is contained in $G$, while the image of $U'$ under the above mapping is contained in $\operatorname{int} G$;

(II.iii) This is the same as (I.iii).

We call the collection $\{(G_1, \phi_1), \ldots, (G_m, \phi_m), (G, \phi)\}$ *regular superposition data,* if it satisfies either (I.i)–(I.iii), or (II.i)–(II.iii).

Under assumptions (I.i)–(I.iii), the function

$$f(x) = \phi(\phi_1(x), \ldots, \phi_m(x))$$

clearly is well defined, convex, and continuous on $U$.

Under assumptions (II.i)–(II.iii), we also can define the above superposition, namely, as follows: If $x \in U$ is such that $\phi_1(x), \ldots, \phi_m(x)$ are finite, then

$$f(x) = \phi(\phi_1(x), \ldots, \phi_m(x));$$

otherwise, $f$ equals to $+\infty$.

**Lemma 5.2.1** *Under assumptions* (II.i)–(II.iii), *the pair* $(U, f)$ *is a functional element.*

**Proof.** We should verify that (a) $U$ is a closed convex domain and $f : U \to \mathbf{R}\bigcup\{+\infty\}$ is convex; (b) $f$ is finite on $\operatorname{int} U$; (c) $f$ is lower semicontinuous on $U$.

(a) The fact that $U$ is closed convex domain is evident. To prove that $f$ is convex, it suffices to verify that, if $x, y \in U$ and $w = \alpha x + (1-\alpha)y$ for certain $\alpha \in [0,1]$, then $f(w) \le \alpha f(x) + (1-\alpha)f(y)$ $(0 \cdot (+\infty) = +\infty)$. In the case when some of the functions $\phi_i$ are infinite at $x$ or at $y$, the above inequality is evident, since then either $f(x) = +\infty$, or $f(y) = +\infty$ (the definition of $f$). If all the functions $\phi_i$ are finite at $x$ and at $y$, then the vector $\Phi(u) = (\phi_1(u), \ldots, \phi_m(u))$ is well defined at $u = x$ and at $u = y$ and therefore is well defined at $w$, and $\Phi(w) \le \alpha\Phi(x) + (1-\alpha)\Phi(y)$ (recall that $\phi_i$ are convex). In particular, all three points $\Phi(x), \Phi(y), \Phi(w)$ belong to $G$ (see (II.ii)), and the inequality $f(w) \le \alpha f(x) + (1-\alpha)f(y)$ immediately follows from $\Phi(w) \le \alpha\Phi(x) + (1-\alpha)\Phi(y)$ and (II.iii).

(b) This part immediately follows from (II.ii).

(c) It suffices to lead to a contradiction the assumption that, for certain sequence of points $x_i \in U$ converging to a point $x$, the sequence $f(x_i)$ converges to certain $a < f(x)$ (here $-\infty \le a \le +\infty$). Under this assumption, $a < \infty$ for all large enough $i$, so that we can assume that the vectors $y_i = (\phi_1(x_i), \ldots, \phi_m(x_i))$ are well defined for all $i$ and that $\phi$ is finite and even above bounded along the sequence $\{y_i\}$ (otherwise, $a = \infty$).

Since $\phi_j$ is convex on $G_j$ and finite on $\operatorname{int} G_j$, $\phi_j$ is below bounded on each bounded subset of $G_j$. In particular, there exists a vector $t$ such that $y_i \ge t$ for all $i$. Also, since $\phi$ is above bounded along $\{y_i\}$, from (II.i) it follows that the sequence $\{y_i\}$ is bounded. Therefore, without loss of generality, we may assume that $\{y_i\}$ converges to certain $y \in G$.

Now $f(x_i) = \phi(y_i)$, and, since $y_i \to y$ and $\phi$ is lower semicontinuous, we have

$$\phi(y) \le a = \lim_{i\to\infty} f(x_i).$$

Furthermore, $\phi_j$ are lower semicontinuous, and $y_i = (\phi_1(x_i), \ldots, \phi_m(x_i))$; since $x_i \to x$, $y_i \to y$ as $i \to \infty$, we conclude that all $\phi_j$ are finite at $x$ and that $(\phi_1(x), \ldots, \phi_m(x)) \le y$. Now, from the definition of $f$ and (II.iii), it follows that $f(x) \le \phi(y)$, and, as it was already proved, $\phi(y) \le a$. Thus, $f(x) \le a$, which is the desired contradiction. $\square$

**Theorem 5.2.1** *Assume that*

$$\{(G_1,\phi_1),\dots,(G_m,\phi_m),(G,\phi)\}$$

*are regular superposition data.*

(i) *Let $\Gamma_i = (E_i^+, G_i^+, F_i^+, \pi_i)$ be $(\vartheta_i, l_i)$-coverings for functional elements $(G_i,\phi_i)$ on $E, 1 \le i \le m$ and let $\Gamma = (E^+, G^+, F^+, \pi)$ be a $(\vartheta, l)$-covering for the functional element $(G,\phi)$ on $\mathbf{R}^m$. Then the coverings $\Gamma_1,\dots,\Gamma_m,\Gamma$ induce a $(\sum_{i=1}^m \vartheta_i + \vartheta, \sum_{i=1}^m l_i + l + m)$-covering $\Gamma^\# = (E^\#, G^\#, F^\#, \pi^\#)$ for the functional element*

$$(U, f(x) = \phi(\phi_1(x),\dots,\phi_m(x))) .$$

*This covering is as follows. $\pi_i$ maps $E_i^+$ onto $\mathbf{R}\times E$,*

$$\pi_i(x_i) = (\tau_i(x_i), \sigma_i(x_i)), \qquad \tau_i(x_i) \in \mathbf{R}, \quad \sigma_i(x_i) \in E.$$

*Similarly, $\pi$ maps $E^+$ onto $\mathbf{R}\times\mathbf{R}^m$,*

$$\pi(x) = (\tau(x), \sigma(x)), \qquad \tau(x) \in \mathbf{R}, \quad \sigma(x) = (\sigma^1(x),\dots,\sigma^m(x))^T \in \mathbf{R}^m.$$

*We set*

$$E^\# = \{(x_1, x_2, \dots, x_m, x) \mid x_i \in E_i^+;\ i = 1,\dots,m,\ x \in E^+;\ \tau \in \mathbf{R}\};$$

$$\sigma_1(x_1) = \sigma_2(x_2) = \cdots = \sigma_m(x_m);\ \tau_1(x_1) = \sigma^1(x),\dots,\tau_m(x_m) = \sigma^m(x)\},$$

$$G^\# = \{(x_1, x_2, \dots, x_m, x) \in E^\# \mid x_i \in G_i^+,\ 1 \le i \le m;\ x \in G^+\},$$

$$F^\#(x_1,\dots,x_m,x) = F_1^+(x_1)+\cdots+F_m^+(x_m)+F^+(x),\ (x_1,x_2,\dots,x_m,x) \in \operatorname{int} G^\#,$$

$$\pi^\#(x_1,\dots,x_m,x) = (\tau(x),\sigma_1(x_1))\ (= (\tau(x),\sigma_2(x_2)) = \cdots = (\tau(x),\sigma_m(x_m))),$$
$$(x_1,x_2,\dots,x_m,x) \in E^\#.$$

(ii) *Let $\mathbf{K}_i = (E_i^+, K_i^+, F_i^+; H_i, \pi_i, \gamma_i)$ be $\vartheta_i$-conic representations of $(G_i,\phi_i)$, $i = 1,\dots,m$ and let $\mathbf{K} = (E^+, K^+, F^+; H, \pi, \gamma)$ be a $\vartheta$-conic representation of $(G,\phi)$. These representations induce a $(\sum_{i=1}^m \vartheta_i + \vartheta)$-conic representation*

$$\mathbf{K}^\# = (E^\#, K^\#, F^\#; H^\#, \pi^\#, \gamma^\#)$$

*for the functional element $(U, f(x) = \phi(\phi_1(x),\dots,\phi_m(x)))$. This representation is as follows. The mappings $\gamma_1,\dots,\gamma_m,\gamma$ map $H_1,\dots,H_m,H$ onto $\mathbf{R}\times E,\dots,\mathbf{R}\times E,\mathbf{R}\times\mathbf{R}^m$, respectively, so that these mappings can be represented as $\gamma_i(\cdot) = (\tau_i(\cdot),\sigma_i(\cdot))$, $\tau_i \in \mathbf{R}$, $\sigma_i \in E$, $i = 1,\dots,m$, and $\gamma(\cdot) = (\tau(\cdot),\sigma(\cdot))$, $\tau(\cdot) \in \mathbf{R}$, $\sigma(\cdot) \in \mathbf{R}^m$. We set*

$$E^\# = E_1^+ \times \cdots \times E_m^+ \times E^+,$$

$$K^\# = K_1^+ \times \cdots \times K_m^+ \times K^+,$$

$$F^{\#}(x_1,\dots,x_m,x) = F_1^+(x_1)+\cdots+F_m^+(x_m)+F^+(x),\ x_i\in\operatorname{int}K_i^+,\ x\in\operatorname{int}K,$$
$$H^{\#} = \{(u_1,\dots,u_m,u)\in H_1\times\cdots\times H_m\times H \mid \sigma_1(u_1)=\cdots=\sigma_m(u_m),$$
$$(\tau_1(u_1),\dots,\tau_m(u_m))=\sigma(u)\},$$
$$\pi^{\#}(u_1,\dots,u_m,u) = (\pi_1(u_1),\dots,\pi_m(u_m),\pi(u)),\ (u_1,\dots,u_m,u)\in H^{\#},$$
$$\gamma^{\#}(u_1,\dots,u_m,u) = (\tau(u),\sigma_1(u_1))\quad (=(\tau(u),\sigma_2(u_2))=\cdots=(\tau(u),\sigma_m(u_m))),$$
$$(u_1,\dots,u_m,u)\in H^{\#}.$$

**Proof.** (i) Let $u'\in\operatorname{int}U$ and let $t_i'>\phi_i(u')$, $1\le i\le m$. Then $(t_i',u')\in\operatorname{int}\mathsf{G}(G_i,\phi_i)$ and therefore there exist $x_i'\in\operatorname{int}G_i^+$ such that $\pi_i(x_i')=(t_i',u')$, $1\le i\le m$. Besides this, in view of (I.ii)–(I.iii) (or (II.ii)–(II.iii)), we have $(t_1',\dots,t_m')\in\operatorname{int}G$. Let $t'>\phi(t_1',\dots,t_m')$; then $(t',t_1',\dots,t_m')\in\operatorname{int}\mathsf{G}(G,\phi)$, so that there exists $x'\in\operatorname{int}G^+$ such that $\pi(x')=(t',t_1',\dots,t_m')$. The point $(x_1',\dots,x_m',x',t')$ clearly belongs to $\operatorname{int}G^{\#}$, so that the latter set (which evidently is closed and convex) is a closed convex domain in $E^{\#}$. A byproduct of our reasoning is the fact that the image of the natural embedding $\omega: E^{\#}\to E_1^+\times\cdots\times E_m^+\times E^+$ intersects $\operatorname{int}D$, where $D=G_1^+\times\cdots\times G_m^+\times G^+$ (the above intersection contains the point $(x_1',\dots,x_m',x')$).

The function $S(x_1,\dots,x_m,x) = F_1^+(x_1)+\cdots+F_m^+(x_m)+F^+(x)$ is a $(\sum_{i=1}^m\vartheta_i+\vartheta)$-self-concordant barrier for $D$ (Proposition 5.1.2); we clearly have $G^{\#}=\omega^{-1}(D)$, and, since $\omega(E^{\#})$ intersects $\operatorname{int}D$, the function $S\circ\omega$ (i.e., $F^{\#}$) is a $(\sum_{i=1}^m\vartheta_i+\vartheta)$-self-concordant barrier for $G^{\#}$.

Now let us prove that $\pi^{\#}(G^{\#})=\mathsf{G}(U,f)$. Let $(t,u)\in\mathsf{G}(U,f)$. Then $u\in G_i$, $1\le i\le m$, the values $t_i=\phi_i(u)$ are well defined, and $(t_i,u)\in\mathsf{G}(G_i,\phi_i)$. The latter relation means that there exists $x_i\in G_i^+$ such that $\pi_i(x_i)=(t_i,u)$. In view of (I.ii)–(I.iii) (or (II.ii)–(II.iii)), the vector $z=(t_1,\dots,t_m)^T$ belongs to $G$, and, since $(t,u)\in\mathsf{G}(U,f)$, we have $(t,z)\in\mathsf{G}(G,\phi)$, so that there exists $x\in G^+$ such that $\pi(x)=(t,z)$. Now the point $(x_1,\dots,x_m,x)$ clearly belongs to $G^{\#}$ and is such that $\pi^{\#}(x_1,\dots,x_m,x)=(t,u)$. Thus, $\pi^{\#}(G^{\#})$ contains $\mathsf{G}(U,f)$. To prove the inverse inclusion, let $(x_1,\dots,x_m,x)\in G^{\#}$. Then $x_i\in G_i^+$, so that the points $(\tau_i(x_i),\sigma_i(x_i))$ belong to $\mathsf{G}(G_i,\phi_i)$, and $u\equiv\sigma_1(x_1)=\cdots=\sigma_m(x_m)$ (the definition of $E^{\#}$). It follows that $u\in U$ and $t_i\equiv\tau_i(x_i)\ge\phi_i(u)$. Furthermore, from $(x_1,\dots,x_m,x)\in G^{\#}$, it also follows that $\pi(x)\equiv(t,z)\in\mathsf{G}(G,\phi)$ and $z=(t_1,\dots,t_m)^T$. The latter fact, combined with (I.iii) ($\equiv$ (II.iii)) and the relation $(t_1,\dots,t_m)\ge(\phi_1(u),\dots,\phi_m(u))$, implies that $t\ge\phi(z)\ge\phi(\phi_1(u),\dots,\phi_m(u))=f(u)$, so that $\pi^{\#}(x_1,\dots,x_m,x)=(t,u)\in\mathsf{G}(U,f)$. Thus, $\pi^{\#}(G^{\#})=\mathsf{G}(U,f)$.

It remains to compute the codimension of $\Gamma^{\#}$. Let us fix a point $w^0=(x_1^0,\dots,x_m^0,x^0)\in E^{\#}$. From the definition of $\pi^{\#}$, it follows that $\pi^{\#}(x_1,\dots,x_m,x)=\pi^{\#}(w^0)$ if and only if $\tau(x^0)=\tau(x)$ and $\pi_i(x_i)=(\sigma^i(x),\sigma_i(x_i^0))$. It follows that $\dim((\pi^{\#})^{-1}(w^0))=\sum_{i=1}^m l_i+m+l$.

(ii) Let us prove that $\gamma^{\#}([\pi^{\#}]^{-1}(K^{\#}))=\mathsf{G}(U,f)$. Let $(t,v)\in\mathsf{G}(U,f)$, so that $(t_i=\phi_i(v),v)\in\mathsf{G}(G_i,\phi_i)$, $i=1,\dots,m$, and $(t,(t_1,\dots,t_m))\in\mathsf{G}(G,\phi)$.

These relations mean that there exist $u_i \in [\pi_i]^{-1}(K_i^+)$, $i = 1, \ldots, m$, and $x \in \pi^{-1}(K^+)$, such that $(t_i, v) = \gamma_i(u_i)$, $i = 1, \ldots, m$, and $(t, (t_1, \ldots, t_m)) = \gamma(u)$. We clearly have $u^{\#} \equiv (u_1, \ldots, u_m, u) \in H^{\#}, \pi^{\#}(u^{\#}) \in K^{\#}$ and $\gamma^{\#}(u^{\#}) = (t, v)$. Thus, $\gamma^{\#}([\pi^{\#}]^{-1}(K^{\#})) \supset \mathsf{G}(U, f)$.

Conversely, let $u^{\#} = (u_1, \ldots, u_m, u) \in H^{\#}$ be such that $\pi^{\#}(u^{\#}) \in K^{\#}$ and let us verify that $\gamma^{\#}(u^{\#}) \in \mathsf{G}(U, f)$. In view of the definition of $H^{\#}$, the points $(t_i, v_i) \equiv \gamma_i(u_i)$, $i = 1, \ldots, m$, have the same $v$-component: $v_1 = \cdots = v_m \equiv v$, and, besides this, we have $\sigma(u) = (t_1, \ldots, t_m)$. Since $\pi^{\#}(u^{\#}) \in K^{\#}$, the $i$th of the above points $(t_i, v)$ belongs to $\mathsf{G}(G_i, \phi_i)$, while the point $\gamma(u) \equiv (t, \sigma(u))$ belongs to $\mathsf{G}(G, \phi)$. Thus, $v \in \bigcap_{i=1}^m G_i = U$, $t_i \geq \phi_i(v)$, $i = 1, \ldots, m$, and $t \geq \phi(\sigma(u)) = \phi(t_1, \ldots, t_m)$. Since $(G_1, \phi_1), \ldots, (G_m, \phi_m), (G, \phi)$ are regular superposition data, it follows that $t \geq f(v)$, so that $(t, v) \in \mathsf{G}(U, f)$. It remains to note that, by construction, $(t, v) = \gamma^{\#}(u^{\#})$, and the desired inclusion $\gamma^{\#}(u^{\#}) \in \mathsf{G}(U, f)$ is proved.

To prove that $\mathbf{K}^{\#}$ is a $(\sum_{i=1}^m \vartheta_i + \vartheta)$-conic representation of $(U, f)$, it suffices to verify, in addition to the already stated relation

$$\mathsf{G}(U, f) = \gamma^{\#}([\pi^{\#}]^{-1}(K^{\#})),$$

that

(1) $K^{\#}$ is a closed convex cone with a nonempty interior in $E^{\#}$;

(2) the image of $\pi^{\#}$ intersects $\operatorname{int} K^{\#}$;

(3) $\gamma^{\#}$ maps $H^{\#}$ onto $\mathbf{R} \times E$;

(4) $F^{\#}$ is a $(\sum_{i=1}^m \vartheta_i + \vartheta)$-logarithmically homogeneous self-concordant barrier for $K^{\#}$. We omit these quite straightforward verifications. □

Let us present a simple application of the superposition theorem.

**Corollary 5.2.3** *Let $\alpha_i > 0$, and let $(G_i, f_i)$ be functional elements on $E$, $1 \leq i \leq m$, such that* $\operatorname{int} (\bigcap_{i=1}^m G_i) \neq \emptyset$. *Then a collection of $(\vartheta_i, l_i)$-coverings (a collection of $\vartheta_i$-conic representations) of $(G_i, f_i)$ induces a $(\sum_{i=1}^m \vartheta_i + 1, \sum_{i=1}^m l_i + m)$-covering (respectively, $(\sum_{i=1}^m \vartheta_i + 1)$-conic representation) for the functional element $(\bigcap_{i=1}^m G_i, \sum_{i=1}^m \alpha_i f_i)$.*

Indeed, the resulting functional element is a superposition of the initial elements and the functional element $(\mathbf{R}^m, \phi(t_1, \ldots, t_m) = \sum_{i=1}^m \alpha_i t_i)$. These superposition data clearly satisfy (II.i)–(II.iii) and therefore are regular, and the external function admits a $(1, 0)$-covering (the function $F(t, t_1, \ldots, t_m) = -\ln (t - \phi(t_1, \ldots, t_m))$ is a 1-self-concordant barrier for its epigraph, see §2.3, Example 2), as well as an 1-conic representation (the latter is defined as $\mathbf{K} = (E^+ = \mathbf{R}, K^+ = \mathbf{R}_+, F^+(t) = -\ln t; H = \mathbf{R}^{m+1}, \pi(t, t_1, \ldots, t_m) = t - \sum_{i=1}^m \alpha_i t_i, \gamma(t, t_1, \ldots, t_m) = (t, t_1, \ldots, t_m))$.

### 5.2.5 How to solve convex problems with regular components

It is now time to explain the role that the above "calculus" plays in the interior-point methods. Assume that we to solve a problem

$$(\mathcal{I}): \qquad f_0(u) \to \min \mid u \in G,\ f_i(u) \le 0,\ 1 \le i \le m,$$

where $G$ is a closed convex domain in $E$ and $f_i$, $0 \le i \le m$ correspond to functional elements $(G_i, f_i)$, $G_i \supset G$. Assume that the constraints satisfy the Slater condition: The set $\{u \in \operatorname{int} G \mid f_i(x) < 0,\ 1 \le i \le m\}$ is nonempty.

There are two typical ways to apply the interior-point machinery to $(\mathcal{I})$. First, we can try to reduce the problem to an equivalent standard problem and to take care of providing the resulting feasible set with a "computable" self-concordant barrier. Second, we can look for an equivalent conic problem, again keeping in mind the necessity to provide the corresponding cone with a logarithmically homogeneous self-concordant barrier. Let us study the corresponding possibilities.

**A. To reduce $(\mathcal{I})$ to a standard problem.** We first should look for coverings for the involved data. Assume that we were successful and have found $(\vartheta_i, l_i)$-coverings $\Gamma_i = (E_i^+, G_i^+, F_i^+, \pi_i)$ for the functional elements $(G_i, f_i)$, $0 \le i \le m$, as well as a $(\vartheta, l)$-covering $\Gamma = (E^+, G^+, F^+, \pi)$ for $G$. Then we can transform $(\mathcal{I})$ into an equivalent standard problem and provide the feasible set of the latter problem with a self-concordant barrier, namely, as follows.

*First step.* We reformulate $(\mathcal{I})$ as

$$(\mathcal{J}) \quad t \to \min \mid (t,x) \in G^* \equiv \bar{G} \bigcap \left( \bigcap_{i=1}^{m} \bar{G}_i \right),$$

where

$$\bar{G} = \mathbf{R} \times G,$$

$$\bar{G}_0 = \mathsf{G}(G_0, f_0),$$

$$\bar{G}_i = \mathbf{R} \times \{u \in G_i \mid f_i(u) \le 0\}, \qquad 1 \le i \le m.$$

*Second step.* Note that $\Gamma$ induces a $(\vartheta, l)$-covering $\Gamma^\#$ for $\bar{G}$ (since $\bar{G}$ is the inverse image of $G$ under the natural projection of $\mathbf{R} \times E$ onto $E$, see Proposition 5.2.2); $\Gamma_0$ is a $(\vartheta_0, l_0)$-covering for $\bar{G}_0$; $\Gamma_i$ induce $(\vartheta_i + 1, l_i)$-coverings $\Gamma_i^\#$ for $\bar{G}_i$, $1 \le i \le m$ (see Propositions 5.2.6 and 5.2.2 and consider that $\{f_i\}$ satisfy the Slater condition). At the second step, we construct the induced coverings.

*Third step.* Since $(\mathcal{I})$ satisfies the Slater condition, $G^*$ is a closed convex domain, so that the coverings $\Gamma^\#, \Gamma_0, \Gamma_1^\#, \ldots, \Gamma_m^\#$ induce a $(\hat{\vartheta}, \hat{l})$-covering $\hat{\Gamma} = (\hat{E}, \hat{G}, \hat{F}, \hat{\pi})$ (see Proposition 5.2.4) for the set $G^*$, where $\hat{\vartheta} = \vartheta + \sum_{i=0}^m \vartheta_i + m$, $\hat{l} = l + \sum_{i=0}^m l_i$. At the third step, we compute the latter covering. Note that $\hat{\pi}(\cdot)$ can be represented as $\hat{\pi}(x) = (\tau(x), \sigma(x))$, $\tau(x) \in \mathbf{R}$, $\sigma(x) \in E$. Clearly, $(\mathcal{J})$ (and therefore $(\mathcal{I})$) is equivalent to the standard problem

$$(\mathcal{S}) \quad \tau(x) \to \min \mid x \in \hat{G},$$

and $\hat{F}$ is a $\hat{\vartheta}$-self-concordant barrier for the feasible domain of the latter problem, so that we can apply to it our interior-point methods.

Note that the updatings required at the second and the third steps are explicit and use "rational" linear algebra routines only.

**B. To reduce $(\mathcal{I})$ to a conic problem.** We might first reduce it to an equivalent standard problem $(\mathcal{S})$ (see above) and then transform $(\mathcal{S})$ into an equivalent conic problem $(\mathsf{P})$, according to the general scheme described in §4.1. Recall that a $\vartheta$-self-concordant barrier $F$ for the feasible domain of $(\mathcal{S})$ can be "explicitly" transformed into an $O(\vartheta)$-logarithmically homogeneous self-concordant barrier $F^{\#}$ for the cone involved into $(\mathsf{P})$ (see Proposition 5.1.4). Nevertheless, the above scheme sometimes is not the best one, since usually it is difficult to find an explicit representation for the Legendre transformation of $F^{\#}$, and consequently it is difficult to apply to $(\mathsf{P})$ those of potential reduction methods that require knowledge of both the primal and the dual barriers. This is the reason for the below reduction scheme, which, in a sense, is compatible with the Legendre transformation.

The scheme is as follows. First, we look for conic representations for the functional elements $(G_i, f_i)$, $i = 0, \ldots, m$ and for a conic representation of $G$. Assume that we were lucky and have found $\vartheta_i$-conic representations $\mathbf{K}_i = (E_i^+, K_i^+, F_i^+; H_i, \pi_i, \gamma_i)$ of $(G_i, f_i)$, as well as a $\vartheta$-conic representation $\mathbf{K} = (E^+, K^+, F^+; H, \pi, \gamma)$ of $G$; assume also that all the cones $K_i^+$, $K^+$ are pointed.

The mappings $\gamma_i$ take values in $\mathbf{R} \times E$; let

$$\gamma_i(\cdot) = (\tau_i(\cdot), \sigma_i(\cdot)), \qquad i = 0, \ldots, m,$$

where $\tau_i(\cdot) \in \mathbf{R}$ and $\sigma_i(\cdot) \in E$. Note that $\gamma(\cdot)$ also takes values in $E$.

Now consider the following problem:

$(\mathsf{P}^*)$ Minimize $\tau_0(u_0)$ by choice of

$$(u_0, u_1, \ldots, u_m, u) \in H^{\#} = H_0^+ \times \cdots \times H_m^+ \times H$$

under restrictions

$$\Pi(u_0, \ldots, u_m, u) \equiv (\pi_0(u_0), \ldots, \pi_m(u_m), \pi(u), (-\tau_1(u_1), \ldots, -\tau_m(u_m)))$$

$$\in K \equiv K_0^+ \times \cdots \times K_m^+ \times K^+ \times \mathbf{R}_+^m, \tag{5.2.1}$$

$$\sigma_0(u_0) = \sigma_1(u_1) = \cdots = \sigma_m(u_m) = \gamma(u). \tag{5.2.2}$$

Note that $K$ is a closed convex pointed cone with a nonempty interior in the space $E^{\#} = E_0^+ \times \cdots \times E_m^+ \times E^+ \times \mathbf{R}^m$.

From the definition of a conic representation, it follows immediately that $(\mathsf{P}^*)$ is equivalent to $(\mathcal{I})$ (if a point $(u_0, \ldots, u_m, u)$ is feasible to $(\mathsf{P}^*)$, then

$x \equiv \gamma(u)$ $(= \sigma_0(u_0) = \cdots = \sigma_m(u_m))$ is feasible for $(\mathcal{I})$ and $\tau_0(u_0) \geq f_0(x))$; conversely, for each feasible solution $x$ of $(\mathcal{I})$, there exists a feasible solution $(u_0, \ldots, u_m, u)$ of $(\mathsf{P}^*)$ such that $\sigma_0(u_0) = \cdots = \sigma_m(u_m) = \gamma(u) = x$ and $\tau_0(u_0) = f_0(x))$. At the same time, $(\mathsf{P}^*)$ is, in fact, a conic problem. Indeed, let $M$ be the affine space formed by $(u_0, \ldots, u_m, u)$ satisfying the *linear* system of equations (5.2.2) and let $N$ be the image of $M$ under the *affine* mapping $\Pi$. Then $(\mathsf{P}^*)$ is a problem of minimizing an affine functional $\tau_0$ over the intersection of the cone $K$ and an affine subspace $N$. The only difference with the definition of a conic problem is that the objective is expressed via certain parameterization of a vector from $N$, not via this vector directly. The latter difficulty can be immediately eliminated. Indeed, let $L$ be (any) level subspace of $\Pi\,|_M$ (all these subspaces are translations of each other). If $\tau_0(\cdot)$ is not constant on $L$, then $(\mathsf{P}^*)$ is either inconsistent or below unbounded (in fact, it is below unbounded, since $(\mathcal{I})$ is assumed to be feasible). Thus, in the case under consideration, $(\mathcal{I})$ is unsolvable, so that this case can be omitted. In the case when $\tau_0(\cdot)$ is constant along the level subspaces of $\Pi\,|_M$, we evidently can find a linear functional $\langle e, \cdot\rangle$ on the space $E^{\#}$ such that $\tau_0(u_0) = \langle e, \Pi(u_0, \ldots, u_m, u)\rangle +$ const, $(u_0, \ldots, u_m, u) \in M$, and the following conic problem:

(P) Minimize $\langle e, w\rangle$ subject to $w \in K \bigcap N$ is equivalent to $(\mathcal{I})$.

Now the function

$$F(z_0, \ldots, z_m, z, t) = \sum_{i=0}^{m} F_i^+(z_i) + F^+(z) - \sum_{i=1}^{m} \ln(t_i),$$

$z_i \in \operatorname{int} K_i^+$, $i = 0, \ldots, m$, $z \in \operatorname{int} K^+$, $t \in \operatorname{int} \mathbf{R}_+^m$, is a $(\sum_{i=0}^m \vartheta_i + \vartheta + m)$-logarithmically homogeneous self-concordant barrier for the cone $K$ (Proposition 5.1.2 combined with Example 2 from §2.3). It is easily seen that the Slater condition in $(\mathcal{I})$ implies that the feasible subspace of (P) intersects $\operatorname{int} K$, which is the main assumption on the problem required by the potential reduction interior point methods.

Now what about the Legendre transformation of the barrier $F^{\#}$? The main advantage of our now scheme is that there is no difficulties with this dual barrier, *provided that, from the very beginning, we are given not only barriers $F_i^+$, $F^+$, but also their Legendre transformations $F_i^*$, $F^*$*. Indeed, the cone dual to $K$ is the direct product of the cones dual to the initial cones and the cone dual to $\mathbf{R}_+^m$ (the latter cone is $\mathbf{R}_+^m$ itself), and the Legendre transformation of $F^{\#}$ is simply

$$F^*(\zeta_0, \ldots, \zeta_m, \zeta, s) = \sum_{i=0}^{m} F_i^*(\zeta_i) + F^*(\zeta) + \left(-\sum_{i=1}^{m} \ln(-s_i)\right) - m,$$

$\zeta_i \in -\operatorname{int}(K_i^+)^*$, $\zeta \in -\operatorname{int} K^*$, $s \in -\operatorname{int} \mathbf{R}_+^m$.

The above schemes motivate our interest to coverings and conic representations: for concrete applications of the schemes, see the next chapter.

## 5.3 Barriers for two-dimensional sets

It is now time to provide the above "equipment" for constructing barriers with "raw materials," i.e., with a list of barriers for concrete convex domains arising in convex programming. In this section, we present a number of barriers for two-dimensional sets (mainly, for the epigraphs of univariate convex functions: the power function, the exponent and the logarithm, the entropy). These barriers are useful at least in the following two situations:

(1) When interested in convex problems involving separable functionals, according to the general scheme of §5.2.5, we need coverings for functions of the type $f(x) = f_1(\langle c_1, x\rangle) + \cdots + f_k(\langle c_k, x\rangle)$, where $f_i$ are convex functions on the axis. Corollary 5.2.3 demonstrates that, to point out a covering for $f$, it suffices to find coverings for all $f_i$;

(2) When applying barrier calculus to obtain a barrier for a complicated function defined by a formula, we need coverings for operations involved into the formula and, in particular, for $\mathbf{R} \to \mathbf{R}$ updatings (like taking square or exponent).

Of course, from the computational viewpoint, there is no serious difficulties with the universal barrier for a two-dimensional convex domain, so that we could avoid any concrete two-dimensional considerations at all. This would, however, be not the best way. Thus, let us list a number of "explicit" two-dimensional barriers.

### 5.3.1 Power functions and the entropy function

**Proposition 5.3.1** *Let $\zeta(t)$ be a nondecreasing $\mathrm{C}^3$-smooth concave function on $(0, \infty)$ such that the quantity*

$$\alpha_\zeta = \min\left\{\alpha \geq 0 \mid \mid \zeta'''(t) \mid\leq \frac{3\alpha}{t} \mid \zeta''(t) \mid \ \forall t > 0\right\} \tag{5.3.1}$$

*is finite. Then the function*

$$\max{}^2\{\alpha_\zeta, 1\}\{-\ln t - \ln(\zeta(t) - x)\} \tag{5.3.2}$$

*is a $2\max^2\{\alpha_\zeta, 1\}$-self-concordant barrier for the set*

$$G^+ = \mathrm{cl}\,\{(t, x) \in \mathbf{R}^2 \mid t > 0,\ \zeta(t) > x\}.$$

*In particular, the epigraphs of the power functions, i.e., the sets*

$$G^p = \{(t, x) \in \mathbf{R}^2 \mid t \geq (x_+)^p\}, \qquad 1 \leq p < \infty$$

*admit 2-self-concordant barriers*

$$\{-\ln t - \ln(t^{1/p} - x)\}.$$

**Proof.** "*General*" *part.* The mapping $t \to \zeta(t) : (\operatorname{int} \mathbf{R}_+) \to \mathbf{R}$ is $(\mathbf{R}_+, \alpha_\zeta)$-compatible with $\mathbf{R}_+$ (§5.1.2.B, Example 3), and the identity mapping $x \to x : \mathbf{R} \to \mathbf{R}$ is $(\{0\}, 0)$-compatible with $\mathbf{R}$ (§5.1.2.B, Example 1). Therefore the mapping

$$\mathcal{A}(t,x) = (\zeta(t), x) : (\operatorname{int} \mathbf{R}_+) \times \mathbf{R} \to \mathbf{R}^2$$

is $(K, \alpha_\zeta)$-compatible with $\mathbf{R}_+ \times \mathbf{R}$, where $K = \mathbf{R}_+ \times \{0\}$ (Lemma 5.1.3(iv)). Now $G^+$ is the inverse image of the half-plane $G = \{(\tau, x) \mid \tau \geq x\}$ under the mapping $\mathcal{A}$, and the recessive cone of this half-plane contains $K$. In view of Proposition 5.1.7 as applied to the standard 1-self-concordant barriers $F(\tau, x) = -\ln(\tau - x)$ for $G$ and $\Phi(t,x) = -\ln t$ for the domain of $\mathcal{A}$, the function

$$\max^2\{\alpha_\zeta, 1\}\{-\ln t - \ln(\zeta(t) - x)\}$$

is a $2\max^2\{\alpha_\zeta, 1\}$-self-concordant barrier for $G^+$, as claimed.

"*Particular*" *part.* It suffices to apply the "general" part to the function $\zeta(t) = t^{1/p}$. □

**Proposition 5.3.2** *Let $f(x)$ be a $C^3$-smooth convex function on $(0, \infty)$ such that the quantity*

$$\alpha_f = \min\left\{\alpha \geq 0 \mid | f'''(x) | \leq \frac{3\alpha}{x} f''(x)\ \forall x > 0\right\} \tag{5.3.3}$$

*is finite. Then the function*

$$\max^2\{\alpha_f, 1\}\{-\ln x - \ln(t - f(x))\}$$

*is a $2\max^2\{\alpha_f, 1\}$-self-concordant barrier for the set*

$$G^+ = \operatorname{cl}\{(t,x) \in \mathbf{R}^2 \mid x > 0,\ t > f(x)\}.$$

*In particular, the epigraphs*

$$G_p = \operatorname{cl}\{(t,x) \in \mathbf{R}^2 \mid x > 0,\ t \geq x^p\}, \qquad p < 0$$

*of the decreasing power functions $x^p$ admit 2-self-concordant barriers*

$$\begin{cases} -\ln x - \ln(t - x^p), & 0 > p > -1, \\ -\ln t - \ln(x - t^{1/p}), & p < -1, \end{cases}$$

*and the epigraph*

$$G = \{(t,x) \in \mathbf{R}^2 \mid x \geq 0,\ t \geq x\ln(x)\}$$

*of the entropy function $x\ln(x)$ admits the 2-self-concordant barrier*

$$\{-\ln(x) - \ln(t - x\ln(x))\}.$$

**Proof.** *"General" part.* The mapping $-f(x) : \operatorname{int} \mathbf{R}_+ \to \mathbf{R}$ is $(\mathbf{R}_+, \alpha_f)$-compatible with $\mathbf{R}_+$ (§5.1.2.B, Example 3), and the identity mapping $t \to t$ is $(\{0\}, 0)$-compatible with $\mathbf{R}$. Therefore the mapping $\mathcal{A}(t,x) = (t, -f(x)) : \mathbf{R} \times (\operatorname{int} \mathbf{R}_+) \to \mathbf{R}^2$ is $(K, \alpha_f)$-compatible with $\mathbf{R} \times \mathbf{R}_+$, where $K = \{0\} \times \mathbf{R}_+$ (Lemma 5.1.3(iv)). Now $G^+$ is the inverse image of the half-plane $G = \{(t,y) \mid t + y \geq 0\}$ under the mapping $\mathcal{A}$, and the recessive cone of this half-plane contains $K$. In view of Proposition 5.1.7 as applied to the standard 1-self-concordant barriers $F(t,y) = -\ln(t+y)$ for $G$ and $\Phi(t,x) = -\ln x$ for the domain of $\mathcal{A}$, the function

$$\max^2\{\alpha_f, 1\}\{-\ln x - \ln(t - f(x))\}$$

is a $2\max^2\{\alpha_f, 1\}$-self-concordant barrier for $G^+$, as claimed.

*"Particular" part.* In the case of $p \in (-1, 0)$, it suffices to apply the "general" part to $f(x) = x^p$ ($\alpha_f = 3^{-1}(2-p) \leq 1$). The case of $p < -1$ can be immediately reduced to the previous one, since, for $t, x > 0$ and $p < 0$, we have $\{t \geq x^p\} \Leftrightarrow \{x \geq t^{1/p}\}$. In the case of the entropy function, it suffices to apply the "general" part to $f(x) = x \ln x$ ($\alpha_f = \frac{1}{3}$). □

### 5.3.2 The exponent and the logarithm

**Proposition 5.3.3** *The function*

$$-\ln(\ln(t-x) - \ln t$$

*is a 2-self-concordant barrier for the epigraph*

$$G^+ = \{t \geq e^x\}$$

*of the exponent, and the function*

$$-\ln(t + \ln x) - \ln x$$

*is a 2-self-concordant barrier for the epigraph*

$$\{x > 0,\ t \geq -\ln x\}$$

*of the function* $-\ln x$.

**Proof.** The first statement follows from Proposition 5.1.3 as applied to $\zeta(t) = \ln t$ ($\alpha_\zeta = \frac{2}{3}$). The second statement is an immediate corollary of the first one, in view of $\{(t,x) \mid t \geq e^x\} = \{(t,x) \mid \ln t \geq x\}$. □

## 5.4 Barriers for multidimensional domains

Below, we present barriers for some multidimensional convex domains (a polytope, a piecewise-quadratic bounded set, the second-order cone, the cone of symmetric positive-semidefinite matrices, the epigraph of the matrix norm).

### 5.4.1 Polytope

The standard logarithmic barrier for a polytope is given by Example 2, §2.3.

**Proposition 5.4.1** *Let*

$$G = \{x \in E \mid \langle a_i, x \rangle \leq b_i, \ 1 \leq i \leq m\}$$

*be a polytope defined by a set of linear constraints satisfying the Slater condition. Then the function*

$$\Psi(x) = -\sum_{i=1}^{m} \ln(b_i - \langle a_i, x \rangle) \tag{5.4.1}$$

*is an $m$-self-concordant barrier for $G$. If $G$ is a cone (i.e., $b_i = 0, \ 1 \leq i \leq m$), then $F$ is $m$-logarithmically homogeneous. If $G = \mathbf{R}^m_+$, so that $-a_i$ are the standard orths of $\mathbf{R}^m$ and $b_i = 0$, then the Legendre transformation of $F$ is, up to an additive constant, the function $F^*(s) = F(-s)$.*

### 5.4.2 Piecewise-quadratically bounded domains

The barriers for quadractically constrained domains are given by the the following statement.

**Proposition 5.4.2** *Let*

$$G = \{x \in E \mid f_i(x) \leq 0, \ 1 \leq i \leq m\},$$

*where all $f_i$ are convex quadratic forms and let the collection of these forms satisfy the Slater condition. Then the function*

$$\mathbf{B}(x) = -\sum_{i=1}^{m} \ln(-f_i(x)) \tag{5.4.2}$$

*is an $m$-self-concordant barrier for $G$.*

**Proof.** This is an immediate corollary of Lemma 3.3.1 and Proposition 5.1.3. Let us present another proof based on Proposition 5.1.7. Let $f$ be a convex quadratic form on $E$ such that the set $\{x \mid f(x) < 0\}$ is nonempty. Then the mapping $x \to -f(x)$ is $(\mathbf{R}_+, 0)$-compatible with $E$ (§5.1.2.B, Example 2), so that Proposition 5.1.7 as applied to the standard barrier $F(s) = -\ln(s)$ for $\mathbf{R}_+$ and the trivial (identically zero) barrier for $E$ implies that the function $-\ln(-f(x))$ is a 1-self-concordant barrier for the set $\mathrm{cl}\{x \mid f(x) < 0\}$, i.e., for the set $\{x \mid f(x) \leq 0\}$. Now the statement under consideration follows from Proposition 5.1.3 (barrier for an intersection). □

### 5.4.3 Second-order cone

Let $K$ be a second-order pointed cone in $\mathbf{R}^{n+1}$; under an appropriate choice of the coordinates, $K$ can be represented as the epigraph of the Euclidean norm,

$$K = \{(t, x) \in \mathbf{R} \times \mathbf{R}^n \mid t \geq \| x \|_2\}. \tag{5.4.3}$$

Note that the standard Euclidean structure on $\mathbf{R}^{n+1}$ allows us to identify $\mathbf{R}^{n+1}$ and $(\mathbf{R}^{n+1})^*$, and, under this identification, $K$ coincides with its dual cone.

**Proposition 5.4.3** *The function*

$$\mathbf{B}(t, x) = -\ln(t^2 - x^T x) \tag{5.4.4}$$

*is a* 2*-logarithmically homogeneous self-concordant barrier for cone* (5.4.3), *and the Legendre transformation of* $\mathbf{B}$ *coincides, up to an additive constant, with the function* $\mathbf{B}^*(t, x) = \mathbf{B}(-t, -x)$.

**Proof.** Let us use Proposition 5.1.8. Let

$$E = \mathbf{R}, \quad K = \mathbf{R}_+, \quad E' = \mathbf{R}^{n+1}, \quad E'' = \mathbf{R}, \quad T = \mathbf{R}_+, \quad G = \mathbf{R}_+,$$

$$\mathcal{Q}(u, v) = (u^T v) : E' \to E$$

and let $A(t)$ be the operator of multiplication by $t$ on $E'$: $A(t)u = tu$.

The above data clearly satisfy assumptions (A.i)–(A.iii) of §5.1.2.B, so that Proposition 5.1.8, as applied to the standard barriers $F(s) = \Phi(s) = -\ln s$ for the half-axes $G$ and $T$, implies that the function

$$\Psi(y, x, t) = -\ln(y - t^{-1} x^T x)) - \ln t = -\ln(ty - x^T x)$$

is a 2-self-concordant barrier for the set

$$\begin{aligned} G^+ &= \mathrm{cl}\,\{(y, x, t) \in E^+ \equiv E \times E' \times E'' \mid t \in \mathrm{int}\, T,\ y - \mathcal{Q}(A^{-1}(t)x, x) \in \mathrm{int}\, G\} \\ &= \mathrm{cl}\,\{(y, x, t) \in E^+ \equiv \mathbf{R} \times \mathbf{R}^{n+1} \times \mathbf{R} \mid t > 0,\ y > t^{-1} x^T x\} \\ &= \{(y, x, t) \in \mathbf{R} \times \mathbf{R}^{n+1} \times \mathbf{R} \mid t \geq 0,\ y \geq 0,\ ty \geq x^T x\}. \end{aligned}$$

We see that $K$ is the inverse image of $G^+$ under the linear mapping

$$(t, x) \to (t, x, t),$$

and Proposition 5.1.1 as applied to $\Psi$ implies that $\mathbf{B}$ is a 2-self-concordant barrier for $K$. The 2-logarithmic homogeneity of the barrier is evident, and an immediate computation demonstrates that the Legendre transformation of $\mathbf{B}$ coincides, up to an additive constant, with $\mathbf{B}(-t, -x)$. □

### 5.4.4 Epigraphs of functions of the Euclidean norm

**Proposition 5.4.4** *Let $G$ be a closed convex domain in $\mathbf{R}^2$ such that $(u,t) \in G$ implies $(v,t) \in G$ for all $v \le u$ (in other words, such that $\mathcal{R}(G)$ contains the vector $(-1,0)^T$) and let $F$ be a $\vartheta$-self-concordant barrier for $G$. Assume that $G$ contains a point with a positive first coordinate. Then the function*

$$\mathbf{B}^{(1)}(t,x) = F(x^T x, t)$$

*is a $\vartheta$-self-concordant barrier for the closed convex domain*

$$G_1^+ = \{(t,x) \in \mathbf{R} \times \mathbf{R}^n \mid (x^T x, t) \in G\}.$$

*If $G$ contains a point with both coordinates being positive, then the function*

$$\mathbf{B}^{(2)}(t,x) = F\left(\frac{x^T x}{t}, t\right) - \ln t$$

*is a $(\vartheta+1)$-self-concordant barrier for the closed convex domain*

$$G_2^+ = \mathrm{cl}\left\{(t,x) \in \mathbf{R} \times \mathbf{R}^n \mid t > 0,\ (\frac{x^T x}{t}, t) \in \mathrm{int}\, G\right\}.$$

*In particular, let $\zeta(t)$ be a nondecreasing concave function on $(0,\infty)$, which is $\mathrm{C}^3$-smooth and is positive for large enough $t$ and satisfies the condition*

$$\alpha_\zeta = \min\left\{\alpha \ge 0 \mid \mid \zeta'''(t) \mid \le \frac{3\alpha}{t} \mid \zeta''(t) \mid \ \forall t > 0\right\} < \infty.$$

*Then the function*

$$\mathbf{B}_\zeta^{(1)}(t,x) = \max{}^2\{1,\alpha_\zeta\}\{-\ln(\zeta(t) - x^T x) - \ln t\}$$

*is a $2\max^2\{1,\alpha_\zeta\}$-self-concordant barrier for the set*

$$\mathrm{cl}\,\{(t,x) \in \mathbf{R} \times \mathbf{R}^n \mid t > 0,\ \| x \|_2^2 \le \zeta(t)\},$$

*and the function*

$$\mathbf{B}_\zeta^{(2)}(t,x) = -\max{}^2\{1,\alpha_\zeta\}\ln(t\zeta(t) - x^T x) - \ln t$$

*is a $(2\max^2\{1,\alpha_\zeta\}+1)$-self-concordant barrier for the set*

$$\mathrm{cl}\,\{(t,x) \in \mathbf{R} \times \mathbf{R}^n \mid t > 0,\ \| x \|_2^2 \le t\zeta(t)\}.$$

*For example, if $p \ge 2$, then the functions*

$$\mathbf{B}_p(t,x) = -\ln(t^{2/p} - x^T x) - \ln t \qquad (5.4.5)$$

*are 2-self-concordant barriers for the epigraphs*

$$\{(t,x) \in \mathbf{R} \times \mathbf{R}^n \mid t \geq \| x \|_2^p\}$$

*of the powers of the Euclidean norm, and if* $1 \leq p < 2$, *then the functions*

$$\mathbf{B}_p(t,x) = -\ln(t^{2/p} - x^T x) - 2\ln t \tag{5.4.6}$$

*are 3-self-concordant barriers for the above epigraphs.*

**Proof.** "*General*" *part.* The mapping $x \to x^T x : \mathbf{R}^n \to \mathbf{R}$ clearly is $(\mathbf{R}_-, 0)$-compatible with $\mathbf{R}^n$ (§5.1.2.B, Example 2), while the mapping $t \to t$ is $(\{0\}, 0)$-compatible with $\mathbf{R}$ (§5.1.2.B, Example 1). It follows that the mapping

$$\mathcal{A}(t,x) = (x^T x, t) : \mathbf{R} \times \mathbf{R}^n \to \mathbf{R}^2$$

is $(K,0)$-compatible with $\mathbf{R} \times \mathbf{R}^n$, where $K$ is the cone $\mathbf{R}_- \times \{0\}$ in $\mathbf{R}^2$ (i.e., $K$ is the ray generated by the vector $(-1,0)^T$). By assumption, the recessive cone of $G$ contains $K$, and the image of $\mathcal{A}$ intersects $\operatorname{int} G$. Therefore, Proposition 5.1.7 implies that $\mathbf{B}^{(1)}(t,x)$ is a $\vartheta$-self-concordant barrier for the set $G_1^+$.

Furthermore, let

$$E = \mathbf{R}^2, \quad K = \mathbf{R}_- \times \{0\}, \quad E' = \mathbf{R}^n, \quad E'' = \mathbf{R}, \quad T = \mathbf{R}_+, \tag{5.4.7}$$

$$\mathcal{Q}(u,v) = (u^T v)(-1,0)^T : \mathbf{R}^n \to \mathbf{R}^2$$

and let $A(t)$ be the operator of multiplication by $t$ on $E'$: $A(t)x = tx$. The above data clearly satisfy assumptions (A.i)–(A.iii) from §5.1.2.B, and the recessive cone of $G$ contains $K$. In view of Proposition 5.1.8(ii) as applied to the self-concordant barriers $F$ for $G$ and $\Phi(s) = -\ln(s)$ for $T$, the function

$$\Psi(y,x,t) = F\left(y - \frac{x^T x}{t}(-1,0)^T\right) - \ln t$$

is a $(\vartheta + 1)$-self-concordant barrier for the set

$$G^+ = \operatorname{cl}\left\{(y,x,t) \in \mathbf{R}^2 \times \mathbf{R}^n \times \mathbf{R} \mid t > 0,\ y - \frac{x^T x}{t}(-1,0)^T \in \operatorname{int} G\right\}.$$

The set $G_2^+$ clearly is the inverse image of $G^+$ under the linear mapping

$$(t,x) \to ((0,t)^T, x, t) : \mathbf{R} \times \mathbf{R}^n \to \mathbf{R}^2 \times \mathbf{R}^n \times \mathbf{R},$$

and the image of this mapping intersects $\operatorname{int} G^+$ (since $G$ contains a point with both coordinates being positive), and Proposition 5.1.1 as applied to $\Psi$ and this mapping implies that $G_2^+$ is a closed convex domain; $\mathbf{B}^{(2)}(t,x)$ is a $(\vartheta + 1)$-self-concordant barrier for this domain, as claimed.

*"Particular" part.* This is an immediate consequence of the "general" part and Proposition 5.3.1. Note that the barriers for the sets

$$\mathrm{cl}\{(t,x)\in\mathbf{R}\times\mathbf{R}^n \mid t>0, \|x\|_2^2\le\zeta(t)\}$$

and, in particular, those for the epigraphs of $\|\cdot\|_2^p$ for $p\ge 2$ are of the type $\mathbf{B}^{(1)}$, and the barriers for the sets

$$\mathrm{cl}\{(t,x)\in\mathbf{R}\times\mathbf{R}^n \mid t>0, \ \|x\|_2^2\le t\zeta(t)\},$$

(in particular, for the epigraphs of $\|\cdot\|_2^p$ for $p<2$) are of the type $\mathbf{B}^{(2)}$. □

### 5.4.5 Cone of positive-semidefinite symmetric matrices

Let $S_n$ denote the space of *symmetric* $n\times n$ matrices with real entries; as usual, $S_n$ is provided with the inner product $\langle x,y\rangle=\mathrm{Tr}\{xy\}$, Tr being the trace. Let $S_n^+$ be the cone of positive-semidefinite matrices from $S_n$. Note that the above Euclidean structure allows us to identify the conjugate to $S_n$ space $S_n^*$ with $S_n$, and, under this identification, the cone $S_n^+$ coincides with its dual cone.

**Proposition 5.4.5** *The function*

$$F(x)=-\ln\mathrm{Det}(x) \tag{5.4.8}$$

*is an $n$-logarithmically homogeneous self-concordant barrier for the cone $S_n^+$, and its Legendre transformation coincides, up to an additive constant, with the function $F^*(x)=F(-x)$.*

**Proof.** The function $F$ clearly is $C^\infty$-smooth on int $S_n^+$ and tends to $\infty$ as the argument approaches a point from $\partial S_n^+$. Besides this, $F$ is $n$-logarithmically homogeneous. Thus, in view of Corollary 2.3.2, to prove the statement, it suffices to verify the inclusion

$$F\in S_1(\mathrm{int}\,S_n^+,S_n). \tag{5.4.9}$$

For $x\in\mathrm{int}\,S_n^+$, $h\in S_n$ we have

$$\begin{aligned}DF(x)[h]&=-\frac{d}{dt}\Big|_{t=0}\ln\mathrm{Det}(x+th)\\&=-\frac{d}{dt}\Big|_{t=0}\{\ln\mathrm{Det}(x)+\ln\mathrm{Det}(I+tx^{-1}h)\}=-\mathrm{Tr}\{x^{-1}h\},\end{aligned}$$

$$D^2F(x)[h,h]=-\frac{d}{dt}\Big|_{t=0}\mathrm{Tr}\{(x+th)^{-1}h\}=\mathrm{Tr}\{x^{-1}hx^{-1}h\},$$

$$D^3F(x)[h,h,h]=\frac{d}{dt}\Big|_{t=0}\mathrm{Tr}\{(x+th)^{-1}h(x+th)^{-1}h\}=-2\,\mathrm{Tr}\{x^{-1}hx^{-1}hx^{-1}h\}.$$

Let $h^* = x^{-1/2}hx^{-1/2}$; then

$$D^2F(x)[h,h] = \mathrm{Tr}\{(h^*)^2\} \geq 0, \qquad | D^3F(x)[h,h,h] |= 2 \mid \mathrm{Tr}\{(h^*)^3\} \mid .$$

In other words, if $\{\nu_i\}$ denote the eigenvalues of the (symmetric) matrix $h^*$, then

$$| D^3F(x)[h,h,h] | \leq 2\sum_{i=1}^{n} | \nu_i |^3 \leq 2\left\{\sum_{i=1}^{n} \nu_i^2\right\}^{3/2} = 2(D^2F(x)[h,h])^{3/2},$$

which leads to (5.4.9). □

**Remark 5.4.1** *Note that* (5.4.8) *is the best possible self-concordant barrier for* $S_n^+$. *Indeed, an appropriate cross section of* $S_n^+$ *is* $\mathbf{R}_+^n$; *therefore, from Propositions* 5.1.1 *and* 2.3.5, *it follows that the value of the parameter of a self-concordant barrier for* $S_n^+$ *cannot be less than* $n$.

**Remark 5.4.2** *The barrier* $-\ln \mathrm{Det}(x)$ *can be not only guessed, but also derived from Proposition* 5.1.8. *Indeed, a symmetric* $n \times n$ *matrix*

$$X = \begin{pmatrix} y & x^T \\ x & t \end{pmatrix}$$

*($t$ is $(n-1)\times(n-1)$) is positive-definite if and only if $t$ is symmetric positive-definite and $y > x^T t^{-1} x$, so that*

$$S_n^+ = \mathrm{cl}\left\{\begin{pmatrix} y & x^T \\ x & t \end{pmatrix} \middle| \mathcal{A}(y,x,t) \equiv y - x^T t^{-1} x \in \mathrm{int}\,\mathbf{R}_+,\ t \in \mathrm{int}\, S_{n-1}^+\right\}.$$

*Proposition* 5.1.8(ii) *states that, if $F_{n-1}$ is a $\vartheta_{n-1}$-self-concordant barrier for $S_{n-1}^+$, then the function*

$$F_n\left(\begin{pmatrix} y & x^T \\ x & t \end{pmatrix}\right) = -\ln(y - x^T t^{-1} x) + F_{n-1}(t)$$

*is a $\vartheta_n$-self-concordant barrier for $S_n^+$, $\vartheta_n = \vartheta_{n-1} + 1$. Since $S_1^+ = \mathbf{R}_+$, we can start this recurrence for barriers for $S^+$ with $F_1(t) = -\ln t$, $\vartheta_1 = 1$, and an immediate induction leads to $F_n(x) = -\ln \mathrm{Det}(x)$, $x \in \mathrm{int}\, S_n^+$.*

### 5.4.6 Epigraph of the matrix norm

Let $L_{n,m}$ denote the space of $n \times m$ matrices with real entries. We provide $L_{n,m}$ with the standard inner product $\langle x, y\rangle = \mathrm{Tr}\{xy^T\}$. Also, let $\| x \|$ denote the standard operator norm of a matrix $x$ associated with the standard Euclidean norms on $\mathbf{R}^m$ and $\mathbf{R}^n$,

$$\| x \| = \max\left\{\frac{\| xs \|_2}{\| s \|_2} \mid s \in \mathbf{R}^m \backslash \{0\}\right\}.$$

Note that Proposition 5.4.5 gives a $(m+n)$-logarithmically homogeneous self-concordant barrier for the epigraph

$$\mathsf{G} = \{(t,x) \in \mathbf{R} \times L_{n,m} \mid t \geq \| x \|\} \tag{5.4.10}$$

of the function $\| \cdot \|$. Indeed, $\mathsf{G}$ clearly is the inverse image of the cone $S^+_{n+m}$ under the linear mapping

$$(t,x) \to \begin{pmatrix} tI_m & x^T \\ x & tI_n \end{pmatrix},$$

where $I_k$ denotes the unit $k \times k$ matrix, and, in view of Proposition 5.1.1, the superposition of barrier (5.4.8) and this mapping is the $(m+n)$-self-concordant barrier for $\mathsf{G}$. Our goal now is to demonstrate that $\mathsf{G}$ admits a better logarithmically homogeneous self-concordant barrier, namely, with the parameter $\min\{n,m\}+1$.

**Proposition 5.4.6** (i) *The epigraph* (5.4.10) *of the matrix norm admits an* $(n+1)$*-logarithmically homogeneous self-concordant barrier, namely,*

$$-\ln \operatorname{Det}\left(tI_n - \frac{xx^T}{t}\right) - \ln t, \tag{5.4.11}$$

*as well as an* $(m+1)$*-logarithmically homogeneous self-concordant barrier, namely,*

$$-\ln \operatorname{Det}\left(tI_m - \frac{x^T x}{t}\right) - \ln t. \tag{5.4.12}$$

(ii) *The epigraph*

$$\mathsf{G}^{\#} = \{(t,x) \in \mathbf{R} \times L_{m,n} \mid t \geq \| x \|^2\}$$

*of the squared matrix norm admits self-concordant barriers*

$$-\ln \operatorname{Det}(tI_n - xx^T) \quad \textit{and} \quad -\ln \operatorname{Det}(tI_m - x^T x)$$

*with the parameters* $n, m$*, respectively.*

**Proof.** (i) The proof of this part immediately follows from Proposition 5.1.8. Indeed, let us set

$$E = S_n, \quad K = G = S_n^+, \quad E' = L_{n,m}, \quad E'' = \mathbf{R}, \quad T = \mathbf{R}_+,$$

$$\mathcal{Q}(x,u) = \tfrac{1}{2}(xu^T + ux^T) : E' \to E$$

and let $A(t)$, $t \in E''$ be the operator of multiplication by $t$ on $E'$.

The above data clearly satisfy assumptions (A.i)–(A.iii) of §5.1.2.B, and, in view of Proposition 5.1.8(ii), the barriers $F(z) = -\ln \operatorname{Det}(z)$ for $S_n^+$ (see

Proposition 5.4.5) and $\Phi(t) = -\ln t$ for $T$ induce the $(n+1)$-self-concordant barrier

$$\Psi(y,x,t) = -\ln \mathrm{Det}\left(y - \frac{xx^T}{t}\right) - \ln t$$

for the set

$$\begin{aligned} G^+ &= \mathrm{cl}\left\{(y,x,t) \in S_n \times L_{n,m} \times \mathbf{R} \mid t > 0, y - \frac{xx^T}{t} \in \mathrm{int}\, S_n^+\right\} \\ &= \{(y,x,t) \in S_n \times L_{n,m} \times \mathbf{R} \mid t \geq 0,\ y \in S_n^+,\ ty - xx^T \in S_n^+\}. \end{aligned}$$

It is clearly seen that the epigraph of the matrix norm is the inverse image of $G^+$ under the affine mapping $(t,x) \to (tI_n, x, t)$, so that Proposition 5.1.1 allows us to conclude that (5.4.11) is an $(n+1)$-self-concordant barrier for the epigraph of the matrix norm.

To prove that (5.4.12) is an $(m+1)$-self-concordant barrier for the set (5.4.10), we can use similar reasoning, starting with $E = S_m$, $K = G = S_m^+$ and $\mathcal{Q}(x,u) = \frac{1}{2}(x^Tu + u^Tx)$.

(ii) Note that $\mathrm{G}^\#$ is the inverse image of the cone $S_n^+$ of positive-semidefinite symmetric $n \times n$ matrices under the mapping $\mathcal{A}(t,x) = (tI_n - xx^T)$, which clearly is concave with respect to $\mathcal{R}(S_n^+) = S_n^+$. Therefore the first statement in (ii) is an immediate consequence of Propositions 5.1.7 and 5.4.5. To prove the second statement, it suffices to replace the above quadratic mapping by $\mathcal{A}(t,x) = tI_m - x^Tx$. □

### 5.4.7 Epigraphs of fractional-quadratic functions

The general fractional-quadratic function of an $n$-dimensional vector $x$ and a symmetric $n \times n$ matrix $X$, $\phi(x,X) = x^TX^{-1}x$, is convex in $(x,X)$ on the set of all pairs $(x,X)$ with positive-definite $X$, and the epigraph $G$ of the lower semicontinuous closure of this function is exactly the set

$$\left\{(u,x,X) \in \mathbf{R} \times \mathbf{R}^n \times S_n \mid \begin{pmatrix} u & x^T \\ x & X \end{pmatrix} \ \text{is positive-semidefinite}\right\};$$

i.e., it is an inverse image of the cone $S_{n+1}^+$ of positive-semidefinite symmetric $(n+1) \times (n+1)$ matrices under an affine mapping. Therefore Propositions 5.1.1 and 5.4.5 imply that the function

$$(5.4.13)\quad F(u,x,X) = -\ln \mathrm{Det}\begin{pmatrix} u & x^T \\ x & X \end{pmatrix} = -\ln(u - x^TX^{-1}x) - \ln \mathrm{Det}\, X$$

is an $(n+1)$-self-concordant barrier for $G$.

In some applications (e.g., in the problem of truss topology design, see [B-TN 92]), we are interested in the epigraphs of more specific quadratic-fractional functions, namely,

$$(5.4.14)\qquad f(x,t) = x^TX^{-1}(t)x,$$

where $X(t) = X_0 + X_1t_1 + \cdots + X_kt_k$ is an affine function of a *nonnegative* $t \in \mathbf{R}^k$ and $X_i$, $i = 0, \ldots, k$ are positive-semidefinite symmetric $n \times n$ matrices with a positive-definite sum. The corresponding epigraph is

$$(5.4.15)\quad G_f = \mathrm{cl}\,\{(u, x, t) \in \mathbf{R} \times \mathbf{R}^n \times \mathbf{R}^k \mid t > 0,\ u > x^T X^{-1}(t)x\}.$$

A self-concordant barrier for $G_f$ could be derived from barrier (5.4.13) with the aid of the barrier calculus,

$$(5.4.16)\quad \begin{aligned} F_f(u,x,t) &= F(u,x,X(t)) - \sum_{i=1}^k \ln(t_i) \\ &= -\ln(u - x^T X^{-1}(t)x) - \ln \operatorname{Det} X(t) - \sum_{i=1}^k \ln t_i \end{aligned}$$

(we have taken into account that $f$ is the restriction onto the nonnegative orthant $\mathbf{R}^k_+$ of the superposition of $\phi$ and the affine substitution $t \to X(t)$; the first two terms of the right-hand side form the barrier for the superposition (Proposition 5.1.1), and the third term $-\sum_{i=1}^k \ln t_i$ is responsible for the restriction of this superposition onto the nonnegative orthant; see Propositions 5.1.3 and 5.4.1). The parameter of the barrier (5.4.16) is $(n + 1 + k)$. Note that, in this barrier, positive semidefiniteness of $X(t)$ is penalized twice: explicitly (the term $-\ln \operatorname{Det} X(t)$) and implicitly (the term $-\sum_{i=1}^k \ln t_i$); recall that the matrix coefficients $X_i$ are assumed to be positive-semidefinite with positive-definite sum, so that positivity of $t$ implies positive definiteness of $X(t)$. It turns out that this redundancy can be eliminated: We can omit the term $-\ln \operatorname{Det} X(t)$, which reduces the parameter of the barrier to $k + 1$ and significantly simplifies the barrier from the computational viewpoint. This is a particular case of the following statement.

**Proposition 5.4.7** (see [B-TN 92]) *Let $q$, $n$, $k$ be a triple of positive integers and let $X_i(t)$, $i = 1, \ldots, q$, be affine functions of $t \in \mathbf{R}^k$ taking values in the space $S_n$ of symmetric $n \times n$ matrices, such that $X_i(t)$ is positive-definite for $t > 0$. Let*

$$f_i(x^i, t) = (x^i)^T X_i^{-1}(t) x^i : \mathbf{R}^n \times (\operatorname{int} \mathbf{R}^k_+) \to \mathbf{R}, \qquad i = 1, \ldots, q$$

*be the quadratic-fractional functions associated with $X_i(\cdot)$. Then the "epigraph" of the vector-valued function $(f_1(x^1, t), \ldots, f_q(x^q, t))$, i.e., the set*

$$G^+ = \mathrm{cl}\, G',$$

*where*

$$\begin{aligned} G' = \{(y, x^1, \ldots x^q, t) \in \mathbf{R}^q \times (\mathbf{R}^n)^q \times \mathbf{R}^k \mid t > 0; \\ y_i \geq (x^i)^T X_i^{-1}(t) x^i,\ i = 1, \ldots, q\}, \end{aligned}$$

*admits a $(q + k)$-self-concordant barrier*

$$(5.4.17)\quad \Psi(y, x^1, \ldots, x^q, t) = -\sum_{i=1}^q \ln(y_i - (x^i)^T X_i^{-1}(t) x^i) - \sum_{i=1}^k \ln t_i.$$

**Proof.** For a direct proof, see [B-TN 92]; below, we present a simple proof based on Proposition 5.1.8. Set

$$E = \mathbf{R}^q, \quad K = G = \mathbf{R}^q_+, \quad E' = (\mathbf{R}^n)^q, \quad E'' = \mathbf{R}^k, \quad T = \mathbf{R}^k_+$$

and let, for $x = (x^1, \dots, x^q)$, $u = (u^1, \dots, u^q) \in (\mathbf{R}^n)^q$

$$\mathcal{Q}(x, u) = \begin{pmatrix} (x^1)^T u^1 \\ \cdots \\ (x^q)^T u^q \end{pmatrix},$$

$$A(t)x = (X_1(t)x^1, \dots, X_q(t)x^q).$$

The above data clearly satisfy assumptions (A.i)–(A.iii) of §5.1.2.B, and the inverse image of $G$ under the mapping

$$\mathcal{A}(y, x, t) = y - \mathcal{Q}(A^{-1}(t)x, x) : E \times E' \times T \to E,$$

i.e., the set

$$\mathrm{cl}\,\{(y, x^1, \dots x^q, t) \in \mathbf{R}^q \times (\mathbf{R}^n)^q \times \mathbf{R}^k \mid t > 0;\ y_i > (x^i)^T X_i^{-1}(t) x^i,\ i = 1, \dots, q\},$$

is precisely the set $G^+$ mentioned in the statement we are proving. It remains to note that (5.4.17) is precisely the barrier associated, by virtue of Proposition 5.1.8(ii), with the standard logarithmic barriers $F(y) = -\sum_{i=1}^q \ln y_i$ for $G$ and $\Phi(t) = -\sum_{i=1}^k \ln t_i$ for $T$. □

## 5.5 Volumetric barrier

### The problem

The problem we study is as follows. The standard logarithmic barrier $F$ (see (5.4.1)) for a $n$-dimensional polytope defined by $m$ linear inequalities has the parameter value equal to $m$. At the same time, we know that every $n$-dimensional closed convex domain admits an $O(n)$-self-concordant barrier (Theorem 2.5.1); unfortunately, the latter barrier, as a rule, is too complicated and cannot be used in practical computations. Thus, in the case of $m \gg n$, which is the usual situation for real-world LP problems, barrier (5.4.1) is far from being optimal: The efficiency estimate of the related interior-point methods turns out to depend on the larger size of a problem instead of its smaller size. Of course, it would be very important to find a "computable" $O(n)$-self-concordant barrier for a $n$-dimensional polytope.

This problem still is unsolved, but recently Vaidya managed to make a breakthrough in this direction. He developed two new barriers for the $n$-dimensional polytope $G$ defined by $m$ linear inequalities: the *volumetric* barrier

$$\Phi(x) = O(1)\sqrt{m} \ln \mathrm{Det}(F''(x)),$$

and the *combined volumetric* barrier

$$\Psi(x) = O(1)\left\{\frac{1}{\sqrt{n}}\Phi(x) + \sqrt{\frac{m}{n}}F(x)\right\},$$

$F$ being the standard logarithmic barrier for $G$. It turns out that the parameters of these barriers are $O(n\sqrt{m})$ and $O(\sqrt{nm})$, respectively. Note that the cost at which we can compute the gradients and the Hessians of these barriers is of the same order as that for $F$.

In this section, we generalize these barriers as follows. Note that the standard logarithmic barrier (5.4.1) for $G$ can be obtained from the $m$-self-concordant barrier $F_*(y) = -\sum_{i=1}^m \ln y_i$ for the nonnegative $m$-dimensional orthant; to obtain $F$, we represent $G$ as the inverse image of $\mathbf{R}^m_+$ under an appropriate embedding $\tau : \mathbf{R}^n \to \mathbf{R}^m$ and set $F = F_* \circ \tau$. Now $F_*$, in turn, is the result of the same construction as applied to the embedding $\rho : \mathbf{R}^m \to S_m$ (an $m$-dimensional vector is regarded as a diagonal $m \times m$-matrix) and the standard $m$-self-concordant barrier (see Proposition 5.4.6) $F^+(\xi) = -\ln \operatorname{Det} \xi$ for the cone $S_m^+$ of positive-semidefinite symmetric $m \times m$ matrices: $F_* = F^+ \circ \rho$, so that

$$F = F^+ \circ \sigma, \qquad \sigma = \rho \circ \tau.$$

It turns out that each barrier of the latter type, i.e., the superposition of the standard barrier $F^+$ for the cone $S_m^+$ of symmetric positive-semidefinite $m \times m$ matrices and an affine embedding of $\mathbf{R}^n$ into $S_m^+$, generates volumetric and combined volumetric barriers with the same as in the case studied by Vaidya values of the parameters. The advantage of this generalization is that the corresponding family of $n$-dimensional convex domains, i.e., the family $F_{n,m}$ of inverse images of $S_m^+$ under affine embeddings, is much wider than the family of $n$-dimensional polytopes and contains some nonpolyhedral domains important for convex programming.

We demonstrate that an affine mapping $\sigma : \mathbf{R}^n \to S_m$ such that $\sigma(\mathbf{R}^n)$ intersects $\operatorname{int} S_m^+$ generates an $O(1)(mn)^{1/2}$-self-concordant barrier for the $n$-dimensional closed convex domain $G = \sigma^{-1}(S_m^+)$. It suffices to establish this result for *homogeneous embeddings* $\sigma$; indeed, the reduction to the case when $\sigma$ is an embedding is evident. Now, if $\sigma(x) = Ax + b$ is an embedding and $b \in \operatorname{Im} A$, then an appropriate translation of the origin in $\mathbf{R}^n$ makes $\sigma$ homogeneous; if $b \notin \operatorname{Im} A$, we can regard $\mathbf{R}^n$ as the hyperplane $\Pi = \{(t, x) \in \mathbf{R} \times \mathbf{R}^n \mid t = 1\}$ in $\mathbf{R}^{n+1} = \mathbf{R} \times \mathbf{R}^n$ and extend $\sigma$ to a homogeneous affine embedding $\sigma^+(t, x) = Ax + bt$. If we could associate with the latter mapping an $O(1)(mn)^{1/2}$-self-concordant barrier $H$ for $G^+ = (\sigma^+)^{-1}(S_m^+)$, we would take as the desired barrier for $G$ the restriction of $H$ onto $G = G^+ \bigcap \Pi$.

Of course, we are interested in the case of $m > n$ only, since $\sigma$ always generates the $m$-self-concordant barrier $F^+ \circ \sigma$ for $G$.

### The main result

**Theorem 5.5.1** *Let $\mathcal{A}$ be a linear homogeneous mapping from $\mathbf{R}^n$ into the space $S_m$ of symmetric $m \times m$ matrices, $m > n$, such that* $\operatorname{Ker} \mathcal{A} = \{0\}$ *and the image of $\mathcal{A}$ intersects the interior of the cone $S_m^+$ of positive-semidefinite symmetric $m \times m$ matrices. Let*

$$F^+(X) = -\ln \operatorname{Det}(X) : \operatorname{int} S_m \to \mathbf{R} \tag{5.5.1}$$

*be the standard $m$-logarithmically homogeneous self-concordant barrier for $S_m^+$. Let*

$$\Phi(x) = F^+(\mathcal{A}x) : \operatorname{int} K \to \mathbf{R}, \qquad K = \mathcal{A}^{-1}(S_m^+),$$

$$\Psi(x) = \ln \operatorname{Det}(\Phi''(x)) : \operatorname{int} K \to \mathbf{R}. \tag{5.5.2}$$

*Then, under an appropriate choice of absolute constants $O(1)$, the function $O(1)\Theta(x)$,*

$$\Theta(x) = \sqrt{m}\Psi(x) \tag{5.5.3}$$

*(the generalized volumetric barrier) is an $O(1)m^{1/2}n-$, and the function $O(1)\Xi(x)$,*

$$\Xi(x) = \rho\Psi(x) + \frac{1}{\rho}\Phi(x), \qquad \rho = \sqrt{\frac{m}{n}} \tag{5.5.4}$$

*(the generalized combined volumetric barrier) is an $O(1)\sqrt{nm}$-logarithmically homogeneous self-concordant barrier for the cone $K$.*

**Application example.** Let $f_i(x)$, $1 \le i \le m$ be convex quadratic forms on $\mathbf{R}^n$ and let $\operatorname{Rank}\{f_i\}$ be the rank of the Hessian of $f_i$. Assume also that the family $\{f_i\}$ satisfies the Slater condition: There exists $x \in \mathbf{R}^n$ such that $f_i(x) < 0$, $1 \le i \le m$. We know (Proposition 5.4.2) that under these assumptions the set

$$G = \{x \mid f_i(x) \le 0, \ 1 \le i \le m\}$$

admits an $m$-self-concordant "computable" barrier. Let us demonstrate that $G$ also admits a computable $O(1)\{n\sum_{i=1}^m[\operatorname{Rank}\{f_i\} + 1]\}^{1/2}$-self-concordant barrier. Note that, in the case of small $\operatorname{Rank}\{f_i\}/n$ and large $m$, the parameter of the latter barrier can be much better than that of the first barrier. Note also that Vaidya's combined volumetric barrier corresponds to the case of $\operatorname{Rank}\{f_i\} \equiv 0$.

A convex quadratic form $f$ of rank $k$, under an appropriate choice of an affine transformation $A_f : \mathbf{R}^n \to \mathbf{R}^{k+1}$, can be represented as

$$f(x) = \phi_k(A_f(x)), \qquad \phi_k(z_1, \ldots, z_k, t) = \sum_{j=1}^k z_j^2 - t.$$

The set $\{z \in \mathbf{R}^{k+1} \mid \phi_k(z) \le 0\}$ can be represented as the inverse image of $S_{k+1}^+$ under the affine mapping $T_k$ as follows:

$$T_k(z) = \begin{pmatrix} I_k & z \\ z^T & t \end{pmatrix}.$$

Thus, there exists an affine mapping $B_f$ ($= T_{\mathrm{Rank}\, f} \circ A_f$) from $\mathbf{R}^n$ into $S_{k+1}$ such that

$$\{x \in \mathbf{R}^n \mid f(x) \le 0\} = B_f^{-1}(S_{k+1}^+).$$

Note that a point $x$ with $f(x) < 0$ under the mapping $B_f$ corresponds to an interior point of $S_{k+1}^+$.

Now, given the family $\{f_i\}$, we can define an affine mapping $T$ from $\mathbf{R}^n$ into $S_M$, where $M = \sum_{i=1}^m \{\mathrm{Rank}\,\{f_i\} + 1\}$, as follows: $T(x)$ is the block-diagonal matrix with $m$ diagonal blocks $B_{f_1}(x), \ldots, B_{f_m}(x)$. Now the set $G$ can be represented as $T^{-1}(S_M^+)$. Note that the Slater condition implies that $T(\mathbf{R}^n)$ intersects $\mathrm{int}\, S_M^+$.

Now we are precisely in the situation of this section: $G$ is represented as an inverse image of $S_M^+$ under the affine mapping $T$, and this representation implies the corresponding combined volumetric barrier with the desired value of the parameter.

**Proof of Theorem 5.5.1.** $1^0$. Let us start with some notation. In what follows, lowercase italics denote elements of $\mathbf{R}^n$, uppercase italics denote $m \times m$ and $n \times n$ matrices, handwritten letters denote linear transformations of $S_m$, and lowercase Greek letters are used for reals. All $O(1)$ are absolute constants (and, in particular, do not depend on $n$ and $m$); if it is necessary to distinguish between these constants, we use subscripts (so that, say, $O_5(1)$ denotes the fifth absolute constant involved in our reasoning).

If $A, B$ are $m \times m$ matrices, not necessarily symmetric, then $\mathcal{S}_A$ denotes the linear mapping

$$X \to AXA^T : S_m \to S_m,$$

$\mathcal{T}_A$ denotes the mapping

$$X \to AX + XA^T : S_m \to S_m,$$

and $\mathcal{P}_{A,B}$ denotes the mapping

$$X \to AXB^T + BXA^T : S_m \to S_m.$$

$S_m$ is regarded as an Euclidean space with the standard inner product $\langle X, Y\rangle = \mathrm{Tr}\{XY\}$; $\| X \|_2 = \mathrm{Tr}^{1/2}\{X^2\}$ denotes the corresponding norm of $X \in S_m$. $\mathbf{R}^n$ is provided with the usual inner product $(x, y) = x^T y$.

$2^0$. For $y \in \mathrm{int}\, K$, $h \in \mathbf{R}^n$ we have

$$\begin{aligned} D\Phi(y)[h] &= -\,\mathrm{Tr}\{(\mathcal{A}y)^{-1}(\mathcal{A}h)\}, \\ D^2\Phi(y)[h,h] &= \mathrm{Tr}\{[(\mathcal{A}y)^{-1}(\mathcal{A}h)]^2\} = (\mathcal{A}^T\mathcal{S}_{(\mathcal{A}y)^{-1}}\mathcal{A}h, h), \\ D^3\Phi(y)[h,h,h] &= -2\,\mathrm{Tr}\{[(\mathcal{A}y)^{-1}(\mathcal{A}h)]^3\} = -2\,\mathrm{Tr}\{[\mathcal{S}_{(\mathcal{A}y)^{-1/2}}\mathcal{A}h]^3\}. \end{aligned} \tag{5.5.5}$$

In particular, $\Phi''(y)$ is nondegenerate (since $\mathrm{Ker}\,\mathcal{A} = \{0\}$), so that $\Psi$ and $\Phi$ are $\mathrm{C}^\infty$-smooth on $\mathrm{int}\, K$.

$3^0$. From (5.5.5), it follows that $D^2\Phi(ty)[h,h] = t^{-2}D^2\Phi(y)[h,h]$, $t > 0$, so that $\Psi$ is $2n$-logarithmically homogeneous (and $\Phi$ is clearly $m$-logarithmically homogeneous). It follows that $\Theta$ is $2n\sqrt{m}$-logarithmically homogeneous and $\Xi$ is $3\sqrt{nm}$-logarithmically homogeneous. Let us prove that, if $y_i \in \operatorname{int} K$ and $\lim_{i\to\infty} y_i = y$ is a boundary point of $K$, then $\Theta(y_i)$ and $\Xi(y_i)$ tend to $\infty$ as $i \to \infty$. Since $\Phi$ is a logarithmically homogeneous self-concordant barrier for $K$ (Proposition 5.1.1 combined with Proposition 5.4.6), $\Phi(y_i) \to \infty$, $i \to \infty$, so that it suffices to verify that the same property holds for $\Psi$.

Since $\Phi$ is a self-concordant barrier for $K$, the ellipsoids

$$W_i = \{y \in \mathbf{R}^n \mid D^2\Phi(y_i)[y - y_i, y - y_i] \le 1\}$$

are contained in $K$ (Proposition 2.3.2(i.1)); since $\operatorname{Ker} \mathcal{A} = \{0\}$, $K$ is a pointed cone, so, from the boundedness of $\{y_i\}$, it follows that all the ellipsoids $W_i$ are uniformly in $i$ bounded. Since $y \notin \operatorname{int} K$, $y$ does not belong to the interior of every $W_i$ and since the centers of $W_i$ converge to $y$, that means that the (smallest) thickness of $W_i$ tends to 0 as $i \to \infty$. The latter fact combined with the uniform boundedness of $W_i$ means that $\operatorname{mes}_n W_i \to 0$, $i \to \infty$, so that

$$\Psi(y_i) = \ln \operatorname{Det} \Phi''(y_i) = -2\ln(\operatorname{mes}_n W_i) + \text{ const } \to \infty, \qquad i \to \infty,$$

as claimed.

The logarithmical homogeneity of $\Theta$ and $\Xi$ and the fact that these functions tend to $\infty$ as the argument belonging to $\operatorname{int} K$ approaches a boundary point of $K$ signify that all we need to prove the theorem is to demonstrate that, for appropriate absolute constants $O_1(1)$ and $O_2(1)$ and all $y \in \operatorname{int} K$, $h \in \mathbf{R}^n$, we have

$$D^2\Psi(y)[h,h] \ge 0 \tag{5.5.6}$$

(this would imply convexity of $\Theta$ and $\Xi$) and

$$| D^3\Theta(y)[h,h,h] | \le O_1(1)\{D^2\Theta(y)[h,h]\}^{3/2}, \tag{5.5.7}$$

$$| D^3\Xi(y)[h,h,h] | \le O_2(1)\{D^2\Xi(y)[h,h]\}^{3/2}, \tag{5.5.8}$$

which would mean that the functions are self-concordant with the parameters of order of 1.

$4^0$. Let us fix $x \in \operatorname{int} K$. Set

$$\bar{\mathcal{A}} = \mathcal{S}_{(\mathcal{A}x)^{-1/2}} \circ \mathcal{A} \circ (\Phi''(x))^{-1/2} : \mathbf{R}^n \to S_m,$$

$$\bar{x} = (\Phi''(x))^{1/2} x.$$

Let

$$\bar{\Phi}(y) = F^+(\bar{\mathcal{A}}y) : \operatorname{int}(\bar{\mathcal{A}})^{-1}(S_m^+) \to \mathbf{R},$$

$$\bar{\Psi}(y) = \ln \operatorname{Det} \bar{\Phi}''(y)$$

and let $\bar{\Theta}$, $\bar{\Xi}$ be the functions associated with $\bar{\Phi}$, $\bar{\Psi}$ in the same way as $\Theta$, $\Xi$ are associated with $\Phi$, $\Psi$. Then we clearly have

$$\bar{\Phi}(y) = \Phi(Py) - \ln \operatorname{Det}(\mathcal{A}x), \qquad P = (\Phi''(x))^{-1/2},$$

whence

$$\bar{\Psi}(y) = \Psi(Py) + \text{ const}(x),$$

so that

$$\bar{\Theta}(y) = \Theta(Py) + \text{ const}_1(x), \qquad \bar{\Xi}(y) = \Xi(Py) + \text{ const}_2(x).$$

Since $\Phi$ is a self-concordant barrier for a pointed cone, $P$ is nondegenerate. We see that establishing (5.5.7), (5.5.8) is the same as establishing the relations obtained from (5.5.7), (5.5.8) by replacing $\Theta \leftarrow \bar{\Theta}$, $\Xi \leftarrow \bar{\Xi}$, $x \leftarrow \bar{x}$ (of course, the constants $O_1(1), O_2(1)$ should not depend on $\bar{x}$). Note that

$$\bar{\mathcal{A}}\bar{x} = I_m, \qquad \bar{\Phi}''(\bar{x}) = I_n.$$

To simplify notation, let us forget about the initial data without bars and let us omit bars in the notation for our new data (so that, say, our present $\mathcal{A}$ is the old $\bar{\mathcal{A}}$); this should not cause any difficulties, since we are no longer interested in the old data.

Thus, it suffices to establish (5.5.7), (5.5.8) under the additional assumption that

$$\mathcal{A}x = I_m, \qquad \Phi''(x) = I_n. \tag{5.5.9}$$

Note that, in view of (5.5.5), the latter relation means that

$$\mathcal{A}^T\mathcal{A} = I_n. \tag{5.5.10}$$

$5^0$. Let us fix $h \in \mathbf{R}^n$ and let $H = \mathcal{A}h$. Then

$$H = \sum_{i=1}^{m} \lambda_i f_i f_i^T,$$

where $\{f_i\}$ is an orthonormal basis of eigenvectors of $H$, and $\lambda_i$ are the corresponding eigenvalues. The matrices

$$E_{ij} = \begin{cases} f_i f_i^T, & i = j, \\ (f_i f_j^T + f_j f_i^T)/\sqrt{2}, & i \neq j, \end{cases}$$

$1 \leq j \leq i \leq m$ form an orthonormal basis in $S_m$, and these matrices are eigenvectors of $\mathcal{S}_{H^k}$, $\mathcal{T}_{H^k}$, so that these operators on $S_m$ commutate and are symmetric. For an operator $\mathcal{B}$ on $S_m$ that has as its eigenvectors the matrices $E_{ij}$, let $\lambda_{ij}(\mathcal{B})$ denote the corresponding eigenvalues.

Note that

$$\lambda_{ij}(\mathcal{T}_{H^2}) = \lambda_i^2 + \lambda_j^2,$$

so that $\mathcal{T}_{H^2}$ is positive-semidefinite.

$6^0$. Let us compute the derivatives of $\Phi$ and $\Psi$ at $x$. Denote

$$Q(y) = \Phi''(y); \tag{5.5.11}$$

then (5.5.5) can be rewritten as

$$Q(y) = \mathcal{A}^T \mathcal{S}_{P(y)} \mathcal{A}, \tag{5.5.12}$$

where

$$P(y) = (\mathcal{A}y)^{-1}; \tag{5.5.13}$$

from (5.5.12), it follows that

$$DQ(y)[h] = \mathcal{A}^T \mathcal{P}_{P(y),DP(y)[h]} \mathcal{A}, \tag{5.5.14}$$

$$D^2Q(y)[h,h] = \mathcal{A}^T \{2\mathcal{S}_{DP(y)[h]} + \mathcal{P}_{P(y),D^2P(y)[h,h]}\} \mathcal{A}, \tag{5.5.15}$$

$$D^3Q(y)[h,h,h] = \mathcal{A}^T \{3\mathcal{P}_{DP(y)[h],D^2P(y)[h,h]} + \mathcal{P}_{P(y),D^3P(y)[h,h,h]}\} \mathcal{A}. \tag{5.5.16}$$

Since $\Psi(y) = \ln \operatorname{Det} Q(y)$, we have

$$D\Psi(y)[h] = \operatorname{Tr}\{Q^{-1}(y) DQ(y)[h]\}, \tag{5.5.17}$$

$$\begin{aligned} D^2\Psi(y)[h,h] = &- \operatorname{Tr}\{[Q^{-1}(y)DQ(y)[h]]^2\} \\ &+ \operatorname{Tr}\{Q^{-1}(y)D^2Q(y)[h,h]\}, \end{aligned} \tag{5.5.18}$$

$$D^3\Psi(y)[h,h,h] = 2\operatorname{Tr}\{[Q^{-1}(y)DQ(y)[h]]^3\} \tag{5.5.19}$$

$$\begin{aligned} &- 3\operatorname{Tr}\{Q^{-1}(y)DQ(y)[h]Q^{-1}(y)D^2Q(y)[h,h]\} \\ &+ \operatorname{Tr}\{Q^{-1}(y)D^3Q(y)[h,h,h]\}. \end{aligned} \tag{5.5.20}$$

Last, from (5.5.13), it follows that

$$DP(y)[h] = -(\mathcal{A}y)^{-1}\mathcal{A}h(\mathcal{A}y)^{-1}, \tag{5.5.21}$$

$$D^2P(y)[h,h] = 2((\mathcal{A}y)^{-1}\mathcal{A}h)^2(\mathcal{A}y)^{-1}, \tag{5.5.22}$$

$$D^3P(y)[h,h,h] = -6((\mathcal{A}y)^{-1}\mathcal{A}h)^3(\mathcal{A}y)^{-1}. \tag{5.5.23}$$

Relations (5.5.21)–(5.5.23), in view of (5.5.9) imply that

$$P(x) = I_m, \tag{5.5.24}$$

$$DP(x)[h] = -H, \tag{5.5.25}$$

$$D^2P(x)[h,h] = 2H^2, \tag{5.5.26}$$

$$D^3P(x)[h,h,h] = -6H^3. \tag{5.5.27}$$

From (5.5.14)–(5.5.16), it follows now that

$$Q(x) = I_n, \tag{5.5.28}$$

$$DQ(x)[h] = -\mathcal{A}^T \mathcal{T}_H \mathcal{A}, \tag{5.5.29}$$

$$D^2Q(x)[h,h] = \mathcal{A}^T\{2\mathcal{S}_H + 2\mathcal{T}_{H^2}\}\mathcal{A}, \tag{5.5.30}$$

$$D^3Q(x) = -6\mathcal{A}^T\{\mathcal{S}_H\mathcal{T}_H + \mathcal{T}_{H^3}\}\mathcal{A}. \tag{5.5.31}$$

Finally, (5.5.28)–(5.5.31) and (5.5.18)–(5.5.20) imply that

$$\alpha \equiv D^2\Psi(x)[h,h] = -\operatorname{Tr}\{\mathcal{A}^T\mathcal{T}_H\mathcal{A}\mathcal{A}^T\mathcal{T}_H\mathcal{A}\} + \operatorname{Tr}\{\mathcal{A}^T\{2\mathcal{S}_H + 2\mathcal{T}_{H^2}\}\mathcal{A}\}.$$

In view of (5.5.10), $\mathcal{A}\mathcal{A}^T$ is an orthoprojector in $S_m$, and the operator $\mathcal{T}_H$ is symmetric on $S_m$ (see $5^0$), so that the operator $\mathcal{A}^T(\mathcal{T}_H)^2\mathcal{A} - \mathcal{A}^T\mathcal{T}_H\mathcal{A}\mathcal{A}^T\mathcal{T}_H\mathcal{A}$ in $\mathbf{R}^n$ is symmetric positive-semidefinite; it follows that

$$\operatorname{Tr}\{\mathcal{A}^T\mathcal{T}_H\mathcal{A}\mathcal{A}^T\mathcal{T}_H\mathcal{A}\} \le \operatorname{Tr}\{\mathcal{A}^T(\mathcal{T}_H)^2\mathcal{A}\} = \operatorname{Tr}\{\mathcal{A}^T\{\mathcal{T}_{H^2} + 2\mathcal{S}_H\}\mathcal{A}\},$$

so that

$$\alpha \ge \operatorname{Tr}\{\mathcal{A}^T\mathcal{T}_{H^2}\mathcal{A}\} \equiv \beta. \tag{5.5.32}$$

Note that $\beta \ge 0$ (see $5^0$); in particular, $\Psi$ is convex. Furthermore, (5.5.18)–(5.5.20) and (5.5.28)–(5.5.31) imply that

$$\begin{aligned} D^3\Psi(y)[h,h,h] = &-2\operatorname{Tr}\{[\mathcal{A}^T\mathcal{T}_H\mathcal{A}]^3\} + 6\operatorname{Tr}\{(\mathcal{A}^T\mathcal{T}_H\mathcal{A})(\mathcal{A}^T\{\mathcal{S}_H + \mathcal{T}_{H^2}\}\mathcal{A})\} \\ &-6\operatorname{Tr}\{\mathcal{A}^T\{\mathcal{S}_H\mathcal{T}_H + \mathcal{T}_{H^3}\}\mathcal{A}\}. \end{aligned} \tag{5.5.33}$$

We also have (see (5.5.5), (5.5.9), (5.5.10))

$$D^2\Phi(x)[h,h] = \operatorname{Tr}\{H^2\}, \qquad D^3\Phi(x([h,h,h] = -2\operatorname{Tr}\{H^3\}. \tag{5.5.34}$$

$7^0$. From (5.5.33), it follows that

$$| D^3\Psi(y)[h,h,h] | \le 6 \sum_{i=1}^{5} \delta_i, \tag{5.5.35}$$

$$\delta_1 = | \operatorname{Tr}\{A^3\} |, \qquad A = \mathcal{A}^T\mathcal{T}_H\mathcal{A}, \tag{5.5.36}$$

$$\delta_2 = | \operatorname{Tr}\{AB\} |, \qquad B = \mathcal{A}^T\mathcal{S}_H\mathcal{A}, \tag{5.5.37}$$

$$\delta_3 = | \operatorname{Tr}\{AC\} |, \qquad C = \mathcal{A}^T\mathcal{T}_{H^2}\mathcal{A}, \tag{5.5.38}$$

$$\delta_4 = | \operatorname{Tr}\{\mathcal{A}^T\mathcal{S}_H\mathcal{T}_H\mathcal{A}\} |, \tag{5.5.39}$$

$$\delta_5 = | \operatorname{Tr}\{\mathcal{A}^T\mathcal{T}_{H^3}\mathcal{A}\} | . \tag{5.5.40}$$

Note that $A$, $B$, $C$ are symmetric $n \times n$ matrices (see $5^0$).

$7^0.1$. Let us verify that

$$(0 \le) \operatorname{Tr}\{A^2\} \le 2\beta. \tag{5.5.41}$$

Indeed, as we have already mentioned,

$$A^2 = \mathcal{A}^T\mathcal{T}_H\mathcal{A}\mathcal{A}^T\mathcal{T}_H\mathcal{A} \le \mathcal{A}^T(\mathcal{T}_H)^2\mathcal{A} \tag{5.5.42}$$

(the inequalities involving symmetric matrices are, as usual, understood in the operator sense: $V \geq 0$ means that $V$ is positive-semidefinite). Furthermore, $\lambda_{ij}(\mathcal{T}_H) = \lambda_i + \lambda_j$, so that $\lambda_{ij}((\mathcal{T}_H)^2) = (\lambda_i + \lambda_j)^2 \leq 2(\lambda_i^2 + \lambda_j^2) = 2\lambda_{ij}(\mathcal{T}_{H^2})$, whence $0 \leq (\mathcal{T}_H)^2 \leq 2\mathcal{T}_{H^2}$, which, combined with (5.5.42), immediately implies (5.5.41).

$7^0$.2. From (5.5.41) and the inequality $|\operatorname{Tr}\{Q^3\}| \leq \operatorname{Tr}^{3/2}\{Q^2\}$ (which clearly holds for an arbitrary symmetric matrix), it immediately follows that

$$\delta_1 \leq 3\beta^{3/2}. \tag{5.5.43}$$

$7^0$.3. We have $\lambda_{ij}(\mathcal{S}_H) = \lambda_i\lambda_j$, whence $-\mathcal{T}_{H^2} \leq 2\mathcal{S}_H \leq \mathcal{T}_{H^2}$, so that

$$-\mathcal{A}^T\mathcal{T}_{H^2}\mathcal{A} \leq 2\mathcal{A}^T\mathcal{S}_H\mathcal{A} \leq \mathcal{A}^T\mathcal{T}_{H^2}\mathcal{A} = C. \tag{5.5.44}$$

Let $\{g_l\}$ be the orthonormal basis in $\mathbf{R}^n$ formed by eigenvectors of the symmetric matrix $A$: $Ag_l = \mu_l g_l$ and let $B_{lq}$, $C_{lq}$ be the elements of $B$, $C$, respectively, in this basis. Then $|\operatorname{Tr}\{AB\}| = |\sum_{l=1}^n \mu_l B_{ll}|$, and, in view of (5.5.44), $|B_{ll}| \leq C_{ll}$, whence $\sum_{l=1}^n |B_{ll}| \leq \operatorname{Tr}\{C\} = \beta$. It follows that $|\sum_{l=1}^n \mu_l B_{ll}| \leq \beta \max_l |\mu_l| \leq \beta \operatorname{Tr}^{1/2}\{A^2\} \leq 2^{1/2}\beta^{3/2}$ (see (5.5.41)). Thus,

$$\delta_2 \leq 2\beta^{3/2}. \tag{5.5.45}$$

$7^0$.4. Now, as in $7^0$.3, we have

$$|\operatorname{Tr}\{AC\}| = \left|\sum_{l=1}^n \mu_l C_{ll}\right| \leq \beta \max_l |\mu_l| \leq \beta \operatorname{Tr}^{1/2}\{A^2\} \leq 2^{1/2}\beta^{3/2},$$

so that

$$\delta_3 \leq 2\beta^{3/2}. \tag{5.5.46}$$

$7^0$.5. Now let us evaluate $\delta_4$ and $\delta_5$. Let $X^{(i)} \in S_m$ denote the $i$th column of $\mathcal{A}$: $X^{(i)} = \mathcal{A}e_i$, where $\{e_i\}_{i=1}^n$ is the standard orthonormal basis in $\mathbf{R}^n$. Let $X_{jk}^{(i)} = (f_j)^T X^{(i)} f_k$. Note that

(5.5.47)
$$\begin{aligned}
\delta_4 &= |\operatorname{Tr}\{\mathcal{A}^T\mathcal{S}_H\mathcal{T}_H\mathcal{A}\}| = \left|\sum_{i=1}^n (\mathcal{A}^T\mathcal{S}_H\mathcal{T}_H\mathcal{A}e_i, e_i)\right| = \left|\sum_{i=1}^n \langle \mathcal{S}_H\mathcal{T}_H\mathcal{A}e_i, \mathcal{A}e_i\rangle\right| \\
&= \left|\sum_{i=1}^n \operatorname{Tr}\{(\mathcal{S}_H\mathcal{T}_H X^{(i)})X^{(i)}\}\right| = \left|\sum_{i=1}^n \operatorname{Tr}\{(HX^{(i)}H^2 + H^2X^{(i)}H)X^{(i)}\}\right| \\
&= 2\left|\sum_{i=1}^n \operatorname{Tr}\{HX^{(i)}H^2X^{(i)}\}\right| = 2\left|\sum_{i=1}^n\sum_{j=1}^m (f_j)^T HX^{(i)}H^2X^{(i)}f_j\right| \\
&= 2\left|\sum_{i=1}^n\sum_{j=1}^m\sum_{k=1}^m \lambda_j\lambda_k^2 (X_{jk}^{(i)})^2\right|,
\end{aligned}$$

while

(5.5.48)
$$\begin{aligned}\delta_5 &= \mid \mathrm{Tr}\{\mathcal{A}^T \mathcal{T}_{H^3} \mathcal{A}\} \mid = \left|\sum_{i=1}^{n} (\mathcal{A}^T \mathcal{T}_{H^3} \mathcal{A} e_i, e_i)\right| = \left|\sum_{i=1}^{n} \langle \mathcal{T}_{H^3} \mathcal{A} e_i, \mathcal{A} e_i\rangle\right| \\ &= \left|\sum_{i=1}^{n} \mathrm{Tr}\{(\mathcal{T}_{H^3} X^{(i)}) X^{(i)}\}\right| = \left|\sum_{i=1}^{n} \mathrm{Tr}\{(X^{(i)} H^3 + H^3 X^{(i)}) X^{(i)}\}\right| \\ &= 2\left|\sum_{i=1}^{n} \mathrm{Tr}\{X^{(i)} H^3 X^{(i)}\}\right| = 2\left|\sum_{i=1}^{n}\sum_{j=1}^{m} (f_j)^T X^{(i)} H^3 X^{(i)} f_j\right| \\ &= 2\left|\sum_{i=1}^{n}\sum_{j=1}^{m}\sum_{k=1}^{m} \lambda_k^3 (X_{jk}^{(i)})^2\right|.\end{aligned}$$

Let us denote

$$\delta = \sum_{i=1}^{n}\sum_{j=1}^{m}\sum_{k=1}^{m} \mid \lambda_k^3 \mid (X_{jk}^{(i)})^2; \tag{5.5.49}$$

then

$$\delta_5 \leq 2\delta. \tag{5.5.50}$$

Furthermore, from (5.5.47), it follows that

$$\delta_4 \leq 2\delta',$$
$$\begin{aligned}\delta' = \sum_{i=1}^{n}\sum_{j=1}^{m}\sum_{k=1}^{m} \mid \lambda_j \mid\mid \lambda_k \mid^2 (X_{jk}^{(i)})^2 &= \sum_{i=1}^{n}\sum_{j=1}^{m}\sum_{k=1}^{m} (\mid \lambda_j \mid\mid \lambda_k \mid) \mid \lambda_k \mid (X_{jk}^{(i)})^2 \\ &\leq \sum_{i=1}^{n}\sum_{j=1}^{m}\sum_{k=1}^{m} \frac{\lambda_j^2 + \lambda_k^2}{2} \mid \lambda_k \mid (X_{jk}^{(i)})^2 = \frac{\delta}{2} + \frac{\delta'}{2}\end{aligned}$$

(we considered that $X^{(i)}$ are symmetric). Thus, $\delta' \leq \delta$, and

$$\delta_4 \leq 2\delta. \tag{5.5.51}$$

Note also that

$$
\begin{aligned}
\beta = \mathrm{Tr}\{\mathcal{A}^T \mathcal{T}_{H^2} \mathcal{A}\} &= \sum_{i=1}^n (\mathcal{A}^T \mathcal{T}_{H^2} \mathcal{A} e_i, e_i) = \sum_{i=1}^n \langle \mathcal{T}_{H^2} \mathcal{A} e_i, \mathcal{A} e_i \rangle \\
&= \sum_{i=1}^n \mathrm{Tr}\{(\mathcal{T}_{H^2} X^{(i)}) X^{(i)}\} \\
&= \sum_{i=1}^n \mathrm{Tr}\{(X^{(i)} H^2 + H^2 X^{(i)}) X^{(i)}\} \\
&= 2 \sum_{i=1}^n \mathrm{Tr}\{X^{(i)} H^2 X^{(i)}\} \\
&= 2 \sum_{i=1}^n \sum_{j=1}^m (f_j)^T X^{(i)} H^2 X^{(i)} f_j \\
&= 2 \sum_{i=1}^n \sum_{j=1}^m \sum_{k=1}^m \lambda_k^2 (X_{jk}^{(i)})^2 \equiv 2\gamma.
\end{aligned}
\tag{5.5.52}
$$

Denote

$$
\sigma_k = \left\{ \sum_{i=1}^n \sum_{j=1}^m (X_{jk}^{(i)})^2 \right\}^{1/2}. \tag{5.5.53}
$$

Then (5.5.50)–(5.5.52) can be rewritten as

$$
\begin{aligned}
&\delta_4,\ \delta_5 \le 2\delta, \qquad \delta \equiv \sum_{k=1}^m \tau_k^3 \sigma_k^2, \qquad \tau_k = \mid \lambda_k \mid; \\
&\beta = 2\gamma, \qquad \gamma = \sum_{k=1}^m \tau_k^2 \sigma_k^2.
\end{aligned}
\tag{5.5.54}
$$

$8^0$. Let us prove that

$$
\begin{aligned}
&\sum_{k=1}^m \tau_k^2 = h^T h, \\
&\tau_k \le \left( \sum_{j=1}^m \tau_j^2 \right)^{1/2} \sigma_k.
\end{aligned}
\tag{5.5.55}
$$

Indeed, in view of (5.5.10), we have $\| \mathcal{A} z \|_2 = (z^T z)^{1/2}$, $z \in \mathbf{R}^n$, whence

$$
\sum_{k=1}^m \tau_k^2 = \| H \|_2^2 = \| \mathcal{A} h \|_2^2 = h^T h;
$$

at the same time,

$$
\begin{aligned}
\tau_k &= \left| \left\langle H, f_k f_k^T \right\rangle \right| = \left| \left\langle \sum_{i=1}^n h_i X^{(i)}, f_k f_k^T \right\rangle \right| \\
&= \left| \sum_{i=1}^n h_i X_{kk}^{(i)} \right| \le \left\{ \sum_{i=1}^n h_i^2 \right\}^{1/2} \left\{ \sum_{i=1}^n (X_{kk}^{(i)})^2 \right\}^{1/2} \le (h^T h)^{1/2} \sigma_k,
\end{aligned}
$$

and (5.5.55) follows.

$9^0$. The summary of (5.5.40), (5.5.43)–(5.5.46), (5.5.54), (5.5.44), and (5.5.34) is as follows. We have established that, for some nonnegative $\tau_k$ and $\sigma_k$ such that

$$\tau_k \leq \left(\sum_{j=1}^{m} \tau_j^2\right)^{1/2} \sigma_k,\ 1 \leq k \leq m, \tag{5.5.56}$$

we have

$$D^2\Psi(x)[h,h] \geq \sum_{k=1}^{m} \tau_k^2\sigma_k^2, \tag{5.5.57}$$

$$D^2\Phi(x)[h,h] = \sum_{k=1}^{m} \tau_k^2, \tag{5.5.58}$$

$$\mid D^3\Psi(x)[h,h,h] \mid \leq 100 \left\{ \left\{ \sum_{k=1}^{m} \tau_k^2\sigma_k^2 \right\}^{3/2} + \sum_{k=1}^{m} \tau_k^3\sigma_k^2 \right\}, \tag{5.5.59}$$

$$\mid D^3\Phi(x)[h,h,h] \mid \leq 100 \sum_{k=1}^{m} \tau_k^3 \tag{5.5.60}$$

(of course, the factor 100 can be reduced, but now we are not interested in the values of absolute constants). We should prove the pair of inequalities (5.5.7), (5.5.8), i.e., the relations

$$\mid \sqrt{m} D^3\Psi(x)[h,h,h] \mid \leq O_1(1)\{\sqrt{m}D^2\Psi(x)[h,h]\}^{3/2}, \tag{5.5.61}$$

$$\begin{aligned} &\mid \rho D^3\Psi(x)[h,h,h] + \rho^{-1}D^3\Phi(x)[h,h,h] \mid \\ &\leq O_2(1)\{\rho D^2\Psi(x)[h,h] + \rho^{-1}D^2\Phi(x)[h,h]\}^{3/2}, \qquad \rho = \sqrt{\frac{m}{n}}; \end{aligned} \tag{5.5.62}$$

where $O_1(1)$ and $O_2(1)$ should be absolute constants (i.e., they cannot depend on any data involved into (5.5.56)–(5.5.60)).

Now we derive (5.5.61), (5.5.62) from (5.5.56)–(5.5.60); the reader can disregard all our preceding considerations.

$10^0$. It is easily seen that, to derive (5.5.61), (5.5.62) from (5.5.56)–(5.5.60), it suffices to verify that, for all nonnegative $\tau_k$ and $\sigma_k$ satisfying (5.5.56), the following inequalities hold:

$$\left\{ \sum_{k=1}^{m} \tau_k^3\sigma_k^2 \right\}^2 \leq O_3(1)\sqrt{m} \left\{ \sum_{k=1}^{m} \tau_k^2\sigma_k^2 \right\}^3, \tag{5.5.63}$$

$$\left\{ \sum_{k=1}^{m} [\tau_k^3\sigma_k^2 + (n/m)\tau_k^3] \right\}^2 \leq O_3(1)\sqrt{\frac{m}{n}} \left\{ \sum_{k=1}^{m} [\tau_k^2\sigma_k^2 + (n/m)\tau_k^2] \right\}^3, \tag{5.5.64}$$

where $O_3(1)$ should be an absolute constant.

Of course, it suffices to consider the case when all $\tau_k$ are positive; besides this, we can assume that all these reals differ from each other. In view of the homogeneity of (5.5.64) and (5.5.56) with respect to $\{\tau_k\}$, it suffices to consider the case when

$$\sum_{k=1}^{m} \tau_k^2 = 1, \tag{5.5.65}$$

so that (see (5.5.56))

$$\sigma_k \geq \tau_k, \qquad 1 \leq k \leq m. \tag{5.5.66}$$

Let us fix $\theta \in [0,1]$ and consider the function

$$\phi_\theta(u) = \frac{\left\{\sum\limits_{k=1}^{m} [\tau_k^3(\tau_k^2 + u_k) + \theta\tau_k^3]\right\}^2}{\left\{\theta + \sum\limits_{k=1}^{m} \tau_k^2(\tau_k^2 + u_k)\right\}^3} : \mathbf{R}_+^m \to \mathbf{R}. \tag{5.5.67}$$

Proving (5.5.63), (5.5.64) is the same as proving that the values of the functions $\phi_0(\cdot)$ and $\phi_{n/m}(\cdot)$ at the point $u^* = (\sigma_1^2 - \tau_1^2, \ldots, \sigma_m^2 - \tau_m^2)^T \in \mathbf{R}_+^m$ (the latter inclusion follows from (5.5.66)) satisfy the relations

$$\phi_0(u^*) \leq O_3(1)\sqrt{m} \tag{5.5.68}$$

and

$$\phi_{n/m}(u^*) \leq O_3(1)\sqrt{\frac{m}{n}}. \tag{5.5.69}$$

To prove these relations, let us evaluate the maximal value of $\phi_\theta(\cdot)$ over $\mathbf{R}_+^m$. First, $\phi_\theta$ is continuous on the latter set and clearly tends to 0 as $\| u \|_2$ tends to $\infty$; thus, $\phi_\theta$ attains its maximum over $\mathbf{R}_+^m$ at certain $u^+ = u^+(\theta)$. From the necessary optimality conditions, we conclude that the following relations hold:

$$\partial\phi_\theta(u^+)/\partial u_k \leq 0, \quad 1 \leq k \leq m; \qquad \partial\phi_\theta(u^+)/\partial u_k < 0 \ \Rightarrow\ u_k^+ = 0. \tag{5.5.70}$$

We have

$$\partial\phi_\theta(u^+)/\partial u_k = \omega\tau_k^2\{\tau_k - \Omega\}, \qquad 1 \leq k \leq m, \tag{5.5.71}$$

where $\omega > 0$ and

$$\Omega = \frac{3}{2} \cdot \frac{\sum\limits_{k=1}^{m} [\tau_k^3(\tau_k^2 + u_k^+) + \theta\tau_k^3]}{\theta + \sum\limits_{k=1}^{m} [\tau_k^2(\tau_k^2 + u_k^+)]}. \tag{5.5.72}$$

According to our assumption that all $\tau_k$ are positive and differ from each other, we conclude from (5.5.70)–(5.5.72) that the only possible cases are the following:

I. $u^+ = 0$;

II. All but one $u_k^+$ are zeros, the index of the remaining coordinate of $u^+$ is the index of the largest $\tau_k$, and the latter, i.e., the largest, $\tau_k$ equals to $\Omega$.

$10^0$.1. Let us start with case II. Without loss of generality, let $\tau_1$ be the largest of $\{\tau_k\}$. It is easily seen that the equality $\tau_1 = \Omega$ implies the relation

$$\phi_\theta(u^+) = \frac{4}{9} \cdot \frac{\tau_1^2}{\theta + \tau_1^4 + \tau_1^2 u_1^+ + \sum\limits_{k=2}^m \tau_k^4} \le \frac{\tau_1^2}{\theta + \tau_1^4 + \sum\limits_{k=2}^m \tau_k^4}. \tag{5.5.73}$$

In view of (5.5.65), we have $\sum_{k=2}^m \tau_k^4 \ge (1-\tau_1^2)^2/(m-1)$. Thus, (5.5.73) implies that

$$\phi_\theta(u^+) \le \frac{\tau_1^2}{\theta + \tau_1^4 + \frac{(1-\tau_1^2)^2}{m-1}}. \tag{5.5.74}$$

$10^0$.2. Now assume that I is the case, so that

$$\phi_\theta(u^+) = \frac{\left\{\sum\limits_{k=1}^m [\tau_k^5 + \theta\tau_k^3]\right\}^2}{\{\theta + \sum\limits_{k=1}^m \tau_k^4\}^3}. \tag{5.5.75}$$

Without loss of generality, let $\tau_1$ be the maximum of $\{\tau_k\}$. Then

$$\left\{\sum_{k=1}^m [\tau_k^5 + \theta\tau_k^3]\right\} \le \tau_1 \left\{\sum_{k=1}^m [\tau_k^4 + \theta\tau_k^2]\right\} = \tau_1 \left\{\sum_{k=1}^m \tau_k^4 + \theta\right\},$$

so that $\phi_\theta(u^+) \le {\tau_1}^2\{\sum_{k=1}^m \tau_k^4 + \theta\}^{-1}$. Besides this, as in $10^0$.1, we have

$$\left\{\sum_{k=1}^m \tau_k^4 + \theta\right\} \ge \theta + {\tau_1}^4 + \frac{(1-{\tau_1}^2)^2}{m-1},$$

and we see that (5.5.74) holds in case I same as it holds in case II.

$10^0$.3. It remains to derive from (5.5.74) the desired relations (5.5.68), (5.5.69). Maximizing the right-hand side of (5.5.74) in $t \equiv \tau_1^2$, we come to

$$\max_{\tau_1} \frac{{\tau_1}^2}{\theta + \tau_1^4 + \frac{(1-\tau_1^2)^2}{m-1}} = \frac{\sqrt{m(1+(m-1)\theta)}+1}{2(m\theta+1)},$$

which, combined with (5.5.74), immediately implies (5.5.68), (5.5.69). □

*This page intentionally left blank*

# Chapter 6

# Applications in convex optimization

In this chapter, we present applications of the interior-point machinery to concrete convex optimization problems of appropriate analytical structure: quadratically constrained convex quadratic programming, geometrical programming, approximation in $\| \cdot \|_p$, minimization of the operator norm of a matrix, optimization over the cone of positive-semidefinite matrices, finding extremal ellipsoids. Below, we do not deal with the most popular and important application areas of the interior-point methods, meaning linear programming and linearly constrained quadratic programming. The reason is that these classes are covered by the class of quadratically constrained quadratic problems. The specific for LP and LCQP issues of the Karmarkar-type acceleration, which allows us to reduce the (average) arithmetic cost of an iteration, are considered in Chapter 8.

## 6.1 Preliminary remarks

In what follows, we start with a problem in the usual form

$$(f) \quad \text{minimize } f_0(x) \quad \text{subject to } f_i(x) \leq 0, \ i = 1, \dots, m, \ x \in G,$$

where $G$ is a "simple" convex set (e.g., a Euclidean ball) in a finite-dimensional real vector space $E$, and fix the analytical structure of the objective and the constraints; then we point out an equivalent standard and/or conic problem and the corresponding barriers, which allows us to solve the problem by the associated interior-point methods. For brevity, we restrict ourselves to the barrier-generated path-following methods for standard problems and the primal-dual potential reduction method for conic problems, with the emphasis on the complexity of the implementation. Namely, we point out the parameter $\vartheta$ of the resulting barrier (which is responsible for the rate of convergence of the method; recall that, for both of the above methods, it requires us to perform $O(\vartheta^{1/2})$ Newton-type steps to improve the accuracy of the current approximate solution by an absolute constant factor) and the arithmetic cost of a step (i.e., the number of operations of exact real arithmetics).

Recall that, to solve the problem by a path-following method, we need an initial strictly feasible solution to the equivalent standard problem; similarly, to apply the primal-dual method, it is required to know an initial pair of strictly feasible primal and dual solutions. There are many ways to find, under mild restrictions on $(f)$, these strictly feasible solutions by applying the same interior-point methods to appropriate auxiliary problems (where no difficulties with initialization occur). Of course, to perform complete complexity analysis, we should consider the cost at which we can solve the auxiliary problems, but detailed study of these issues would be too time-consuming. To simplify our considerations, we present complete complexity analysis only for the path-following method combined with a special initialization scheme based on the "regularization of the constraints." This scheme is as follows.

Assume that (i) the set $G$ involved into $(f)$ is a bounded and closed convex domain, and we are given a $\vartheta$-self-concordant barrier $F$ for $G$ and an interior point $x^+$ of $G$ such that asymmetry of $G$ with respect to $x^+$ is not worse than a given $\delta > 0$,

$$x^+ + \delta(x^+ - G) \subset G;$$

(ii) We are given a constant $V$ such that

$$f_i(x) \leq V, \qquad i = 1, \ldots, m, \quad x \in G;$$

$$| f_0(x) | \leq V, \qquad x \in G;$$

(iii) The functions $f_i$ are represented by functional elements $(G_i, f_i)$ (see §5.2.1), $G_i \supset G$, and we know $\vartheta_i$-self-concordant barriers $F_i$ for the epigraphs of these functional elements, $i = 0, \ldots, m$.

Besides this, we assume that $(f)$ is consistent (and therefore, in view of (i), (iii) is solvable).

Let $\varepsilon \in (0, V)$. Let us define an $\varepsilon$-solution to $(f)$ as a point $x$ such that

$$x \in G;$$

$$f_i(x) \leq \varepsilon, \qquad i = 1, \ldots, m;$$

$$f_0(x) - f^* \leq \varepsilon,$$

where $f^*$ is the optimal value of the objective in $(f)$.

The *constraint regularization scheme* for finding an $\varepsilon$-solution to $(f)$ ($\varepsilon$ is given in advance) is as follows. Set

$$\Omega(\varepsilon) = \frac{\varepsilon}{3V}$$

and

$$f_\varepsilon(x) = \max\{\Omega(\varepsilon)(f_0(x) + V), f_1(x), \ldots, f_m(x)\}.$$

Consider the problem

$$(f_\varepsilon): \qquad \text{minimize } f_\varepsilon(x) \quad \text{s.t. } x \in G.$$

Let us prove that each point $x \in G$ satisfying the relation

$$(6.1.1) \qquad f_\varepsilon(x) - \min_G f_\varepsilon \le \frac{\varepsilon^2}{3V}$$

is an $\varepsilon$-solution to $(f)$. Indeed, let $x^*$ be an optimal solution to $(f)$. Then, in view of (ii), we have $\min_G f_\varepsilon \le f_\varepsilon(x^*) = \Omega(\varepsilon)(f^* + V)$. Thus, if $x \in G$ satisfies (6.1.1), then

$$f_i(x) \le \Omega(\varepsilon)(f^* + V) + \frac{\varepsilon^2}{3V} \le \frac{2\varepsilon V}{3V} + \frac{\varepsilon^2}{3V} \le \varepsilon, \qquad i = 1, \dots, m,$$

and

$$\Omega(\varepsilon)(f_0(x) + V) \le \Omega(\varepsilon)(f^* + V) + \frac{\varepsilon^2}{3V},$$

whence

$$f_0(x) \le f^* + [\varepsilon^2/(3V)]/[\varepsilon/(3V)] = f^* + \varepsilon.$$

We see that, to find an $\varepsilon$-solution to $(f)$, it suffices to provide (6.1.1). This can be done as follows. Consider the standard problem

$$(f_\varepsilon^+): \qquad \text{minimize } t \quad \text{s.t. } (t, x) \in G_\varepsilon^+,$$

where

$$G_\varepsilon^+ = \{(t, x) \in \mathbf{R} \times E \mid x \in G,\ f_\varepsilon(x) \le t \le 2V\}.$$

From (ii), it follows that $f_\varepsilon(x) \le V$, $x \in G$, so that $(f_\varepsilon^+)$ is equivalent to $(f_\varepsilon)$. Now the function

$$F^+(t, x) = F_0\left(\frac{t}{\Omega(\varepsilon)} - V, x\right) + \sum_{i=1}^m F_i(t, x) + F(x) - \ln(2V - t)$$

is a $\vartheta^+$-self-concordant barrier for the bounded closed convex domain $G^+$, where

$$\vartheta^+ = \sum_{i=0}^m \vartheta_i + \vartheta + 1$$

(see Propositions 5.1.3 and 5.4.1).

It remains to note that we can easily point out an interior point of $G^+$. Indeed, let

$$w = \left(\frac{3}{2}V, x^+\right);$$

since $x^+ \in \operatorname{int} G$ and $f_\varepsilon(x^+) \le V$, we have $w \in \operatorname{int} G^+$. Let us verify that the asymmetry of $G^+$ with respect to $w$ is not worse than $\delta^+ \equiv \min\{\frac{1}{3}, \delta\}$. Indeed, if $(t, x) \in G^+$, then $t \ge f_\varepsilon(x) \ge 0$ (the latter relation follows immediately from (ii)) and $t \le 2V$, so that the point $(t(s), x(s)) \equiv w + s\{w - (t, x)\}$ satisfies the relation $V \le t(s) \le 2V$ for all $s \in [0, \frac{1}{3}]$. For $0 \le s \le \delta$, we also have $x(s) \in G$ (see (i)), so that, for $0 \le s \le \min\{\frac{1}{3}, \delta\}$, we have $x(s) \in G$ and

$f_\varepsilon(x) \le t(s) \le 2V$ (recall that $f_\varepsilon(x) \le V, x \in G$). Thus, for $s = \delta^+$, the point $w + s\{w - (t,x)\}$ does belong to $G^+$ whenever $(t,x) \in G^+$.

Thus, we can solve $(f_\varepsilon^+)$ by the two-stage path-following method associated with the barrier $F^+$ and the starting point $w$ (see §3.2). From Theorem 3.2.1, it follows that, to solve $(f_\varepsilon^+)$ to the accuracy given by (6.1.1) (which results in an $\varepsilon$-solution to $(f)$), it requires to perform no more than

$$M(\varepsilon) = O(1)(\vartheta^+)^{1/2} \ln \frac{V\vartheta^+}{\varepsilon\delta}$$

steps of the preliminary and the main stages, where $O(1)$ is an absolute constant.

To complete the complexity analysis of the path-following method, it suffices to combine the latter efficiency estimate (expressed in terms of the amount of steps) with the evaluation of the arithmetic cost of a step.

Now we are ready to present concrete application examples. In all of these, it is assumed that the problem under consideration is solvable and that we are given the constant $V$ involved into (ii).

## 6.2 Quadratically constrained quadratic problems

Let the functionals involved into $(f)$ be the following convex quadratic forms:

$$f_i(x) = \tfrac{1}{2}x^T A_i x - b_i^T x + c_i : \mathbf{R}^n \to \mathbf{R}, \qquad i = 0, \dots, m,$$

where $A_i$ are positive-semidefinite symmetric $n \times n$ matrices, $b_i \in \mathbf{R}^n$, and let $G = \{x \in \mathbf{R}^n \mid \| x \|_2 \le R\}$ be an Euclidean ball.

### 6.2.1 Path-following approach

To apply the scheme described in §6.1, let us note that, in the case under consideration, (i) is valid with

$$\vartheta = 1, \qquad F(x) = -\ln(R^2 - \| x \|_2^2)$$

(see Proposition 5.4.2) and

$$\delta = 1, \qquad x^+ = 0.$$

To provide (ii), it suffices to set

$$F_i(t,x) = -\ln(t - f_i(x)) \qquad (\vartheta_i = 1,\ i = 0, \dots, m)$$

(see Proposition 5.4.2).

Thus, the barrier associated with problem $(f_\varepsilon^+)$ is

$$F^+(t,x) = -\ln\left(\tfrac{t}{\Omega(\varepsilon)} - V - f_0(x)\right) - \sum_{i=1}^{m} \ln(t - f_i(x))$$
$$- \ln(R^2 - \| x \|_2^2) - \ln(2V - t),$$

the corresponding parameter is $\vartheta^+ = m + 3$, which implies (assuming that $m > 0$) that an $\varepsilon$-solution to $(f)$ can be found in no more than

$$M(\varepsilon) \le O(1)m^{1/2} \ln \frac{2mV}{\varepsilon}$$

steps of the preliminary and the main stages (henceforth, all $O(1)$ denote appropriate *absolute* constants).

The arithmetic cost of a step clearly does not exceed $O(1)(mn^2 + n^3)$ (it costs $O(mn^2)$ operations to form the Newton system and $O(n^3)$ operations to solve it), so that the total arithmetic cost of an $\varepsilon$-solution to $(f)$ is

$$N(\varepsilon) \le O(1)m^{1/2}(mn^2 + n^3) \ln \frac{2mV}{\varepsilon}.$$

### 6.2.2 Potential reduction approach

Quadratically constrained quadratic problems form a nice field for the potential reduction methods, since these problems can be naturally reformulated in the conic form. Indeed, let us first replace $(f)$ with the equivalent problem

$$(f'): \qquad \text{minimize } t \quad \text{s.t. } g_i(t,x) \le 0,\ i = 0,\ldots,m+1,$$

where

$$g_0(t,x) = f_0(x) - t,\ g_i(t,x) = f_i(x),\ i = 1,\ldots,m,\ g_{m+1}(t,x) = x^Tx - R^2$$

are convex quadratic forms of the argument $(t,x) \in \mathbf{R}^{n+1}$.

Now let us construct conic representations for the Lebesgue sets of the constraints. Given a convex quadratic form

$$g(y) = y^TAy + b^Ty + c : \mathbf{R}^s \to \mathbf{R},$$

let us find a decomposition $A = D^TD$ of the matrix $A$, $D$ being an $r \times s$ matrix (the smallest possible value of $r$ is Rank $A$); note that such a decomposition costs $O(s^3)$ arithmetic operations.

Consider the affine mapping

$$\mathcal{B}(y) : \mathbf{R}^s \to \mathbf{R}^{r+2} : \mathcal{B}(y) = \begin{pmatrix} 2Dy \\ 1 + b^Ty + c \\ 1 - b^Ty - c \end{pmatrix}.$$

It is easily seen that $g(y) \le 0$ if and only if $\mathcal{B}(y) \in K^2_{r+2}$, where $K^2_m$ is the standard $m$-dimensional second-order cone

$$\text{(6.2.1)} \qquad K^2_m = \{z \in \mathbf{R}^m \mid z_m \ge (z_1^2 + \cdots + z_{m-1}^2)^{1/2}\}.$$

If the set $\{y \mid g(y) < 0\}$ is nonempty, then the image of the affine mapping $\mathcal{B}$ intersects int $K^2_{r+2}$, so that $\mathcal{B}$ defines a conic representation of the Lebesgue set $\{y \mid g(y) \le 0\}$ of $g$.

Thus, given the quadratic forms $g_0, \dots, g_{m+1}$ involved into $(f')$, we can find (at the total arithmetic cost $O(1)mn^3$) $m+2$ affine mappings $\mathcal{B}_i(t,x)$, $i = 0, \dots, m+1$, from $\mathbf{R}^{n+1}$ into the spaces $\mathbf{R}^{r_i+2}$, $r_i = \text{Rank } g_i''$, such that

$$\{(t,x) \mid g_i(t,x) \leq 0\} = \{(t,x) \mid \mathcal{B}_i(t,x) \in K^2_{r_i+2}\}.$$

Now let

$$E = \prod_{i=0}^{m+1} \mathbf{R}^{r_i+2}, \qquad K = \prod_{i=0}^{m+1} K^2_{r_i+2},$$

$$\mathcal{B}(t,x) = (\mathcal{B}_0(t,x), \dots, \mathcal{B}_{m+1}(t,x)) : \mathbf{R}^{n+1} \to E.$$

Then $K$ is a closed pointed convex cone in $E$ with a nonempty interior, and the feasible set of $(f')$ can be represented as $\{(t,x) \mid \mathcal{B}(t,x) \in K\}$. Note that the linear functional $\phi(t,x) = t$ is constant along Ker $\mathcal{B}$ (since $(f)$ is assumed to be solvable), so that there exists a linear functional $c$ on $E$ such that

$$\langle c, \mathcal{B}(t,x)\rangle = t + \text{const}.$$

Note that $c$ can be found at the arithmetic cost $O(1)mn^3$.

We see that $(f')$ (and therefore $(f)$) is equivalent to the conic problem

$$(f^{\#}): \qquad \text{minimize } \langle c, u\rangle \quad \text{s.t. } u \in K \bigcap \text{Im } \mathcal{B},$$

and it costs $O(1)mn^3$ arithmetic operations to transform the initial data into an explicit representation of the data involved into $(f^{\#})$. Note also that, if $(f)$ satisfies the Slater condition, then $(f^{\#})$ admits strictly feasible solutions.

Now there are no problems with solving $(f^{\#})$ by a potential reduction method. Indeed, let

$$F(u) = -\sum_{i=0}^{m+1} \ln \left( \{u(i)\}^2_{r_i+2} - \sum_{j=1}^{r_i+1} \{u(i)\}^2_j \right), \tag{6.2.2}$$

where $u(i)$ denotes the projection of $u \in E = \prod_{i=0}^{m+1} \mathbf{R}^{r_i+2}$ onto the $i$th direct factor and $\{u(i)\}_j$ denotes the $j$th coordinate of this projection. From Proposition 5.4.3 combined with Proposition 5.1.2, it follows that $F$ is a $\vartheta$-logarithmically homogeneous self-concordant barrier for the cone $K$, where $\vartheta = 2(m+2)$. It is easily seen that the cone dual to $K$ is $K$ itself (of course, we provide $E$ with the standard Euclidean structure, which allows us to identify $E$ and $E^*$), and the Legendre transformation of $F$ is, up to an additive constant (which is of no interest), $F(-s)$, so that $F$ can be used as both the primal and the dual barrier required by the potential reduction methods.

Thus, we can solve $(f)$ by applying to $(f^{\#})$, say, the primal-dual method associated with the barrier $F$, and, to improve the current primal-dual gap by an absolute constant factor, it suffices to perform $O(1)m^{1/2}$ steps of the method.

Now let us evaluate the arithmetic cost of a step. From the description of the method, it is easily seen that this cost does not exceed $O(1)(mn^2 + n^3)$, provided that, at the beginning of the solution process, we compute and store the matrices $B_i^T B_i$, $i = 0, \dots, m+1$, where $B_i$ are the homogeneous parts of the affine mappings $\mathcal{B}_i$. Thus, the total arithmetic cost of the preprocessing (i.e., of computing the data for $(f^\#)$ and the matrices $B_i^T B_i$) is $O(1)mn^3$, while the arithmetic cost at which we can reduce the initial primal-dual gap by a factor of the form $2^{-L}$ is

$$O(1)m^{1/2}(mn^2 + n^3)L.$$

### 6.2.3 Conic problems involving second-order cones

Let $\mu = (m_1, \dots, m_k)$ be a collection of positive integers; such a collection is called *a structure*. For a structure $\mu$, let

$$\mu^s = (m_1^s, \dots, m_k^s), \qquad | \mu | = \sum_{i=1}^{k} m_i$$

and let

$$\mathbf{R}_\mu = \mathbf{R}^{m_1} \times \cdots \times \mathbf{R}^{m_k}.$$

By $K_\mu^2$, we denote the direct product $K_{m_1}^2 \times \cdots \times K_{m_k}^2$ of second-order cones (6.2.1) of the dimensions $m_1, \dots, m_k$; this is a closed convex pointed cone with a nonempty interior, and the cone dual to it is $K_\mu^2$ itself. Conic problems

$$(\mathrm{Q}): \qquad \text{minimize } \langle c, x\rangle \quad \text{s.t. } x \in K_\mu^2 \bigcap (L + b)$$

($L$ is a linear subspace in $\mathbf{R}_\mu$, $b \in \mathbf{R}_\mu$) are called *quadratic* ($q$-problems for short). Usually, we deal with slightly different representation of (Q), namely, with the formulation

$$(\mathrm{Q}'): \qquad \text{minimize } \sigma^T \xi \quad \text{s.t. } \xi \in \mathbf{R}^n, \ \mathcal{A}(\xi) \in K_\mu^2,$$

where $\mathcal{A}$ is an affine mapping from $\mathbf{R}^n$ into $\mathbf{R}_\mu$ (compare with §6.2.2). Similarly to §6.2.2, (Q$'$) can be easily rewritten as (Q), provided that (Q$'$) is solvable.

As we just have indicated, the function

$$(6.2.3) \qquad F(x) = -\sum_{i=1}^{k} \ln \left( \{x(i)\}_{m_i}^2 - \sum_{j=1}^{m_i - 1} \{x(i)\}_j^2 \right) : \operatorname{int} K_\mu^2 \to \mathbf{R},$$

$x(i)$ being the projection of $x \in \mathbf{R}_\mu = \mathbf{R}^{m_1} \times \cdots \times \mathbf{R}^{m_k}$ onto the $i$th direct factor $\mathbf{R}^{m_i}$, is a $2 \mid \mu \mid$-logarithmically homogeneous self-concordant barrier for $K_\mu^2$ and the Legendre transformation $F^*(s)$ of $F$ coincides with $F(-s)$ within an additive constant. As applied to $(\mathcal{Q}')$, the potential reduction methods associated with $F$ improve the accuracy of a given solution by an absolute constant factor in no more than $O(1) \mid \mu \mid^{1/\nu}$ iterations, where $\nu = 1$ for the method of Karmarkar and the projective method and $\nu = 2$ for the primal-dual method. It is easily seen that the arithmetic cost of an iteration does not exceed the quantity

$$O(1)(kn^2 + \mid \mu \mid n + n^3),$$

provided that we henceforth compute and memorize $k$ matrices $A_i^T A_i$, $A_i$ being the $i$th block of the homogeneous part of $\mathcal{A}(\cdot)$ (i.e., the block corresponding to the $i$th direct factor $\mathbf{R}^{m_i}$ in the image space of $\mathcal{A}$); the arithmetic cost of this preprocessing is $O(1) \mid \mu \mid n^2$.

*Problems that can be reformulated as quadratic ones.* Consider a convex programming problem

$$(f): \qquad \text{minimize } f_0(\xi) \quad \text{s.t. } f_i(\xi) \leq 0, \; i = 1, \ldots, m, \;\; \xi \in G,$$

where $G$ is a closed convex domain in $\mathbf{R}^n$ and $f_i$, $i = 0, \ldots, m$, are convex functions represented by functional elements $(G_i, f_i)$ (see §5.2.1) with $G \subseteq G_i$.

The question we now consider is when $(f)$ can be reduced to a $q$-problem. We have just seen that this property is shared by quadratically constrained quadratic problems, but the latter class, in fact, is "significantly smaller" than the whole class of $q$-problems. Indeed, the simplest problem

$$\text{minimize } ax + by \quad \text{s.t. } xy \geq 1, \; x, y \geq 0$$

is, of course, a $q$-problem, since the corresponding feasible set—the interior of a branch of a hyperbola—can be easily represented as the inverse image of the three-dimensional second-order cone under an appropriate embedding of $\mathbf{R}^2$ into $\mathbf{R}^3$; at the same time, this problem, from the standard viewpoint, is *not* a convex quadratically constrained problem, since the constraint $1 - xy \leq 0$ involves a nonconvex quadratic functional.

The natural description of the class of $q$-reducible problems is as follows: To reduce problem $(f)$ to a $q$-problem, it suffices to know *second-order conic representations* (SO-representations) of $G$ and of the f.e. $(G_i, f_i)$, $i = 1, \ldots, m$, i.e., conic representations of the data involving the cones of the type $K^2_{(\cdot)}$. Indeed, given these representations, i.e., affine mappings

$$\gamma_i(\cdot) = (\tau_i(\cdot), \chi_i(\cdot)) : \mathbf{R}^{l_i} \to \mathbf{R} \times \mathbf{R}^n$$

and

$$\pi_i : \mathbf{R}^{l_i} \to \mathbf{R}_{\mu(i)}, \; i = 0, \ldots, m+1,$$

such that

$$(\{(t, \xi) \in \mathbf{R} \times \mathbf{R}^n \mid \xi \in G_i, \, t \geq f_i(\xi)\} \equiv)$$

$$\mathsf{G}(G_i, f_i) = \{(t, \xi) \mid \exists z_i \in \mathbf{R}^{l_i} : \; (t, \xi) = \gamma_i(z_i), \; \pi_i(z_i) \in K^2_{\mu(i)}\}, \; i = 0, \ldots, m,$$

$$G = \{\xi \mid \exists z_{m+1} \in \mathbf{R}^{l_{m+1}} : \; \xi = \gamma_{m+1}(z_{m+1}), \; \pi_{m+1}(z_{m+1}) \in K^2_{\mu(m+1)}\},$$

we can apply to $(f)$ the scheme from §5.2.5.B; this scheme leads to an equivalent conic reformulation of $(f)$. The latter reformulation is

$$(\mathsf{Q}_f): \qquad \text{minimize } t(Z) = \tau_0(z_0) \text{ over } Z = (z_0, \dots, z_{m+1}) \in E \quad \text{s.t. } \mathcal{A}(z) \in K^2_\mu,$$

where

$$E = \{Z = (z_0, \dots, z_{m+1}) \in \mathbf{R}^{l_0} \times \cdots \times \mathbf{R}^{l_{m+1}} \mid \chi_0(z_0) = \chi_1(z_1) = \cdots = \chi_{m+1}(z_{m+1})\},$$

$$\mu = \{\mu(0), \dots, \mu(m+1), \overbrace{1, \dots, 1}^{m \text{ times}}\},$$

$$\mathcal{A}(Z) = (\pi_0(z_0), \dots, \pi_{m+1}(z_{m+1}), -\tau_1(z_1), \dots, -\tau_m(z_m)).$$

The mapping

$$(z_0, \dots, z_{m+1}) \to (\tau_0(z_0), \pi_0(z_0))$$

transforms a feasible solution to $(\mathsf{Q}_f)$ into a feasible solution to the problem

$$(f'): \qquad \text{minimize } t \quad \text{s.t. } t \geq f_0(\xi),\ f_i(\xi) \leq 0,\ i = 1, \dots, m,\ \xi \in G$$

(the latter problem clearly is equivalent to $(f)$), and every feasible solution to $(f')$ can be obtained as the image of a feasible solution to the former problem.

In connection with the above observation, it is worth noting that the class of convex domains (functional elements) that admit SO-representations is closed with respect to the calculus of conic representations described in §5.2 and, in particular, with respect to superpositions (Theorem 5.2.1, as applied to SO-representations of the initial functional elements, leads to an SO-representation of the superposition).

*Examples of* SO*-representable functions.* Let us give a number of examples of SO-representable functions; these are the functions that can occur as the objective and the constraints in problems reducible to quadratic problems.

**1.** *Linear functional.* A linear functional $f(y) = a^T y + b$ admits an SO-representation with $\mu = \{1\}$. It suffices to note that

$$\{(t, y) \mid t \geq f(y)\} = \mathcal{A}^{-1}\left(K^2_{\{1\}}\right), \qquad \mathcal{A}(t, y) = t - f(y).$$

**2.** *Quadratic functional.* A convex quadratic form $f$ of $n$ variables admits an SO-representation with $\mu = \{\operatorname{Rank}\{f''\} + 2\}$; this representation is

$$\{(t, x) \mid t \geq x^T A x + b^T x + c\} = \mathcal{A}^{-1}(K^2_\mu),$$

where (compare with §6.2.2)

$$\mathcal{A}(t, x) = \begin{pmatrix} 2Dx \\ 1 + b^T y + c - t \\ 1 - b^T y - c + t \end{pmatrix},$$

$D$ being a $\operatorname{Rank}\{f''\} \times n$ matrix such that $D^T D = A$.

**3.** *Maximum.* The function $\max\{y_1,\dots,y_l\}$ admits an SO-representation with

$$\mu=\{\overbrace{1,\dots,1}^{m \text{ times}}\},$$

$$\{(t,y) \mid t\geq \max\{y_1,\dots,y_l\}\}=\mathcal{A}^{-1}(K^2_\mu), \qquad \mathcal{A}(t,y)=(t-y_1,\dots,t-y_l).$$

**4.** *Euclidean norm.* The function $\|y\|_2,\ y\in\mathbf{R}^n$ admits the trivial SO-representation with $\mu=\{n+1\}$,

$$\{(t,y) \mid t\geq \|y\|_2\}=K^2_\mu.$$

**5.** *Geometrical mean.* Let $q$ be a positive integer and let $n=2^q$. Consider the function $f(y_1,\dots,y_n)=-(y_1\cdots y_n)^{1/n}:\mathbf{R}^n_+\to\mathbf{R}$. Let us describe an SO-representation of this function. Let the vector of additional variables $s$ be comprised of $n-1$ variables $s_{i,j},\ i=0,\dots,q-1,\ j=1,\dots,2^i$ and let

$$\mu=\{\overbrace{3,\dots,3}^{n-1 \text{ times}},1\}.$$

Define an affine mapping $\mathcal{A}(t;y,s)$ taking values in $\mathbf{R}_\mu$ as follows. $\mathcal{A}$ is comprised of $(n-1)$ components with three-dimensional images and a single component with one-dimensional image. The three-dimensional components are enumerated by the pairs $(i,j),\ i=0,\dots,q-1,\ j=1,\dots,2^i$; the component $(q-1,j)$ is of the form

$$\begin{pmatrix} 2s_{q-1,j}\\ y_{2j-1}-y_{2j}\\ y_{2j-1}+y_{2j}\end{pmatrix};$$

this three-dimensional vector belongs to $K^2_{\{3\}}$ if and only if $y_{2j-1}$ and $y_{2j}$ are both nonnegative and $\mid s_{q-1,j}\mid\leq\sqrt{y_{2j-1}y_{2j}}$.

The component $(i,j)$ with $i<q-1$ is

$$\begin{pmatrix} 2s_{i,j}\\ s_{i+1,2j-1}-s_{i+1,2j}\\ s_{i+1,2j-1}+s_{i+1,2j}\end{pmatrix};$$

it belongs to $K^2_{\{3\}}$ if and only if its "parents" $s_{i+1,2j-1}$ and $s_{i+1,2j}$ are both nonnegative and $\mid s_{i,j}\mid\leq\sqrt{s_{i+1,2j-1}s_{i+1,2j}}$. The (unique) one-dimensional component of $\mathcal{A}(\cdot)$ is $t+s_{0,1}$.

It is clear that $\mathcal{A}(t;y,s)$ belongs to $K^2_\mu$ if and only if the relations

$$\begin{array}{rl}
y_j\geq 0, & 1\leq j\leq 2^q;\\
0\leq s_{q-1,j}\leq (y_{2j-1}y_{2j})^{1/2}, & 1\leq j\leq 2^{q-1};\\
0\leq s_{q-2,j}\leq (s_{q-1,2j-1}s_{q-1,2j})^{1/2}, & 1\leq j\leq 2^{q-2};\\
\dots\dots\dots\dots\dots\dots & \\
\mid s_{0,j}\mid\leq (s_{1,2j-1}s_{1,2j})^{1/2}; & j=1\\
t+s_{0,1}\geq 0 &
\end{array}$$

hold.

Given $y_1, \dots, y_n, t$, we can find $s_{i,j}$ satisfying the above inequalities if and only if all $y_s$ are nonnegative and $t \geq f(y)$, so that the constraint $\{\exists s : \mathcal{A}(t; y, s) \in K^2_\mu\}$ is equivalent to $\{y \geq 0,\ t \geq f(y)\}$. The image of $\mathcal{A}$ clearly intersects int $K^2_\mu$, so that we have constructed the desired representation.

**6.** *Fractional-quadratic function.* What follows originates from the reference [B-TN 92]. Let $A_i$, $i = 0, \dots, k$ be positive-semidefinite $l \times l$ symmetric matrices with a positive-definite sum and let

$$A(s) = A_0 + A_1 s_1 + \cdots + A_k s_k : \mathbf{R}^k \to S_l,$$

where $S_l$ denotes the space of symmetric $l \times l$ matrices. Since the sum of $A_i$ is positive-definite, $A(s)$ is positive-definite for positive $s$, so that the following fractional-quadratic function is well defined:

$$f(x, s) = x^T [A(s)]^{-1} x : Q \equiv \mathbf{R}^l \times \{s \in \mathbf{R}^k, s > 0\} \to \mathbf{R}.$$

Extending $f$ from $Q$ onto $G = \mathrm{cl}\ Q$ as a lower semicontinuous function taking values in the extended axis $\mathbf{R} \bigcup \{+\infty\}$, we obtain a functional element $(G, f)$; problems involving constraints and objective of the above fractional-quadratic type arise in some applications, e.g., to Truss Topology Design [B-TN 92]). It is easily seen that the element $(G, f)$ can be expressed as

$$\begin{aligned} G &= \{(x, s) \in \mathbf{R}^l \times \mathbf{R}^k \mid s \geq 0\}, \\ f(x, s) &= \sup\{2x^T u - u^T A(s) u \mid u \in \mathbf{R}^l\}. \end{aligned} \tag{6.2.4}$$

Let us construct an SO-representation of the f.e. $(G, f)$.

Let $D_i$ be $r_i \times l$ matrices $r_i = \mathrm{Rank}\,\{A_i\}$, such that $D_i^T D_i = A_i$, $i = 0, \dots, k$, and let us define a linear space $H$ and a structure $\mu$ as

$$\begin{aligned} H = \{z = (y_0, \dots, y_k, x, s, \tau, t) &\in \mathbf{R}^{r_0} \times \cdots \times \mathbf{R}^{r_k} \times \mathbf{R}^l \times \mathbf{R}^k \times \mathbf{R}^k \times \mathbf{R} \mid \\ D_0^T y_0 + \cdots + D_k^T y_k &= x\}, \\ \mu &= \{r_0 + 2, \dots, r_k + 2, 1\}. \end{aligned}$$

Let us define the affine mapping $\gamma$ from $H$ into $\mathbf{R} \times \mathbf{R}^l \times \mathbf{R}^k$ and $\pi$ from $H$ into $\mathbf{R}_\mu$ as follows:

$$\gamma(y_0, \dots, y_k, x, s, \tau, t) \equiv (t, x, s),$$

$$\pi(y_0, \dots, y_k, x, s, \tau, t) = (\pi_0(y_0, s, \tau), \dots, \pi_k(y_k, s, \tau), \pi_{k+1}(\tau, t)),$$

where

$$\pi_0(y_0, s, \tau) = \begin{pmatrix} 2y_0 \\ \tau_0 - 1 \\ \tau_0 + 1 \end{pmatrix}, \qquad \pi_i(y_i, s, \tau) = \begin{pmatrix} 2y_i \\ \tau_i - s_i \\ \tau_i + s_i \end{pmatrix}, \quad i = 1, \dots, k,$$

$$\pi_{k+1}(\tau, t) = t - \tau_0 - \cdots - \tau_k.$$

We claim that $H$, $\pi$, $\gamma$ define a conic representation of the epigraph of $(G, f)$, i.e., that

$$\mathsf{G}(G,f) \equiv \{(t,x,s) \,|, (x,s) \in G, t \geq f(x,s)\}! = \{\gamma(z) \mid z \in H, \pi(z) \in K^2_\mu\} \tag{6.2.5}$$

and that Im $\pi$ intersects int $K^2_\mu$. The latter statement is evident. To prove (6.2.5), denote the right-hand side of this relation by $Q$ and note that, by construction, $Q$ is the set of all $z = (y_0, \ldots, y_k, x, s, \tau, t)$ satisfying the system of equations and inequalities as follows:

$$D_0^T y_0 + \cdots + D_k^T y_k = x, \tag{6.2.6}$$

$$y_i^T y_i \leq s_i \tau_i, \ \tau_i + s_i \geq 0, \qquad i = 0, \ldots, k, \tag{6.2.7}$$

where $s_0 \equiv 1$,

$$t \geq \tau_0 + \cdots + \tau_k. \tag{6.2.8}$$

Equation (6.2.6) is the definition of $H$, relations (6.2.7) express the inclusions $\pi_i(z) \in K^2_{r_i+2}$, and (6.2.8) is the same as the relation $\pi_{k+1}(z) \in K^2_1$.

To prove (6.2.5), note that in view of (6.2.4) $(t, x, s) \in \mathsf{G}(G, f)$ if and only if $s \geq 0$ and the quadratic form of $u \in \mathbf{R}^l 2x^T u - u^T A(s) u$ does not exceed $t$, or, which is the same, if and only if

$(*)$ : $s \geq 0$ and the linear system with respect to $u$

$$A(s)u = x \tag{6.2.9}$$

is solvable and for some (equivalently, for any) solution to this system we have

$$u^T x \leq t. \tag{6.2.10}$$

Let $(t, x, s) \in \mathsf{G}(G, f)$ and let $u$ be a solution to the corresponding system (6.2.9). Let us set

$$y_i = s_i D_i u \in \mathbf{R}^{r_i}, \qquad i = 0, \ldots, k \tag{6.2.11}$$

(henceforth, $s_0 \equiv 1$).

Relation (6.2.9) means that

$$D_0^T y_0 + \cdots + D_k^T y_k = x, \tag{6.2.12}$$

so that $(y_0, \ldots, y_k, x)$ satisfies (6.2.6); multiplying (6.2.12) by $u^T$ and taking into account (6.2.11), we obtain

$$x^T u = \sum_{i=0}^{k} \frac{y_i^T y_i}{s_i}$$

(henceforth $a/0 = +\infty$ for $a > 0$ and $a/0 = 0$ for $a = 0$), so that the reals

$$\tau_i = \frac{y_i^T y_i}{s_i}, \qquad i = 0, \ldots, m$$

extend the vector $s$ to a collection satisfying (6.2.7); in view of (6.2.10), the above $\tau_i$ satisfy (6.2.8). We conclude that the vector $z = (y_0, \ldots, y_k, x, s, \tau, t)$ is feasible for (6.2.6)–(6.2.8), since, by construction, $\gamma(z) = (t, x, s)$, we have proved that the left-hand side of (6.2.5) is contained in the right-hand side $Q$.

To prove (6.2.5), it remains to establish the inclusion $Q \subseteq \mathsf{G}(G, f)$. Let

$$z = (y_0, \ldots, y_k, x, s, \tau, t) \in Q,$$

so that $z$ satisfies (6.2.6)–(6.2.8). We should prove then that $(t, x, s) \in \mathsf{G}(G, f)$. Let us first note that from (6.2.7) it follows $s \geq 0$, so that $(x, s) \in G$. Furthermore, let $I$ be the set of indices of nonzero $s_i$, $i = 0, \ldots, k$; from (6.2.7), it follows that $y_i = 0$, $i \notin I$. Consider the auxiliary problem

$$\text{(6.2.13)} \qquad \text{minimize} \sum_{i \in I} \frac{v_i^T v_i}{s_i} \text{ over } v_i,\ i \in I \quad \text{s.t.} \sum_{i \in I} D_i^T v_i = x;$$

this problem is feasible. A feasible solution is given by $v_i = y_i$, $i \in I$, and the objective at the latter solution is $\leq t$ (see (6.2.7), (6.2.8)). It follows that the problem is solvable, and, for its optimal solution $y_i^*$, $i \in I$, we have

$$\text{(6.2.14)} \qquad \frac{(y_i^*)^T y_i^*}{s_i} \leq t.$$

Now optimality conditions for (6.2.13) imply that, for some $u \in \mathbf{R}^l$,

$$y_i^* = s_i D_i u, \qquad i \in I.$$

From the latter relation and the fact that $y_i^*$ are feasible for (6.2.13), it follows that

$$A(s)u = \sum_{i \in I} s_i D_i^T D_i u = x;$$

multiplying the latter equation by $u^T$, we obtain $x^T u = \sum_{i \in I} (y_i^*)^T y_i^* / s_i$, and (6.2.14) implies that the latter quantity is $\leq t$, so that $(t, x, s) \in \mathsf{G}(G, f)$ (see the description of the latter set given by (*)). Thus, (6.2.5) is proved.

## 6.3 More of structured nonlinear problems

### 6.3.1 Geometrical programming problem (exponential form)

The class of geometrical programming problems corresponds to the case when the functionals $f_i$, $i = 0, \ldots, m$ involved into $(f)$ are of the form

$$f_i(x) = \sum_{j=1}^{r_i} c_{ij} \exp\{a(i,j)^T x\} + d_i : \mathbf{R}^n \to \mathbf{R}, \qquad i = 0, \ldots, m,$$

where $a(i,j)$ are vectors from $\mathbf{R}^n$ and all the coefficients $c_{ij}$ are positive; as in the previous section, we assume that $G = \{x \mid \| x \|_2 \leq R\}$ is an Euclidean ball (with minor modifications, what follows is also valid for, say, a polytope $G$).

To apply to $(f)$ the scheme from §6.1, it is reasonable to reformulate the problem as follows. Let

$$A = \{a(i,j) \mid 0 \leq i \leq m,\ 1 \leq j \leq r_i\}$$

and let $k$ be the number of elements in $A$ (i.e., the number of different exponentuals $a(i,j)$ involved into the problem). Let $a_s$, $1 \leq s \leq k$ be the $s$th element of $A$ with respect to some ordering, let $s(i,j)$ be defined by the relation

$$a(i,j) = a_{s(i,j)},$$

and let $c_s$ be the largest of those coefficients $c_{ij}$ for which $s(i,j) = s$. Set $\sigma_{ij} = c_{ij}/c_{s(i,j)}$, so that $0 < \sigma_{ij} \leq 1$.

Adding $k$ new variables $\tau_1, \ldots, \tau_k$, we can rewrite $(f)$ in the equivalent form

$$(f^{\#}): \qquad \text{minimize } g_0(\tau) = \sum_{j=1}^{r_0} \sigma_{0j}\tau_{s(0,j)} + d_0$$

subject to

$$g_i(\tau) = \sum_{j=1}^{r_i} \sigma_{ij}\tau_{s(i,j)} + d_i \leq 0, \qquad i = 1, \ldots, m,$$

$$g_{m+s}(\tau, x) = c_s \exp\{a_s^T x\} - \tau_s \leq 0, \qquad s = 1, \ldots, k,$$

$$(\tau, x) \in G^{\#} = \{(\tau, x) \mid \| x \|_2 \leq R, 0 \leq \tau_s \leq V^+,\ s = 1, \ldots, k\},$$

where

$$V^+ = V + \max_{0 \leq i \leq m} \mid d_i \mid .$$

First, let us verify that $(f^{\#})$ is equivalent to $(f)$. It suffices to prove that, if $x \in G$, then $c_s \exp\{a_s^T x\} \leq V^+$ for all $s$. Indeed, since, by the definition of $V$, $f_i(x) \leq V$, $x \in G$, we have $c_{ij} \exp\{a_{s(i,j)}^T x\} \leq V - d_i \leq V^+$ for all $i$ and $j$, and the desired inequalities do hold.

Second, now we are in the situation required by the scheme from §6.1. Indeed, $G^{\#}$ clearly is bounded, and the point

$$x^+ = \left(\tau_1 = \frac{V^+}{2}, \ldots, \tau_k = \frac{V^+}{2}, x = 0\right)$$

is the symmetry center of $G^{\#}$ (so that we can set in (i) $\delta = 1$). From Propositions 5.1.3, 5.4.1, and 5.4.2, it follows that the function

$$F(\tau, x) = -\sum_{s=1}^{k} \{\ln \tau_s + \ln(V^+ - \tau_s)\} - \ln(R^2 - x^T x)$$

is a $\vartheta$-self-concordant barrier for $G^{\#}$ required in (i), where $\vartheta = 2k+1$. Furthermore, for $(\tau, x) \in G^{\#}$, we clearly have

$$g_i(\tau, x) \le (1 + \max_{0 \le i \le m} r_j)V^+ \equiv V^{\#}, \qquad i = 1, \ldots, m+k,$$

$$| \, g_0(\tau, x) \, | \le V^{\#},$$

as is required in (ii). It remains to provide (iii), i.e., to point out $\vartheta_i$-self-concordant barriers $F_i$ for the epigraphs of the functions $g_i$. In view of Proposition 5.4.1, we can take

$$\vartheta_i = 1, \;\; F_i(t; \tau, x) = -\ln(t - g_i(\tau, x)), \qquad i = 0, \ldots, m;$$

from Proposition 5.3.3, it follows that we can also take

$$\vartheta_{m+s} = 2,$$

$$F_{m+s}(t; \tau, x) = \{-\ln(t + \tau_s) - \ln(\ln(t + \tau_s) - \ln c_s - a_s^T x)\}.$$

Thus, to find an $\varepsilon$-solution to $(f)$, we can apply to $(f^{\#})$ the scheme from §6.1. The barrier $F^+$ associated with the resulting problem $((f^{\#})^+_\varepsilon)$ is as follows:

(6.3.1)
$$\begin{aligned} F^+(t; \tau, x) = & -\ln\left(\frac{t}{\Omega^{\#}(\varepsilon)} - V^{\#} - g_0(\tau)\right) - \sum_{i=1}^{m} \ln(t - g_i(\tau)) \\ & - \sum_{s=1}^{k} \left\{ \ln(t + \tau_s) + \ln\left(\ln(t + \tau_s) - \ln c_s - a_s^T x\right)\right\} \\ & - \sum_{s=1}^{k} \{\ln \tau_s + \ln(V^+ - \tau_s)\} - \ln(R^2 - x^T x) - \ln(2V^{\#} - t), \end{aligned}$$

where
$$\Omega^{\#}(\varepsilon) = \frac{\varepsilon}{3V^{\#}}.$$

The parameter of the barrier is $\vartheta^+ = 4k + m + 3$.

From the results of §6.1, it follows that an $\varepsilon$-solution to $(f)$, $0 < \varepsilon \le V^{\#}$ can be found in no more than

$$M(\varepsilon) = O(1)(k+m)^{1/2} \ln \frac{(k+m)V^{\#}}{\varepsilon}$$

steps of the preliminary and the main stages of the barrier-generated path-following method associated with $F^+$.

Let us evaluate the arithmetic cost of a step. The computational effort at a step is dominated by solving the Newton system with the $(n+k+1)\times(n+k+1)$ matrix (the Hessian of $F^+$), and the right-hand side is comprised of a fixed vector and the gradient of $F^+$. It is easily seen that the right-hand side can be computed in $O(1)(m+n)k$ operations. Furthermore, from (6.3.1) it follows that, to multiply a given vector by the Hessian of $F^+$, it requires no more than $O(1)k(m+n)$ operations. Thus, the system can be solved by the conjugate gradient method in $O(1)k(m+n)(n+k)$ operations, so that the total arithmetic cost of an $\varepsilon$-solution to $(f)$ does not exceed the quantity

$$O(1)k(m+n)(n+k)(m+k)^{1/2}\ln\frac{(m+k)V^{\#}}{\varepsilon}.$$

### 6.3.2 Approximation in $L_p$-norm

The problem of approximation in $L_p$-norm can be formulated as follows: Given an affine subspace $L$ in $\mathbf{R}^k$ and a vector $b$ from $\mathbf{R}^k$, find the element of $L$ closest to $b$ with respect to the norm $\|\cdot\|_p$ . We study a slightly more general problem, in which the approximation is subject to some additional (for the sake of simplicity, quadratic) constraints. In other words, we consider problem $(f)$ under the assumptions that

$$f_0(x) = \sum_{j=1}^{k} | a_j^T x - b_j |^p, \; x, a_j \in R^n \qquad (p \in [1,\infty))$$

and all $f_i(x)$, $i = 1, \dots, m$ are convex quadratic forms. As usual, $G$ is supposed to be the ball $\{\| x \|_2 \le R\}$.

Similarly to above, to apply the scheme from §6.1, we should reformulate the problem. Namely, introducing $k$-dimensional vector $\tau = (\tau_1, \dots, \tau_k)^T$ of additional variables, we can rewrite $(f)$ as the problem

$$(f^{\#}): \qquad \text{minimize } g_0(\tau) \equiv \sum_{j=1}^{k} \tau_j$$

subject to

$$g_j(\tau, x) \equiv | a_j^T x - b_j |^p - \tau_j \le 0, \qquad 1 \le j \le k,$$
$$g_{k+i}(\tau, x) \equiv f_i(x) \le 0, \qquad 1 \le i \le m,$$
$$(\tau, x) \in G^{\#} =] \left\{ (\tau, x) \mid \| x \|_2 \le R, \; 0 \le \tau_j, 1 \le j \le k, \; \sum_{j=1}^{k} \tau_j \le V \right\},$$

which is clearly equivalent to $(f)$.

Let us indicate the data required by the scheme of §6.1 as applied to $(f^{\#})$. First, the set $G^{\#}$ admits a $(k+2)$-self-concordant barrier

$$F(\tau, x) = -\ln(R^2 - x^T x) - \sum_{j=1}^{k} \ln \tau_j - \ln\left(V - \sum_{j=1}^{k} \tau_j\right)$$

and the asymmetry coefficient of this set with respect to the point

$$x^+ = \left(\tau_1 = \frac{V}{k+1}, \dots, \tau_k = \frac{V}{k+1}, x = 0\right),$$

so that (i) is satisfied with $\vartheta = k+2$ and $\delta = 1/(k+1)$. Part (ii) clearly holds for $(f^{\#})$ with the same constant $V$ as in $(f)$. It remains to provide the barriers required in (iii). By Proposition 5.3.1, the function

$$\psi_{(p)}(t,u) = -\{2\ln t + \ln(t^{2/p} - u^2)\}$$

is a 4-self-concordant barrier for the set $\{(t,u) \in \mathbf{R}^2 \mid t \geq |u|^p\}$. Thus, we can take

$$\vartheta_0 = 1,\ F_0(t;\tau,x) = -\ln(t - g_0(\tau))$$

(see Proposition 5.4.1),

$$\vartheta_j = 4, \quad F_j(t;\tau,x) = \psi_{(p)}(t+\tau_j, a_j^T x - b_j), \quad j = 1,\ldots,k,$$

$$\vartheta_{k+i} = 1, \quad F_{k+i}(t;\tau,x) = -\ln(t - f_i(x)), \quad i = 1,\ldots,m$$

(see Proposition 5.4.2).

Thus, to find an $\varepsilon$-solution to $(f)$, we can apply to $(f^{\#})$ the scheme from §6.1. The barrier $F^+$ associated with the resulting problem $((f^{\#})^+_\varepsilon)$ is as follows:

(6.3.2)
$$\begin{aligned} F^+(t;\tau,x) = &-\ln\left(\frac{t}{\Omega(\varepsilon)} - V - g_0(\tau)\right) - \sum_{j=1}^m \ln(t - f_j(x)) \\ &-\sum_{j=1}^k\{2\ln(t+\tau_j) + \ln\left((t+\tau_j)^{2/p} - (a_j^T x - b_j)^2\right)\} \\ &-\sum_{s=1}^k \ln\tau_s - \ln\left(V - \sum_{s=1}^k \tau_s\right) - \ln(R^2 - x^T x) - \ln(2V - t), \end{aligned}$$

where

$$\Omega(\varepsilon) = \frac{\varepsilon}{3V}.$$

The parameter of the barrier is $\vartheta^+ = 5k + m + 4$.

From the results of §6.1, it follows that an $\varepsilon$-solution to $(f)$, $0 < \varepsilon \leq V$ can be found in no more than

$$M(\varepsilon) = O(1)(k+m)^{1/2}\ln\frac{(k+m)V}{\varepsilon}$$

steps of the preliminary and the main stages of the barrier-generated path-following method associated with $F^+$.

It is easily seen that the arithmetic cost of a step does not exceed $O(1)((m+k)n^2 + (k+n)^3)$. Thus, the arithmetic cost of an $\varepsilon$-solution to $(f)$ does not exceed the quantity

$$O(1)\left((m+k)n^2 + (k+n)^3\right)(m+k)^{1/2}\ln\frac{(m+k)V}{\varepsilon}.$$

Note that the approach used in §§6.3 and 6.3.2 can, in fact, be applied to any convex problem $(f)$ with the functionals of the type

$$f_i(x) = \sum_{j=1}^{r_i} \phi_{ij}(a_{ij}^T x),$$

where $\phi_{ij}(t)$ are convex univariate functions, provided that we can point out explicit self-concordant barriers for the epigraphs of $\phi_{ij}$. As was mentioned in §5.3, it is not too difficult to find the latter barriers, since even the universal barrier for a two-dimensional closed convex domain can be regarded as "computable."

### 6.3.3 Minimization of the matrix norm

The problem we study is as follows: Given an affine function $A(x)$ taking values in the space $L_{k,l}$ of real $k \times l$ matrices, minimize the operator norm $\| A(x) \|$ of it. Here the operator norm of a $k \times l$ matrix $A$ is defined as

$$\| A \| = \max\{\| Au \|_2 \mid u \in \mathbf{R}^l, \ \| u \|_2 \leq 1\},$$

or, which is the same, as the maximal eigenvalue of the matrix $| A | = (A^T A)^{1/2}$. In addition, we will restrict $x$ by a number of (for simplicity, quadratic) constraints. Thus, we are considering problem $(f)$ in the case when

$$f_0(x) = \| A(x) \|, \qquad A(\cdot) : \mathbf{R}^n \to L_{k,l},$$

where the elements of $A(x)$ are affine functions of $x$, and all $f_i$ are convex quadratic forms, $i = 1, \ldots, m$. As usual, we take $G = \{x \mid \| x \|_2 \leq R\}$. Without loss of generality, we can assume that $k \leq l$ (otherwise, we could replace $A(x)$ with $A^T(x)$).

Let us apply the scheme from §6.1, provided that we are given the constant $V$ involved into (ii). The function

$$F(x) = -\ln(R^2 - x^T x)$$

is an 1-self-concordant barrier for $G$ (Proposition 5.4.2), and $G$ is symmetric with respect to $x^+ = 0$ ($\delta = 1$). According to Proposition 5.4.6, the function

$$F_0(t, x) = -\ln \mathrm{Det}\left(tI_k - \frac{1}{t} A(x) A^T(x)\right) - \ln t$$

is a $(k+1)$-self-concordant barrier for the epigraph $\{t \geq \| A(x) \|\}$ of the objective; here $I_k$ denotes the unit $k \times k$ matrix. In view of Proposition 5.4.2, the functions

$$F_i(t, x) = -\ln(t - f_i(x)), \qquad i = 1, \ldots, m$$

are 1-self-concordant barriers for the epigraphs of the constraints. Thus, we can apply to $(f)$ the scheme of §6.1. The resulting barrier is

$$F^+(t,x) = -\ln \operatorname{Det}\left\{\left(\frac{t}{\Omega(\varepsilon)} - V\right) I_k - \frac{\Omega(\varepsilon)}{t-\Omega(\varepsilon)V} A(x)A^T(x)\right\}$$
$$-\ln\left(\frac{t}{\Omega(\varepsilon)} - V\right) - \sum_{i=1}^{m} \ln(t - f_i(x)) - \ln(R^2 - \| x \|_2^2) - \ln(2V - t),$$

and its parameter is $\vartheta^+ = k + m + 3$. Therefore, to find an $\varepsilon$-solution to $(f)$, it suffices to perform

$$M(\varepsilon) \le O(1)(m+k)^{1/2} \ln \frac{(m+k)V}{\varepsilon}$$

steps of the barrier-generated path-following method associated with $F^+$.

The arithmetic cost of a step evidently does not exceed $O(n^2(m+n) + nkl(n+k))$ operations, so that the arithmetic cost of an $\varepsilon$-solution to $(f)$ is

$$O(1)(m+k)^{1/2}\left(n^2(m+n) + nkl(n+k)\right) \ln \frac{(m+k)V}{\varepsilon}.$$

## 6.4 Semidefinite programming

Under the wording "semidefinite programming," we understand conic problems on the cone of positive-semidefinite symmetric matrices. As we see, this class of problems is a very nice field for the interior-point machinery. First, many important convex problems can be naturally reformulated as the problems from the above class. Second, there are no difficulties with fine primal and dual logarithmically homogeneous self-concordant barriers for the corresponding cones, so that, to solve the problems, we can use the most attractive potential reduction interior-point methods.

### 6.4.1 Preliminary remarks

Recall that a structure $\mu = (m_1, \ldots, m_k)$ is a collection of positive integers and that, for a structure $\mu$, we set $\mu^s = (m_1^s, \ldots, m_k^s)$, $\mid \mu \mid = \sum_{i=1}^k m_i$. Let $\mu$ be a structure and let $S_\mu$ denote the space of all symmetric block-diagonal $\mid \mu \mid \times \mid \mu \mid$ matrices with $k$ diagonal blocks of the row sizes $m_1, \ldots, m_k$. The space is provided with the scalar product $\langle x, y\rangle = \operatorname{Tr}\{xy\}$. By $S_\mu^+$, we denote the cone in $S_\mu$ formed by all positive-semidefinite matrices from $S_\mu$; this is a closed convex pointed cone with a nonempty interior, and the cone dual to it is $S_\mu^+$ itself. The function

$$F(x) = -\ln \operatorname{Det} x : \operatorname{int} S_\mu^+ \to \mathbf{R}$$

is a $\mid \mu \mid$-logarithmically homogeneous self-concordant barrier for $S_\mu^+$ (see **Proposition** 5.4.5), and it is easily seen that the Legendre transformation $F^*(s)$

of $F$ coincides with $F(-s)$ within an additive constant

$$F^*(s) = F(-s) + \text{ const.}$$

Thus, to solve a conic problem on $S^+_\mu$, i.e., the problem of the form

$$(\mathsf{P}): \qquad \text{minimize } \langle c, x\rangle \quad \text{s.t. } x \in S^+_\mu \bigcap (L+b)$$

($L$ is a linear subspace of $S_\mu$, $b \in S_\mu$), we can use each of the associated with $F$ potential reduction methods described in Chapter 4. The amount of steps of the methods that are required to improve the accuracy by an absolute constant factor is $O(1) \mid \mu \mid^{1/\nu}$, where $\nu = 1$ for the method of Karmarkar and the projective method and $\nu = 2$ for the primal-dual method.

Usually, we deal with slightly different representation of (P), namely, with the formulation (compare with §6.2.3)

$$(\mathsf{P}'): \qquad \text{minimize } \sigma^T\xi \quad \text{s.t. } \xi \in \mathbf{R}^n,\ \mathcal{A}(\xi) \in S^+_\mu,$$

where $\mathcal{A}$ is an affine mapping from $\mathbf{R}^n$ into $S_\mu$. Note that (P$'$) can be easily rewritten as (P), provided that (P$'$) is solvable. In what follows, we refer to (P) (or, which is the same, to (P$'$)) as to a *pd-problem of structure* $\mu$.

As already mentioned, to improve the accuracy of the current approximate solution to (P$'$) by an absolute constant factor, it suffices to perform $O(1) \mid \mu \mid^{1/\nu}$ steps of a potential reduction interior-point method. What is the arithmetic cost of a step? This quantity is dominated by the computational effort of finding the Newton direction of a given linear functional reduced to the feasible plane of (P$'$), i.e., by the cost at which, given $d \in S_\mu$ and $x \in \operatorname{int} S^+_\mu$, we can minimize the quadratic form

$$p(\xi) = \langle d, A\xi\rangle + \tfrac{1}{2}\left\langle A^T F''(x) A\xi, \xi\right\rangle$$

over $\xi \in \mathbf{R}^n$, where $A$ is the matrix involved into $\mathcal{A}$. If $a_i \in S_\mu$, $i = 1, \dots, n$, denote the columns of $A$; then the elements of the $n\times n$ matrix $Q = A^T F''(x)A$ are $Q_{ij} = \mathrm{Tr}\{x^{-1}a_i x^{-1}a_j\}$, so that, to form $Q$, it costs $O(1) \mid \mu^3 \mid$ operations to compute $x^{-1}$, $O(1)n \mid \mu^3 \mid$ operations more to compute all the matrices $x^{-1}a_i x^{-1}$, $i = 1, \dots, n$, and then $O(1)n^2 \mid \mu^2 \mid$ operations to compute all $Q_{ij}$. Since minimizing $p$ is the same as computing $Q^{-1}A^T d$, the arithmetic cost of a step is

$$O(1)\{n^3 + n^2 \mid \mu^2 \mid + n \mid \mu^3 \mid\}.$$

Of course this cost can be reduced, if the $a_i$ are, say, sparse enough.

### 6.4.2 Positive-definite representable functions

Consider a convex programming problem

$$(f): \qquad \text{minimize } f_0(\xi) \quad \text{s.t. } f_i(\xi) \le 0,\ i = 1, \dots, m,\ \xi \in G,$$

where $G$ is a closed convex domain in $\mathbf{R}^n$ and $f_i$, $i = 0, \ldots, m$ are convex functions represented by functional elements $(G_i, f_i)$ (see §5.2.1) with $G \subseteq G_i$. Similarly to the case of quadratic problems (see §6.2.3), to reduce $(f)$ to $(\mathsf{P}')$, it suffices to find *positive-semidefinite representations* (PD-representations) of $G$ and of the f.e. $(G_i, f_i)$, $i = 1, \ldots, m$, i.e., conic representations of the above data involving the cones of the type $S^+_\mu$.

Similarly to the quadratic case, the class of convex domains (functional elements) that admit PD-representations is closed with respect to the calculus of conic representations described in §5.2 and, in particular, with respect to superpositions (Theorem 5.2.1, as applied to positive-definite conic representations of the initial functional elements, leads to a positive-semidefinite conic representation of the superposition).

In what follows, we present some concrete applications; in particular, we give a number of important examples of positive-semidefinite representable functions. These functions can be involved into $(f)$ as the objective and the constraints.

### 6.4.3 Examples of PD-representable functions

**1.** *Euclidean norm.* The function $\| y \|_2, y \in \mathbf{R}^n$ admits a PD-representation with $\mu = \{n+1\}$,

$$\{(t, y) \mid t \geq \| y \|_2\} = \mathcal{B}^{-1}(S^+_\mu), \qquad \mathcal{B}(t, y) = \begin{pmatrix} t & y^T \\ y & tI_n \end{pmatrix}$$

(as usual, $I_l$ denotes the $l \times l$ unit matrix).

The above statement means precisely that the second-order cone $K^2_{n+1}$ can be represented as an intersection of the cone of positive-semidefinite symmetric matrices $S^+_{\{n+1\}}$ and an appropriate linear subspace passing through the origin and intersecting $\operatorname{int} S^+_{\{n+1\}}$. It follows that any function that admits a second-order representation (see §6.2.3) also admits a PD-representation. This observation underlies part of the below examples.

**2.** *Linear functional.* A linear functional $f(y) = a^T y + b$ admits a PD-representation with $\mu = \{1\}$. It suffices to note that

$$\{(t, y) \mid t \geq f(y)\} = \mathcal{A}^{-1}(S^+_{\{1\}}), \qquad \mathcal{A}(t, y) = t - f(y).$$

**3.** *Quadratic functional.* A convex quadratic form $f$ of $n$ variables admits a PD-representation with $\mu = \{\text{Rank } f'' + 2\}$. Indeed, the epigraph of $f$ is of the form $\{(t, y) \mid g(t, y) \leq 0\}$, where $g(t, y)$ is a convex quadratic form of $(t, y)$ and the set $Q = \{(t, y) \mid g(t, y) < 0\}$ is nonempty. As we have seen in §6.2.2, there exists (and can be effectively found) an affine mapping $\mathcal{C}(t, y)$ from $\mathbf{R} \times \mathbf{R}^n$ into $\mathbf{R}^{r+2}$, $r = \text{Rank } g'' = \text{Rank } f''$, such that the image of $\mathcal{C}$ intersects $\operatorname{int} K^2_{r+2}$ (see (6.2.1)) and

$$\{(t, y) \mid t \geq f(y)\} = \mathcal{C}^{-1}(K^2_{r+1}).$$

In turn, the cone $K^2_{r+2}$ can be represented as the inverse image of $S^+_{\{r+2\}}$ under a linear mapping $\mathcal{B}$ (see §6.4.3.1). Thus,

$$\{(t,y) \mid t \geq f(y)\} = (\mathcal{B} \circ \mathcal{C})^{-1}(S^+_{\{r+2\}}),$$

and this is the desired PD-representation of the epigraph of $f$, since the mapping $\mathcal{B}$ presented in §6.4.3.1 is such that the image of the mapping $\mathcal{B} \circ \mathcal{C}$ intersects int $S^+_{\{r+2\}}$.

**4.** *Maximum.* The function $\max\{y_1, \ldots, y_l\}$ admits a PD-representation with

$$\mu = \{\overbrace{1, \ldots, 1}^{l \text{ times}}\},$$

as follows:

$$\{(t,y) \mid t \geq \max\{y_1, \ldots, y_l\}\} = \mathcal{A}^{-1}(S^+_\mu), \qquad \mathcal{A}(t,y) = \mathrm{Diag}\,\{t-y_1, \ldots, t-y_l\}.$$

**5.** *Matrix norm.* The matrix norm $\| y \|$ on the space $L_{p,q}$ of $p \times q$ matrices admits a PD-representation with $\mu = \{p+q\}$,

$$\{(t,y) \mid t \geq \| y \|\} = \mathcal{A}^{-1}(S^+_\mu), \qquad \mathcal{A}(t,y) = \begin{pmatrix} tI_q & y^T \\ y & tI_p \end{pmatrix}.$$

**6.** *Maximal eigenvalue.* The maximal eigenvalue $\lambda_{\max}(y)$ of a symmetric $p \times p$ matrix $y$ admits a PD-representation with $\mu = \{p\}$,

$$\{(t,y) \mid t \geq \lambda_{\max}(y)\} = \mathcal{A}^{-1}(S^+_\mu), \qquad \mathcal{A}(t,y) = tI_p - y.$$

**7.** *The sum of k largest eigenvalues.* Let $y$ denote a symmetric $p \times p$ matrix and let $\lambda_1(y) \geq \lambda_2(y) \geq \cdots \geq \lambda_p(y)$ be the eigenvalues of $y$ written in the descent order. For $1 \leq k \leq p$, set

$$\sigma_k(y) = \sum_{i=1}^{k} \lambda_i(y).$$

This function admits the following PD-representation with $\mu = \{1, p, p\}$:

$$\{(t,y) \mid t \geq \sigma_k(y)\} = \mathcal{A}^{-1}(S^+_\mu),$$

where

$$\mathcal{A} : \mathbf{R} \times \mathbf{R} \times S_{\{p\}} \times S_{\{p\}} \to S_\mu : \mathcal{A}(t,s,y,z) = \begin{pmatrix} (t-ks) - \mathrm{Tr}\{z\} & & \\ & z & \\ & & z - y + sI_p \end{pmatrix}$$

(blank spaces correspond to zero entries).

Let us verify that the above mapping does represent $\sigma_k$. In other words, we should prove that, first, $t \geq \sigma_k(y)$ if and only if there exist real $s$ and $z \in S_{\{p\}}$

such that $z \geq 0$, $z - y + sI_p \geq 0$ and at the same time $t - ks - \mathrm{Tr}\{z\} \geq 0$ and, second, that the image of $\mathcal{A}$ intersects $\mathrm{int}\, S^+_\mu$.

*"If" part.* Recall that, if $u$, and $v$ are symmetric and $u - v$ is positive-semidefinite, then the eigenvalues of $u$ are not less that those of $v$, both of the sequences being considered in the descent order (this statement immediately follows from Weil's characterization of eigenvalues,

$$\lambda_i(u) = \sup_E \inf\{\langle ux, x\rangle \mid x \in S(E)\},$$

where $E$ runs over the family of all $i$-dimensional subspaces of $\mathbf{R}^p$ and $S(E)$ denotes the unit $\|\cdot\|_2$-sphere in $E$). Now assume that, for a given $(t, y)$, the above $s$ and $z$ do exist. That means that $y \leq z + sI_p$, whence, in view of the previous remark, $\sigma_k(y) \leq \sigma_k(z + sI_p) = \sigma_k(z) + sk$. Since $z \geq 0$, we have $\sigma_k(z) \leq \mathrm{Tr}\{z\}$, so that $\sigma_k(y) \leq \mathrm{Tr}\{z\} + sk \leq t$.

*"Only if" part.* Let $t \geq \sigma_k(y)$. Set $s = \lambda_k(y)$; the $k$ largest eigenvalues of the matrix $y - sI_p$ are nonnegative, while the remaining are nonpositive. We clearly can represent this matrix as $z - w$, where both $z$ and $w$ are positive-semidefinite, the $k$ largest eigenvalues of $z$ being the same as those of $y - sI_p$ and the remaining $p - k$ eigenvalues of $z$ being equal to 0. Now $w = z - y + sI_p \geq 0$ and $\mathrm{Tr}\{z\} = \sigma_k(z) = \sigma_k(y - sI_p) = \sigma_k(y) - sk \leq t - sk$, so that $\mathcal{A}(t, s, y, z) \in S^+_\mu$.

To verify that $\mathrm{Im}\, \mathcal{A}$ intersects $S^+_\mu$, note that $\mathcal{A}(t, s, y, z)$ is positive-definite when, say, $t = 10p$, $s = 0$, $y = 0$, $z = I_p$.

**8.** *Geometrical mean.* Let $q$ be a positive integer and let $n = 2^q$. Consider the function

$$f(y_1, \ldots, y_n) = -(y_1 \ldots y_n)^{1/n} : \mathbf{R}^n_+ \to \mathbf{R}.$$

We already know that this function is SO-representable (see §6.2.3.5), and therefore it is PD-representable (see §6.4.3.1). In fact, the SO-representation given in §6.2.3.5 is a PD-representation as well. Indeed, this representation involves a direct product of $\mathbf{R}_+$ and three-dimensional second-order cones, and a three-dimensional second-order cone

$$\left\{(x, y, z) \in \mathbf{R}^3 \mid z \geq \sqrt{x^2 + y^2}\right\}$$

is nothing but the cone of symmetric positive-semidefinite $2 \times 2$ matrices; the isomorphism is given by the mapping

$$(x, y, z) \mapsto \begin{pmatrix} z - x & y \\ y & z + x \end{pmatrix}.$$

**9.** *Determinant of a symmetric positive-semidefinite matrix.* Let $n \geq 2$ be integer and let $V_n(y) = \mathrm{Det}\, y : S^+_{\{n\}} \to \mathbf{R}$. It is well known that for $l \geq n$ the function $V_n^{1/l}(y)$ is concave. Set $\nu = \rfloor \log_2 n \lfloor$ (the smallest integer not less than $\log_2 n$), $k = k(n) = 2^\nu$, so that $n \leq k < 2n$ and let $v_n(y) = V_n^{1/(2k)}(y)$. To construct a PD-representation of $(-v_n)$, we introduce $k(k+1)/2$ additional variables $\sigma_{ij}$, $1 \le j \le i \le k$ and $k-1$ more additional variables $\tau$; let $\sigma'$ be the collection comprised of $\sigma_{1,1}, \sigma_{2,2}, \dots, \sigma_{k,k}$. Also, let $Y(y)$ be a $k \times k$ block-diagonal matrix with $y$ being $n \times n$ diagonal block and $I_{k-n}$ being the $(k-n) \times (k-n)$ diagonal block. Consider the affine mapping

$$\mathcal{A}(t; y, \sigma, \tau) = \begin{pmatrix} \mathcal{B}(t, \sigma', \tau) & & \\ & I_k & \Delta^T(\sigma) \\ & \Delta(\sigma) & Y(y) \end{pmatrix},$$

where $\Delta(\sigma)$ denotes the lower triangular matrix with the nonzero entries $\sigma_{i,j}$ and the $k \times k$ block $\mathcal{B}(t, \sigma', \tau)$ "represents" the relation $t \ge -(\sigma_{1,1} \ldots \sigma_{k,k})^{1/k}$, i.e., (see §6.4.3.8)

$$\{(t, \sigma') \mid \exists \tau : \mathcal{B}(t, \sigma', \tau) \in S^+_{\{k\}}\} = \{(t, \sigma') \mid \sigma' \ge 0, t \ge -(\sigma_{1,1} \ldots \sigma_{k,k})^{1/k}\}.$$

Note that $\mathcal{A}(\cdot)$ takes values in the space of symmetric $(4k-1) \times (4k-1)$ matrices.

We claim that the set

$$Q \equiv \{(t, y) \mid \exists (\sigma, \tau) : \mathcal{A}(t; y, \sigma, \tau) \text{ is positive-semidefinite}\}$$

coincides with the epigraph of $(-v_n)$, so that $\mathcal{A}$ defines the desired PD-representation of $(-v_n)$. Indeed, the block of $\mathcal{A}$ comprised of $I_k$, $\Delta$, $\Delta^T$, and $Y(y)$ is positive-semidefinite if and only if $\sigma$ is such that $\Delta(\sigma)\Delta^T(\sigma) \le Y(y)$. Combining the latter observation with the above property of $\mathcal{B}$, we conclude that if, given $(t, y)$, we can find $(\sigma, \tau)$ in such a way that $\mathcal{A}(t; y, \sigma, \tau)$ is positive-semidefinite, then $y$ is positive-semidefinite and $t \ge -v_n(y)$. In other words, $Q$ is contained in the epigraph of $(-v_n)$. Conversely, if $(t, y)$ satisfies the latter relations, then we can define $\sigma$ in such a way that $\sigma_{i,i}$ are nonnegative and $Y(y) = \Delta(\sigma)\Delta^T(\sigma)$, so that $\operatorname{Det} y = (\sigma_{1,1} \ldots \sigma_{k,k})^2$ and therefore $t \ge -(\sigma_{1,1} \ldots \sigma_{k,k})^{1/k}$. The latter relation, by construction of $\mathcal{B}$, means that there exists $\tau$ such that $\mathcal{B}(t, \sigma', \tau)$ is positive-semidefinite, so that $\mathcal{A}(t; y, \sigma, \tau)$ is positive-semidefinite. Thus, $Q$ contains the epigraph of $(-v_n)$, which, combined with the (already proved) opposite inclusion, means that $Q$ coincides with the epigraph of $(-v_n)$. The fact that $\operatorname{Im} \mathcal{A}$ intersects $\operatorname{int} S^+_\mu$ is evident.

**10.** *General fractional-quadratic function.* Let

$$\tilde{f}(x, X) = x^T X^{-1} x : \mathbf{R}^n \times \{\operatorname{int} S^+_n\} \to \mathbf{R};$$

passing to the lower semicontinuous closure of this function, we obtain the functional element $(G = \mathbf{R}^n \times S^+_n, f)$, where, as it is easily seen,

$$f(x, X) = \sup\{2x^T y - y^T X y \mid y \in \mathbf{R}^n\}$$

(compare with §6.2.3.6). We can verify immediately that the functional element $(G, f)$ (we dealt with this "general fractional-quadratic function" in §5.4.7) admits the following PD-representation:

$$\mathsf{G}(G, f) = \mathcal{A}^{-1}(S^+_{\{n+1\}}), \qquad \mathcal{A}(t, x, X) = \begin{pmatrix} t & x^T \\ x & X \end{pmatrix}.$$

Recall that the restriction of $f$ onto the cone

$$\{X = A_0 + s_1 A_1 + \cdots + s_k A_k \mid s_i \geq 0\}$$

generated by *positive-semidefinite symmetric* $A_0, \ldots, A_k$ with positive-definite sum admits the SO-representation given in item 6 of §6.2.3.

### 6.4.4 Applications

**1.** *Linearly and quadratically constrained problems.* As we have already seen, a convex quadratic form $f$ admits a PD-representation of the structure $\{r+2\}$, $r$ being the rank of $f$. Thus, the scheme of §5.2.5.B allows us to find a *pd*-reformulation of any quadratically constrained convex quadratic problem. The size $\mid \mu \mid$ of the resulting structure $\mu$ equals to $\sum_{i=0}^{m}\{r_i + 2\}$, where $m$ denotes the number of constraints and $r_i$ is rank of the $i$th constraint (for $i = 0$, rank of the objective). Of course, the parameter $\mid \mu \mid$ of the resulting barrier is worse than the one (namely, $2(m+1)$) corresponding to the conic reformulation of a quadratically constrained problem (see §6.2.2).

**2.** *Minimization of the matrix norm.* Although this problem has already been studied (see §6.3.3), it is reasonable to note that the epigraph of the matrix norm on the space $L_{k,l}$ of $k \times l$ matrices admits a PD-representation (see §6.4.3.5), so that the problem studied in §6.3.3 can also be reformulated as a *pd*-problem.

**3.** *Minimization of the largest eigenvalue of a symmetric matrix. Lovasz capacity number of a graph.* Let $A(y) = A_0 + \sum_{i=1}^{p} A_i y_i$ be an affine function of $y \in \mathbf{R}^p$ taking values in $S_{\{q\}}$ and let $\lambda_{\max}(A)$ be the largest eigenvalue of a symmetric matrix $A$. Assume that we desire to minimize the function $\lambda_{\max}(A(y))$ over $y$. The PD-representation of $\lambda_{\max}(\cdot)$ (see §6.4.3.6) immediately implies the *pd*-reformulation of the latter problem,

$$\text{minimize } f(u) = t \text{ by choice of } u = (t, y) \quad \text{s.t. } tI_q - A(y) \in S^+_{\{q\}}.$$

The problem of minimizing the largest eigenvalue of a symmetric matrix linearly depending on the control vector occurs, for example, in connection with computation of the Lovasz capacity number of a graph. Let $\Gamma$ be a graph with the set of vertices $V = \{1, 2, \ldots, N\}$ and the set of arcs $E$. Consider the following characteristics of $\Gamma$:

$\alpha(\Gamma)$—the maximal cardinality of independent subsets of $V$, i.e., the subsets in which no pairs of vertices are connected by an arc in $\Gamma$;

$\sigma(\Gamma)$—the Shannon capacity number of $\Gamma$, defined as follows. Regard $V$ as an alphabet and let $\Gamma^k$ be the graph with the set of vertices being the set of all $k$-letter words in the alphabet; a pair of these words is adjacent in $\Gamma^k$ if and only if, for each $i \leq k$, the $i$th letters of these words (vertices of $\Gamma$) are adjacent in $\Gamma$. By definition, $\theta(\Gamma) = \lim_{k\to\infty} \left(\alpha(\Gamma^k)\right)^{1/k}$;

$\chi(\Gamma)$—the size of minimal vertex coloring in the complement to $\Gamma$ graph $\Gamma'$ (the minimal number of sets in partitioning the nodes of $\Gamma'$ under the requirement that no vertices of different sets are adjacent in $\Gamma'$).

The characteristics $\alpha(\Gamma)$ and $\chi(\Gamma)$ are very important for many combinatorial problems; unfortunately, to compute them for an arbitrary graph, it is an NP-hard problem. It is also very difficult to compute the Shannon capacity number (it is easily seen that it is an upper bound for $\alpha(\Gamma)$). Lovasz [Lo 79] suggested a nontrivial upper bound $\theta(\Gamma)$ for $\sigma(\Gamma)$, which can be effectively computed. One of equivalent definitions of $\theta(\Gamma)$ is as follows. Let us set into correspondence to each (indirected) arc $\gamma$ of the graph $\Gamma$ its own control variable $y_\gamma$ and consider the following function $A(y)$ of the variables $y = \{y_\gamma,\, \gamma \in E\}$, taking values in the space of symmetric $N \times N$ matrices: The $ij$th entry of $A(y)$ is 1 if either $i = j$ or the vertices $i$ and $j$ are not adjacent; otherwise, they are linked by an arc $\gamma$, and the $ij$th entry of $A(y)$ is, by definition, $y_\gamma$. Now, by definition,

$$\theta(\Gamma) = \min_y \lambda\,(A(y)).$$

Lovasz proves that

$$\chi(\Gamma) \geq \theta(\Gamma) \geq \sigma(\Gamma)\ (\geq \alpha(\Gamma)).$$

Thus, the quantity $\theta(\Gamma)$ that can be computed by solving the associated *pd*-problem is a bound (and, in many interesting cases, very "strong") one for important characteristic numbers of a graph.

**4.** *Dual bounds in Boolean programming.* Consider a quadratic programming problem with equality constraints

$$(\mathsf{QE}): \qquad \text{minimize } f_0(u) \quad \text{s.t. } u \in \mathbf{R}^l,\ f_i(u) = 0,\ 1 \leq i \leq q,$$

where all the functions $f_i$, $0 \leq i \leq q$ are quadratic forms.

This is a nonconvex and a very difficult problem; indeed, each mathematical programming problem with polynomial objective and constraints is easily reduced to (QE) (to represent, say, the monomial $x_1^3 x_2^2$ via quadratic equalities, it suffices to introduce variables $x_{1,2}$, $x_{1,3}$, $x_{2,2}$, $x_{1,3;2,2}$, and constraints $x_{1,2} = x_1^2$, $x_{1,3} = x_1 x_{1,2}$, $x_{2,2} = x_2^2$, $x_{1,3;2,2} = x_{1,3}x_{2,2}$ $(= x_1^3 x_2^2)$). Boolean constraint $x \in \{0; 1\}$ also can be represented via the quadratic equality $x^2 = x$. Thus, (QE) covers "almost everything."

In many cases (say, in the branch-and-bound algorithms as applied to (QE)), it is important to evaluate from below the optimal value QE* of (QE) (as usual, QE* $= +\infty$, if (QE) is not consistent; otherwise, QE* is the infimum of the objective values over the feasible set of (QE)). Shor [ShD 85] suggested a lower bound for QE* based on the duality theory, namely the following.

Consider the following Lagrange function for (QE):

$$L(u, y) \equiv f_0(u) + \sum_{i=1}^{q} y_i f_i(u) = -2b^T(y)u + u^T A(y) u + d(y).$$

Herein $b(y)$, $A(y)$, $d(y)$ are linear in $y$ functions taking values, respectively, in $\mathbf{R}^l$, $S_{\{l\}}$, and $\mathbf{R}$. It is clear that the quantity

$$\phi(y) = \inf\{L(u,y) \mid u \in \mathbf{R}^l\} \geq -\infty$$

for each $y$ does not exceed $\mathsf{QE}^*$, so that the quantity

$$\phi^* = \sup\{\phi(y) \mid y \in \mathbf{R}^q\} \in [-\infty, +\infty]$$

is a lower bound for $\mathsf{QE}^*$.

The advantage of this lower bound is that, to compute it, it suffices to solve a convex problem (even a *pd* one), and therefore this bound can be found effectively. At the same time, for many real world problems, this bound is reported to be "strong" enough.

Let us present a *pd*-reformulation of the problem of maximizing $\phi$. Set $\mu = \{l+1\}$ and define the function $\alpha(t,y) = Au+b$, $u = (t,y)$ ($t$ is an additional scalar variable), taking values in the space of $(l+1)\times(l+1)$ symmetric matrices as follows:

$$\alpha(t,y) = \begin{pmatrix} d(y)-t & b^T(y) \\ b(y) & A(y) \end{pmatrix}.$$

It is easily seen (compare with §6.4.3.10) that the matrix $\alpha(t,y)$ is positive semidefinite if and only if $t \leq \phi(y)$. Therefore computing $\phi^*$ is the same as solving the following problem

$$(\mathsf{QE}'): \qquad \text{minimize } f(u) = -t \text{ by choice of } u = (t,y) \quad \text{s.t. } x \equiv \alpha(u) \in S_\mu,$$

where $\mu = \{l+1\}$.

**5.** *Inscribing the maximal ellipsoid into a convex polytope.* This problem (studied in detail in §6.5, although *not* as a *pd*-problem) is as follows. Given a convex polytope

$$Q = \{x \in \mathbf{R}^n \mid a_i^T x \leq b_i,\ i = 1,\dots,m\},$$

we should find an ellipsoid

$$W(B,u) = \{x = By + u \mid \| y \|_2 \leq 1\}$$

of the maximal volume contained in $Q$. The control variables here are $B$ (which should be a positive-definite symmetric $n \times n$ matrix) and $u \in \mathbf{R}^n$.

We assume that $Q$ is bounded with a nonempty interior and that $a_i$ are nonzeros. Under these assumptions, the problem evidently is solvable.

To find a *pd*-reformulation of the problem, note that an ellipsoid $W(B,u)$ defined by a positive-semidefinite matrix $B$ is contained in $Q$ if and only if $(B,u)$ satisfies the set of constraints

$$\| Ba_i \|_2 + a_i^T u - b_i \leq 0, \qquad i = 1,\dots,m;$$

instead of maximizing the volume of $W(B,u)$, i.e., of $\operatorname{Det} B$, we can minimize the function

$$V_n(B) = -(\operatorname{Det} B)^{1/(2k)}, \quad k = 2^{\rfloor \log_2 n \lfloor}.$$

As we already know (§6.4.3.1), the Euclidean norm of a vector admits a PD-representation: There exists an affine mapping

$$\mathcal{B}(t,v) : \mathbf{R}^{n+1} \to S_{\{n+1\}}$$

such that

$$\{(t,x) \mid t \geq \| x \|_2\} = \mathcal{B}^{-1}(S^+_{\{n+1\}}).$$

In particular, the affine mapping

$$\mathcal{B}_i(B,u) = \mathcal{B}(b_i - a_i^T u, Ba_i) : S_{\{n\}} \times \mathbf{R}^n \to S_{\{n+1\}}$$

represents the constraint

$$\| Ba_i \|_2 + a_i^T u - b_i \leq 0$$

(i.e., a pair comprised of an $n \times n$ matrix $B$ and a vector $u \in \mathbf{R}^n$ satisfies the latter constraint if and only if $\mathcal{B}_i(B,u) \in S^+_{\{n+1\}}$). In §6.4.3.9, it was explained that an appropriate affine mapping $\mathcal{C}(t; B, \sigma, \tau)$ taking values in the space of symmetric $(4k-1) \times (4k-1)$ matrices ($\sigma$ and $\tau$ are additional vectors of variables of the dimensions $k(k+1)/2$, $k-1$, respectively) represents the epigraph of the function $V_n(\cdot)$ in the sense that $B$ is symmetric positive-semidefinite and $t \geq V_n(B)$ if and only if there exist $\sigma$ and $\tau$ such that $\mathcal{C}(t; B, \sigma, \tau)$ is positive-semidefinite.

Now let

$$\mu = \{4k-1, \overbrace{n+1, \ldots, n+1}^{m \text{ times}}\}$$

and let $\mathcal{A}(t; B, u, \sigma, \tau)$ be the affine mapping from the space of corresponding variables ($t \in \mathbf{R}$, $B \in S_{\{n\}}, u \in \mathbf{R}^n, \sigma \in \mathbf{R}^{k(k+1)/2}, \tau \in \mathbf{R}^{k-1}$) into $S_\mu$ defined as follows: The unique $(4k-1) \times (4k-1)$ diagonal block of $\mathcal{A}(t; B, u, \sigma, \tau)$ is $\mathcal{C}(t; B, \sigma, \tau)$, and the remaining $m$ of $(n+1) \times (n+1)$ diagonal blocks are $\mathcal{B}_i(B,u)$. From the above remarks, it follows immediately that the problem

$$\text{minimize } t \text{ by choice of } (t, B, u, \sigma, \tau) \quad \text{s.t. } \mathcal{A}(t; B, u, \sigma, \tau) \in S^+_\mu$$

is the desired *pd*-reformulation of the problem of finding the maximal volume ellipsoid inscribed into the polytope $Q$.

**6.** *Applications in control theory.* Semidefinite programming has many applications in modern control theory (see [BB 90], [BBr 91], [BBK 89], [BG 92], [BY 89], [DPZ 91], [Do 82], [FN 91], [FT 86], [FT 88], [FT 91], [FTD 91], [GB 86], [KR 91], and especially [BGFB 93]). Let us present an example arising from Lyapunov stability analysis of systems subject to uncertainty (see Boyd and El Ghaoui [BG 92]). In many cases, such a system can be described by a *differential inclusion*

$$x'(t) \in X(x(t)), \tag{6.4.1}$$

defined by a given multivalued mapping $X(x)$ ($x \in \mathbf{R}^n$, $X(x) \subseteq \mathbf{R}^n$). The problem is to bound from above asymptotic behaviour of the trajectories of (6.4.1). To this purpose, we might look for a quadratic Lyapunov function $x^T Lx$ for the inclusion, $L$ being a positive definite symmetric matrix. If

$$2x^T Ly \le \lambda x^T Lx \tag{6.4.2}$$

for certain $\lambda$ and all $x \in \mathbf{R}^n$, $y \in X(x)$, then clearly

$$x^T(t)Lx(t) \le \exp\{\lambda t\}\{x^T(0)Lx(0)\} \tag{6.4.3}$$

for all solutions to (6.4.1).

We henceforth restrict ourselves to the case when

$$\{X_1 x; \dots, X_k x\} \subseteq X(x) \subseteq \operatorname{conv}\{X_1 x; \dots; X_k x\}$$

for certain given matrices $X_i$ (this assumption is satisfied in many applications in control theory). Then, checking whether (6.4.2) is satisfied for *a given* $\lambda$ is the same as checking whether the system of "linear matrix inequalities"

$$\lambda L - (X_i^T L + LX_i) \in S_n^+, \qquad i = 1, \dots, k; L - I_n \in S_n^+ \tag{6.4.4}$$

with the unknown $L \in S_n$ is solvable. The latter problem is, of course, a *pd*-problem with trivial objective.

A more interesting question is to find *the best* (with the smallest possible $\lambda$) quadratic Lyapunov function for (6.4.1). This is, of course, the same as solving the following generalized linear-fractional problem (see §4.4) involving the cone of positive-semidefinite matrices:

$$\text{minimize } \lambda \quad \text{s.t. } \lambda B(L) - A(L) \in S^+_{\underbrace{\{n, \dots, n\}}_{n \text{ times}}}, \ C(L) \in S^+_{\{n\}},$$

where $B(L)$ is the block-diagonal matrix with $k$ ($n \times n$) diagonal blocks, each of them being $L$, $A(L)$ is the block-diagonal matrix with the diagonal blocks $X_i^T L + LX_i$, $i = 1, \dots, k$, and $C(L) = L - I_n$. To solve the resulting problem, we can use the projective method (§4.4) or reduce it via dichotomy to a "small" series of the usual *pd*-problems of the type (6.4.4).

## Concluding remarks

We have presented a number of interesting convex problems that admit a *pd*-reformulation. For some of them (minimizing the largest eigenvalue, Shor's bounding), *pd*-formulation seems to be very natural, while, for some others (e.g., for quadratically constrained quadratic programs), *pd*-formulation might look too sophisticated. Nevertheless, we note some advantages of a *pd*-reformulation.

**A.** *Possibility to use the generalized combined volumetric barrier.* An important (at least from the theoretical viewpoint) consequence of the *pd*-representation (P′) of a convex problem is the possibility to provide the feasible set of (P′) with the generalized combined volumetric barrier (§5.5). Recall that the value of the parameter of the latter barrier is $O(1)(n \mid \mu \mid)^{1/2}$, and this quantity can be essentially less than the values associated with the standard barriers. For example, the parameter of the "standard" barrier $-\sum_{i=1}^m \ln(-f_i)$ for the domain $G$ in $\mathbf{R}^n$ defined by $m$ convex quadratic constraints $f_i \leq 0$ is $m$. The parameter of the barrier induced by the standard conic representation of $G$ (see §3.2) is $2m$. The parameter of the standard barrier associated with the *pd*-conic representation of $G$ is $M = \sum_{i=1}^m \{\text{Rank}\,\{f_i\} + 1\}$, which, of course, is worse than $m$. The parameter $O(1)(Mn)^{1/2}$ of the combined volumetric barrier induced by the latter representation, however, can be much smaller that $m$.

**B.** *Nice anticipated behaviour of the potential reduction methods.* In many experiments with "large-step" versions of the Karmarkar method as applied to LP problems, it was found that the number $N$ of Newton steps required to solve an LP problem to a fixed accuracy is "almost independent" on the (larger) size of the problem $m$: The dependence looks like $N = O(\ln m)$. Recall that the theoretical upper bound is essentially worse: $N \leq O(m)$. Our own experiments with the projective method as applied to *pd*-problems also demonstrate that the number of steps is almost independent of the size of the problem. Although no rigorous justification of this phenomenon is known, there is certain plausible explanation of it, which appears as follows. Consider a step of, say, the Karmarkar method as applied to (P) (see §6.4.1) and let $\bar{x}$ be the strictly feasible solution that is updated at the step into a new solution $\bar{x}'$. Let us perform *scaling*, i.e., instead of variables $x$ taking values in the space $S_\mu$, let us use variables $\xi = \bar{x}^{-1/2} x \bar{x}^{-1/2}$. It is easily seen that the step of the method of Karmarkar in the $\xi$-variables looks precisely as it looked in the initial variables, excluding the fact that the current strictly feasible solution now is the unit matrix $I$ (and, of course, the feasible plane and the objective now should be subject to certain transformation).

The step looks as follows. We are given a positive-semidefinite matrix $\sigma \in S_\mu$ (the updated objective divided by its value at $I$) and a linear subspace $L$ in $S_\mu$ ($L$ is the updated homogeneous feasible subspace of the problem intersected with the null space of the gradient of the barrier at $I$). It is known that

(a) $\text{Tr}\{\sigma\} = 1$;

(b) $L$ is contained in the subspace formed by matrices with zero trace;

(c) The set $(I + L) \bigcap S_\mu^+$ contains a matrix $\xi^*$ with $\text{Tr}\{\xi^* \sigma\} = 0$ ($\xi^*$ is the updated optimal solution to the problem; recall that, in Karmarkar's setting, the optimal value is zero).

At the step, we compute the orthoprojection $\chi$ of $\sigma$ onto $L$ (with respect to the Euclidean structure on $S_\mu$ defined by the Hessian of the barrier $F(\xi) =$

$-\ln(\mathrm{Det}\,\xi)$ at $I$; this is simply the standard Euclidean structure on $S_\mu$), and then we minimize the potential

$$v(t) = F(I - t\chi) + m \ln \langle \sigma, I - t\chi \rangle \,, \ m = \mid \mu \mid$$

by choice of $t > 0$ subject to the constraint that $I - t\chi$ should be positive-definite. The next strictly feasible solution to (P) is an appropriate normalization of the matrix $I - t^*\chi$, where $t^*$ is the minimizer of the potential $v$. The amount $\Delta$ by which the Karmarkar potential function is decreased at the step is precisely $v(0) - v(t^*)$.

As we have mentioned, $\langle \sigma, I \rangle = 1$, and, since $\chi$ is the orthoprojection of $\sigma$ onto $L$, we have $\langle \sigma, \chi \rangle = \langle \chi, \chi \rangle$ . Thus,

$$\begin{aligned} v(t) &= -\ln \mathrm{Det}\{I - t\chi\} + m \ln(1 - t\,\langle \chi, \chi \rangle) \\ &= -\ln\left(\prod_{i=1}^{m}(1 - ts_i)\right) + m \ln\left(1 - t \parallel s \parallel_2^2\right), \end{aligned}$$

where $s = (s_1, \ldots, s_m)$ is the vector comprised of the eigenvalues of $\chi$. What we know about $s$ is

($\alpha$) $\sum_{i=1}^m s_i = 0$ (see (b)) and

($\beta$) $\parallel s \parallel_\infty \geq 1/m$ (indeed, in view of (c), we have $\langle \xi^*, \sigma \rangle = 0$ for certain positive-semidefinite $\xi^* = I - \delta$, $\delta \in L$. Since $\langle \sigma, I \rangle = 1$ (see (a)), it follows that $\langle \delta, \sigma \rangle = 1$. Since $\delta \in L$ and $\chi$ is the orthoprojection of $\sigma$ onto $L$, we have $\langle \delta, \chi \rangle = \langle \delta, \sigma \rangle = 1$. Thus, $\langle \chi, I - \xi^* \rangle = 1$ or, in view of ($\alpha$), $\langle \chi, \xi^* \rangle = -1$. The latter relation, in view of positive semidefiniteness of $\xi^*$ and the relation $\mathrm{Tr}\,\xi^* = m$ (see (c) and (b)) implies ($\beta$).

Now, for $0 < t <\parallel s \parallel_\infty^{-1}$, we have

$$\begin{aligned} v(t) &\leq -\sum_{i=1}^{m}\sum_{j=1}^{\infty}\frac{1}{j}(ts_i)^j - tm \parallel s \parallel_2^2 = -\sum_{i=1}^{m}\sum_{j=2}^{\infty}\frac{1}{j}(ts_i)^j - tm \parallel s \parallel_2^2 \\ &\leq \frac{\parallel s \parallel_2^2}{\parallel s \parallel_\infty^2}\sum_{j=2}^{\infty}\frac{1}{j}(t \parallel s \parallel_\infty)^j - tm \parallel s \parallel_2^2 \\ &= -\frac{\parallel s \parallel_2^2}{\parallel s \parallel_\infty^{-2}}\{\ln(1 - t \parallel s \parallel_\infty) + t \parallel s \parallel_\infty\} - tm \parallel s \parallel_2^2 \equiv \gamma(t) \end{aligned}$$

(we have considered ($\alpha$)). The minimizer of $\gamma$ on $0 < t <\parallel s \parallel_\infty^{-1}$ is the point $\tau = m/(1 + m \parallel s \parallel_\infty)$, and

$$\gamma(\tau) = -\left(\frac{\parallel s \parallel_2}{\parallel s \parallel_\infty}\right)^2 \{m \parallel s \parallel_\infty - \ln(1 + m \parallel s \parallel_\infty)\} \leq -(1 - \ln 2)\left(\frac{\parallel s \parallel_2}{\parallel s \parallel_\infty}\right)^2$$

(the latter inequality follows from ($\beta$)). Since $v(0) = 0$ and $v(\tau) \leq \gamma(\tau)$, we conclude that a large step in the method of Karmarkar decreases the potential at least by the quantity

$$\Delta(s) = \kappa \left( \frac{\| s \|_2}{\| s \|_\infty} \right)^2, \qquad \kappa = 1 - \ln 2;$$

here $s$ is a certain nonzero vector from $\mathbf{R}^m$ (depending on the data and the step) with zero mean of the coordinates.

The worst-case efficiency estimate for the method of Karmarkar follows from the fact that $\Delta(s)$ is always $\geq \kappa$ (evident). At the same time, the "typical" value of the ratio $(\| s \|_2 / \| s \|_\infty)^2$ for a $n$-dimensional vector $s$ is much larger than the worst-case value 1. For example, let $s$ be a random vector taking values in the space of $m$-dimensional vectors with zero mean and let the distribution of the direction of $s$ be uniform on the corresponding sphere. Then "typical" value of $(\| s \|_2 / \| s \|_\infty)^2$ is $O(m(\ln m)^{-1})$ (more precisely, the probability of the event $\{(\| s \|_2 / \| s \|_\infty)^2 > O(1)m(\ln m)^{-1}\}$ under an appropriate choice of an *absolute* constant $O(1)$ tends to 1 as $m \to \infty$). If the decreasing of the potential at each step of the method of Karmarkar were of the above "typical" order $m/(\ln m)$, then, to improve the accuracy of the current approximate solution by an absolute constant factor, it would be sufficient to perform $O(\ln m)$ steps (instead of $O(m)$ steps prescribed by the worst-case analysis). Of course, it seems to be impossible to prove something rigorous here: The directions $s$ occurring at the sequential steps of the method heavily depend on each other, and there is no hope to provide the directional symmetry of all of them by a consistent choice of a probabilistic distribution on the set of problem instances.

Nevertheless, in view of the above analysis of the "anticipated behaviour" of the Karmarkar method as applied to (P), the aforementioned empiric phenomenon does not seem too surprising. Note that the anticipated behaviour of the projective method as applied to *pd*-problems is the same as that of the method of Karmarkar.

## 6.5 Extremal ellipsoids

### 6.5.1 Inscribed ellipsoids: Geometric formulation of the problem

In this section, we study in detail the following geometric problem mentioned in §6.4.4.5:

$\mathcal{P}(\mathbf{K})$ : Given a polytope
$\mathbf{K} = \{x \in \mathbf{R}^n \mid a_i^T x \leq b_i, 1 \leq i \leq m\}$,
find the ellipsoid of maximal volume contained in $\mathbf{K}$.

This problem arises in connection with the *inscribed ellipsoid method* (IEM) (see Khachiyan, Tarasov, and Erlikh [KhTE 88]) for convex nondifferentiable optimization. The method minimizes a convex function $f$, say, over an $n$-dimensional cube to a relative accuracy $\nu$ in $O(n \ln(n/\nu))$ steps ($\equiv$ evaluations of $f$ and $f'$). Note that this number of steps cannot be reduced (for each $\nu < \frac{1}{2}$) more than by an absolute constant factor (for precise formulation of the latter statement, see [NYu 79]). Each step of the IEM requires finding an $\varepsilon$-solution to the above geometrical problem (it is necessary to find an inscribed ellipsoid of the volume not less than

$$\{\text{maximum volume of the inscribed ellipsoids}\}\, e^{-\varepsilon},$$

$\varepsilon$ being an appropriate absolute constant). In [KhTE 88], the latter problem is solved by the ellipsoid method, which requires nearly $O(m^8)$ arithmetic operations per step. It turns out that the barrier method from Chapter 2 reduces this amount to $O(m^{4.5}\ln m)$. In this section, we describe the implementation of the barrier method.

We study $\mathcal{P}(\mathbf{K})$ under the following assumptions.

(I) $\mathbf{K}$ is presented by the list of corresponding linear inequalities;

(II) $\mathbf{K}$ is bounded with a nonempty interior, and $a_i \neq 0$, $1 \le i \le m$;

(III) $\mathbf{K}$ contains the unit Euclidean ball $V$ centered at 0 and is contained in the concentric ball $W$ of a given radius $\mathcal{R}$.

Note that all these assumptions are satisfied in the case of IEM, where, without loss of generality, we can take $\mathcal{R} = 10n$ and maintain with the aid of a simple restart strategy the inequality $m \le O(n \ln n)$.

### 6.5.2 Algebraic formulation of the problem

We can reformulate $\mathcal{P}(\mathbf{K})$ as follows. Let $L_n$ be the space of real $n \times n$ matrices and $L_n^+$ be the domain in $L_n$ formed by matrices of positive determinant. Each ellipsoid in $\mathbf{R}^n$ can be represented as

$$W(B,u) = \{x = By + u \mid \| y \|_2 \le 1\},$$

where $u \in \mathbf{R}^n$ is the center of the ellipsoid and $B \in L_n^+$.

Note that, under an appropriate choice of the volume unit, the volume $|\cdot|$ of an ellipsoid $W(B,u)$ is

$$| W(B,u) | = \operatorname{Det} B,$$

and the inclusion $W(B,u) \subset \mathbf{K}$ holds if and only if $(B,u)$ satisfies the system of inequalities

$$\mathcal{Q}(a^m, b^m): \quad \| B^T a_i \|_2 \le b_i - a_i^T u, \ 1 \le i \le m$$

($a^m$ denotes the collection of vectors $a_i$, $1 \le i \le m$, and $b^m$ denotes the collection of numbers $b_i$, $1 \le i \le m$).

Let

$$\mathcal{V}(B,u) = -\ln \operatorname{Det} B : L_n^+ \times \mathbf{R}^n \to \mathbf{R}.$$

Problem $\mathcal{P}(\mathbf{K})$ can be reformulated as follows:

$$\mathcal{P}(\mathbf{K}): \qquad \begin{array}{l} \text{minimize } \mathcal{V}(B,u) \text{ by choice of } (B,u) \in L_n^+ \times \mathbf{R}^n \\ \text{s.t. constraints } \mathcal{Q}(a^m, b^m). \end{array}$$

An ellipsoid $W(B,u)$ is called *$\varepsilon$-optimal* if it is contained in $\mathbf{K}$ and its volume is not less than $V^* \mathrm{e}^{-\varepsilon}$, where $V^*$ is the maximal volume of ellipsoids contained in $\mathbf{K}$.

*$\mathcal{P}(\mathbf{K})$ as a convex programming problem.* The quantities $B$ and $u$ in the representation of an ellipsoid in the form $W(B,u)$ are not uniquely defined. If $U$ is an orthogonal $n \times n$ matrix, then

$$W(B,u) = W(BU,u).$$

Hence we can assume the $B$-component of the variable $z = (B,u)$ involved in $\mathcal{P}(\mathbf{K})$ to be symmetric positive-definite. This additional restriction leads to a convex programming problem. In fact, there are many convex programming problems equivalent to $\mathcal{P}(\mathbf{K})$. Let us describe these.

Let $S_n$ be the space of symmetric real $n \times n$ matrices and $S_n^\circ$ be the interior of the cone $S_n^+$ of positive-semidefinite $n \times n$ matrices.

Let $\Lambda \in L_n^+$ and let $\Lambda\mathcal{Q}(a^m, b^m)$ denote the following system of inequalities with unknowns $(B,u) \in L_n \times \mathbf{R}^n$ :

$$\| B^T \Lambda^T a_i \|_2 \le b_i - a_i^T u, \qquad 1 \le i \le m.$$

Consider the problem

$$\mathcal{P}(\Lambda,\mathbf{K}): \qquad \begin{array}{l} \text{minimize } \mathcal{V}(z) \text{ by choice of } z = (B,u) \in S_n^+ \times \mathbf{R}^n \\ \text{s.t. } \Lambda\mathcal{Q}(a^m, b^m). \end{array}$$

Problems $\mathcal{P}(\mathbf{K})$ and $\mathcal{P}(\Lambda,\mathbf{K})$ are clearly consistent. Let the optimal values of their objectives be $v^*$, $v_\Lambda^*$, respectively, and let

$$\Delta(z) = \mathcal{V}(z) - v^* \quad \text{for a } \mathcal{P}(\mathbf{K})\text{-feasible point } z,$$

$$\Delta_\Lambda(z) = \mathcal{V}(z) - v_\Lambda^*$$

for a $\mathcal{P}(\Lambda,\mathbf{K})$-feasible point $z$.

**Lemma 6.5.1** *Let $\Lambda \in L_n^+$. If $z = (B,u)$ is $\mathcal{P}(\Lambda,\mathbf{K})$-feasible, then $\Lambda z \equiv (\Lambda B, u)$ is $\mathcal{P}(\mathbf{K})$-feasible and*

$$\Delta(\Lambda z) = \Delta_\Lambda(z). \tag{6.5.1}$$

**Proof.** $\Lambda z$ clearly is feasible for $\mathcal{P}(\mathbf{K})$. To verify (6.5.1), note that

$$\mathcal{V}(\Lambda z) - \mathcal{V}(z) = -\ln \operatorname{Det} \Lambda$$

does not depend on $z$. Thus, each point that is feasible for $\mathcal{P}(\Lambda, \mathbf{K})$ corresponds to a point that is feasible for $\mathcal{P}(\mathbf{K})$, the values of the objectives at the points being equal to each other within a constant term $(-\ln \operatorname{Det} \Lambda)$. In particular,

$$v^* - v^*_\Lambda \leq -\ln \operatorname{Det} \Lambda. \tag{6.5.2}$$

To prove (6.5.1), it suffices to show that the latter inequality is an equality. Let $(B^*, u^*)$ be the solution to $\mathcal{P}(\mathbf{K})$. The polar factorization of the matrix $(\Lambda^{-1}B^*)$ allows us to represent $B^*$ as $B^* = \Lambda BU$, $U$ being orthogonal and $B$ being symmetric positive-definite. Since $(B^*, u^*)$ satisfies the constraints $\mathcal{Q}(a^m, b^m)$, the point $z^+ = (B, u^*)$ satisfies the constraints $\Lambda\mathcal{Q}(a^m, b^m)$, so that $z^+$ is feasible for $\mathcal{P}(\Lambda, \mathbf{K})$. Since $U$ is orthogonal, we have

$$\begin{aligned} v^* = \mathcal{V}(B^*, u^*) = \mathcal{V}(\Lambda B, u^*) = \mathcal{V}(B, u^*) - \ln \operatorname{Det} \Lambda = \mathcal{V}(z^+) - \ln \operatorname{Det} \Lambda \\ \geq v^*_\Lambda - \ln \operatorname{Det} \Lambda; \end{aligned}$$

thus $v^* - v^*_\Lambda \geq -\ln \operatorname{Det} \Lambda$. This inequality, combined with (6.5.2), proves the lemma. □

The lemma shows that solving $\mathcal{P}(\mathbf{K})$ is the same as solving any of the convex programming problems $\mathcal{P}(\Lambda, \mathbf{K})$.

### 6.5.3 Path-following method

Let us discuss how to solve $\mathcal{P}(\Lambda, \mathbf{K})$ by the path-following barrier-generated method. Let $E = S_n \times \mathbf{R}^n$. This space is provided by the standard Euclidean structure, given by the scalar product

$$((B, u), (C, v)) = \operatorname{Tr}\{BC\} + u^T v.$$

Let $G(\Lambda)$ denote the following closure of the feasible region of $\mathcal{P}(\Lambda, \mathbf{K})$ :

$$G(\Lambda) = \{z = (B, u) \in E \mid B \in S_n, \| B\Lambda^T a_i \|_2 \leq b_i - a_i^T u,\ 1 \leq i \leq m\}.$$

It is clear that $G(\Lambda)$ is a bounded closed convex domain in $E$ and that the function

$$F^\Lambda(z) = 2\mathcal{V}(z) - \sum_1 \ln\{(b_i - a_i^T u)^2 - \| B(\Lambda^T a_i) \|_2^2\} \equiv 2\mathcal{V}(z) + \Phi^\Lambda(z)$$

$(z = (B, u))$ is a $\vartheta$-self-concordant barrier for $G(\Lambda)$, where

$$\vartheta = 2n + 2m < 4m$$

(see Propositions 5.4.3, and 5.4.5; since $\mathbf{K}$ is a compact set, we have $n < m$). Note that $\mathcal{V}(\cdot)$ is 1-compatible with this barrier (Proposition 3.2.1(ii)). By condition (III), the point $z = ((1/2)I_n, 0)$ is a good starting point for $\mathcal{P}(\Lambda, \mathbf{K})$; thus, the problem can be solved by the basic barrier method associated with the barrier $F^\Lambda$. It can be verified that, to find an $\varepsilon$-solution to the problem, it suffices to perform no more than

$$O\left(m^{1/2}\ln\frac{\mathcal{R}m}{\varepsilon}\right)$$

steps of the preliminary and the main stages of the method. Each of these steps requires us to form and to solve some linear system with $\dim E$ ($\leq O(m^2)$) equations and unknowns. It is easy to show that the standard implementation of a step costs no more than $O(m^6)$ arithmetic operations (in fact, $O(m^5)$ operations if the conjugate gradient method is applied, since the matrix of the resulting system is sparse). Thus, the straightforward application of the barrier method to $\mathcal{P}(\mathbf{K})$ finds an $\varepsilon$-solution at the cost $O(m^{5.5}\ln(\mathcal{R}m/\varepsilon))$. Our aim now is to demonstrate that the intrinsic symmetry of the problem allows us to reduce the cost by a factor of $O(m)$.

The idea of the speed-up can be easily described for the main stage, where we must compute the Newton directions for the functions $F_t^\Lambda(z) = (2+t)\mathcal{V}(z)+\Phi^\Lambda(z)$. This computation (at a "general" point $z$) costs $O(m^5)$ operations (for simplicity, we replace the powers of $n$ by the same powers of $m > n$). At a "special" $z$, namely, at $z = (I_n, u)$, however, the computation costs only $O(m^4)$ operations. So we would prefer to deal only with "special" points. Namely, assume we have performed $i$ iterations of the main stage and have a value $t_i$ of the penalty parameter and an approximate solution $(B_i, u_i)$, which is $\mathcal{P}(\mathbf{K})$-feasible.

Consider the problem $\mathcal{P}(B_i, \mathbf{K})$. Since $(B_i, u_i)$ is $\mathcal{P}(\mathbf{K})$-feasible, the point $(I_n, u_i)$ is $\mathcal{P}(B_i, \mathbf{K})$-feasible. Let us compute the Newton iterate $(B_i', u_{i+1})$ of $(I_n, u_i)$ (the Newton method is applied to the function $F_{t_i}^{B_i}$). Then let us increase the value of the penalty parameter in the same manner as in the basic barrier method. Thereby we find a new approximate solution $(B_{i+1} \equiv B_iB_i', u_{i+1})$ to $\mathcal{P}(\mathbf{K})$ and a new value, $t_{i+1}$, of the penalty parameter. Now we can perform the next iteration, and so on. Note that the described procedure needs justification, because now we have no convex programming problem to which we can apply the barrier method. The main idea of the justification is the following. Similarly to the case of the basic barrier method, our aim is to prove that the procedure maintains the relation

$$\lambda(F_{t_i}^{B_i}, (I_n, u_i)) \leq 0.1$$

(of course, 0.1 could be replaced by some other constant).

Assume that this relation holds for some $i$. Our usual arguments, when applied to $\mathcal{P}(B_i, \mathbf{K})$, imply that

$$\lambda(F_{t_i}^{B_i}, (B_i', u_{i+1})) \leq \frac{0.1^2}{(1-0.1)^2} \leq 0.02.$$

The latter inequality in the case of

$$t_{i+1} = \left(1 + \frac{O(1)}{\vartheta^{1/2}}\right)t_i$$

implies that

$$\lambda(F_{t_{i+1}}^{B_i}, (B_i', u_{i+1})) \leq 0.05. \tag{6.5.3}$$

We wish to conclude from the latter relation that

$$\lambda(F_{t_{i+1}}^{B_{i+1}}, (I_n, u_{i+1})) \leq 0.1; \tag{6.5.4}$$

the obstacle is the fact that (6.5.3) and (6.5.4) involve *different* self-concordant functions. Let us use the following observation. It is not difficult to show that there is a *nonlinear* one-to-one correspondence between the feasible domains $G(\Lambda)$, $G(\Lambda')$ of problems $\mathcal{P}(\Lambda, \mathbf{K})$, $\mathcal{P}(\Lambda', \mathbf{K})$ such that the values of $\mathcal{V}$ (as well as the values of $\Phi^\Lambda$ and $\Phi^{\Lambda'}$) at two points corresponding to each other coincide (within an additive constant that depends on $\Lambda$, $\Lambda'$ only). Therefore $F_t^\Lambda$ and $F_t^{\Lambda'}$ at two points corresponding to each other differ by a constant which depends on $t$, $\Lambda$, $\Lambda'$ only. It turns out that the point $(B_i', u_{i+1})$ of the set $G(B_i)$ corresponds to the point $(I_n, u_{i+1})$ of the set $G(B_{i+1})$. By Theorem 2.2.2(iii), relation (6.5.3) means that the value of the function $F_{t_{i+1}}^{B_i}$ at the point $(B_i', u_{i+1})$ differs from the minimum of this function over $G(B_i)$ by no more than $(0.06)^2/2$. It follows that the value of $F_{t_{i+1}}^{B_{i+1}}$ at $(I_n, u_{i+1})$ differs from the minimum of the latter function over $G(B_{i+1})$ by no more than $(0.06)^2/2$. By Theorem 2.2.2(iv), the latter relation implies (6.5.4).

To use the same trick at the preliminary stage, we need some special effort. Indeed, at this stage, we deal with families of functions of the type

$$t\{\text{linear form}\} + F^\Lambda(z).$$

We wish these functions to be "almost invariant" under the above correspondence between $G(\Lambda)$ and $G(\Lambda')$. This condition is satisfied if the {linear form} depends on $u$-component of $z$ only, and we should provide the latter property. To this purpose, we use the *prepreliminary stage*, which we include in the method. At this stage, we are seeking such a point $z^\#$ that the partial derivative of the barrier with respect to $B$-component at $z^\#$ is close to zero. Assume that such a point is found. Then we can take the restriction of the first-order differential of the barrier onto $\mathbf{R}^n$ as the {linear form} and use $z^\#$ as the starting point for the preliminary stage. Now, to obtain an appropriate $z^\#$, we set $u = 0$ and minimize the barrier over the $B$-component only. This subproblem proves to be relatively simple, and we manage to solve it (this is the prepreliminary stage) with the aid of the barrier method at the cost of $O(m^{3.5}\ln(m\mathcal{R}))$ operations.

Now let us describe the *three-stage* version of the barrier method for $\mathcal{P}(\mathbf{K})$. Let us start with the description of the correspondence between $G(\Lambda)$ and $G(\Lambda')$.

**Lemma 6.5.2** *Let $\Lambda, \Lambda' \in L_n^+$. Consider the mapping $Z_{\Lambda,\Lambda'}$, which transforms a point $z = (B, u) \in S_n \times \mathbf{R}^n$ into the point $z' = (B', u) \in S_n \times \mathbf{R}^n$ such that*

$$\Lambda B^2 \Lambda^T = (\Lambda')(B')^2(\Lambda')^T$$

*(it is clear that the latter relation, for given $\Lambda, \Lambda'$, defines a positive-semidefinite symmetric $B'$ as a function of a positive semidefinite symmetric $B$).*

*The mapping $Z_{\Lambda,\Lambda'}$ is a one-to-one mapping from $G(\Lambda)$ onto $G(\Lambda')$ : $z \in G(\Lambda) \Rightarrow z' \in G(\Lambda')$, such that $\mathcal{V}(z) - \mathcal{V}(z')$ does not depend on $z$ and $\Phi^{\Lambda}(z) = \Phi^{\Lambda'}(z')$. The inverse of $Z_{\Lambda,\Lambda'}$ is $Z_{\Lambda',\Lambda}$.*

The proof of the lemma is straightforward, and we omit it.

**Description of the method.** To simplify our considerations, we choose the parameters of the method as concrete numeric constants.

*Prepreliminary stage.* At this stage, we deal with the problem

$$\mathcal{P}_1(\mathbf{K}) : \qquad \begin{array}{l} \text{minimize } R(C) = -2\ln \operatorname{Det} C - \sum\limits_{i=1}^{m} \ln\{b_i^2 - \langle C, A_i\rangle\} \\ \text{s.t. } C \in S_n^+,\ \langle C, A_i\rangle \le b_i^2,\ 1 \le i \le m, \end{array}$$

where $A_i = a_i a_i^T$ and $\langle Q, X\rangle = \operatorname{Tr}\{QX\}$ is the usual scalar product on $S_n$.

Let $G$ be the feasible domain of $\mathcal{P}_1(\mathbf{K})$. It is clear that $G$ is a bounded closed convex domain and $R$ is a $\vartheta'$-self-concordant barrier for $G$, $\vartheta' \equiv 2(n + 2m) \le 6m$. Let $C_0 = 0.5I_n$. By (III), $C_0$ is an interior point of $G$. At the prepreliminary stage, we apply the preliminary stage of the basic barrier method associated with $R$, $C_0$ being the starting point. The parameters $\lambda_1'$ ,$\lambda_1$, $\lambda_2$, $\lambda_3'$, $\lambda_3$ of the method are subject to the inequalities (compare with (3.2.5), (3.2.6))

$$0 < \lambda^+(\lambda_1) \le \lambda_1' < \lambda_2 < \lambda_3 < 0.01,$$

$$\lambda_1' < \lambda_1 < 0.01, \quad \lambda^+(\lambda_3) \le \lambda_3' < \lambda_3,$$

$$\zeta(\lambda_1') \le 0.01, \quad \frac{\omega^2(\lambda_3')}{(1-\omega(\lambda_3')^2} < 0.1,$$

$$\frac{\omega^2(\lambda_2)}{(1-\omega(\lambda_2))^2} \le 0.01;$$

recall that

$$\lambda^+(\lambda) = \frac{\lambda^2}{(1-\lambda)^2}, \quad \omega(\lambda) = 1 - (1-3\lambda)^{1/3}, \quad \zeta(\lambda) = \frac{\omega^2(\lambda)(1+\omega(\lambda))}{1-\omega(\lambda)}.$$

Let $C^*$ be the result of the prepreliminary stage. This point belongs to $\operatorname{int} G$ and satisfies the inequality

$$\lambda(R, C^*) \le \lambda \equiv 0.01. \tag{6.5.5}$$

Note that the number of iterations required to find $C^*$ does not exceed

$$N_1 = O(1)m^{1/2} \ln \frac{2m}{1 - \pi_R(C_0)} \tag{6.5.6}$$

(Proposition 3.2.3); herein $\pi_R$ is the Minkowsky function of $G$ with the pole at the minimizer of $R$ over $\operatorname{int} G$.

**Proposition 6.5.1** *We have*

$$N_1 \le O(m^{1/2}\ln(\mathcal{R}m)); \tag{6.5.7}$$

$$R(C^*) - \min_G R \le 0.6\lambda^2. \tag{6.5.8}$$

*The arithmetic cost of the prepreliminary stage does not exceed*

$$M_1 \le O\left(m^{3.5}\ln(\mathcal{R}m)\right). \tag{6.5.9}$$

**Proof.** In view of (6.5.6), to verify (6.5.7), we should prove that

$$\alpha \equiv 1 - \pi_R(C_0) \ge O\left(\frac{1}{(\mathcal{R}m)^s}\right)$$

for certain *absolute constant* $s$.

Recall that $\mathbf{K}$ contains the unit ball centered at 0 and is contained in the ball of the radius $\mathcal{R}$ centered at 0; moreover, $C$ is feasible for $\mathcal{P}_1(\mathbf{K})$ if and only if the ellipsoid $W(C^{1/2}, 0)$ is contained in $\mathbf{K}$. The above arguments show that the ball (in $S_n$) of radius $\frac{1}{4}$ centered at $C_0$ is contained in $G$ and that the diameter of $G$ does not exceed $4n\mathcal{R}^2$ (the latter relation holds since the semi axes of the ellipsoid $W(C^{1/2}, 0)$ for $C \in G$, i.e., the eigenvalues of the matrix $C^{1/2}$, do not exceed $\mathcal{R}$). Hence

$$\alpha \ge \frac{1}{16n\mathcal{R}^2} \ge \frac{O(1)}{(\mathcal{R}m)^2},$$

Relation (6.5.8) immediately follows from (6.5.5) and Theorem 2.2.2(iii).

In view of (6.5.8), to prove (6.5.9), it suffices to show that the Newton minimization step for a function

$$\{\text{a linear function of } C\} + R(C)$$

can be performed at the cost $O(m^3)$. It is easily seen that the gradient $2H$ of the function at a given point $C \in \operatorname{int} G$ can be computed at the above cost. A straightforward computation shows that the Hessian $2\mathcal{W}$ of the function at $C$ transforms $X \in S_n$ into the matrix

$$2\mathcal{W}X = 2C^{-1}XC^{-1} + \sum_{i=1}^{m} 2d_i \langle A_i, X\rangle A_i,$$

where the collection of numbers

$$2d_i = \frac{1}{(b_i^2 - \langle A_i, C\rangle)^2}, \qquad 1 \le i \le m$$

can be computed at the cost $O(mn^2)$. The Newton displacement $X$ is the solution to the system $\mathcal{W}X = H$; hence, it can be represented as

$$X = C\left\{H + \sum_{i=1}^{m} x_i A_i\right\}C, \tag{6.5.10}$$

where $x_i$, $1 \le i \le m$ are (unknown) scalars. Let us derive the system of linear equations for these scalars (the solution of the latter system, after substitution into (6.5.10), gives the desired $X$). To derive the system, let us substitute (6.5.10) into the equation $\mathcal{W}X = H$. After some simple transformations, we obtain the equation

$$\sum_{i=1}^{m} x_i A_i + \sum_{i=1}^{m} d_i \langle A_i, CHC\rangle A_i + \sum_{i=1}^{m} d_i \sum_{j=1}^{m} x_j \langle CA_jC, A_i\rangle A_i = 0. \tag{6.5.11}$$

This matrix equation is equivalent to a system (let it be (*)) of $m$ scalar linear equations with $m$ unknowns $x_i$. The system is obtained by taking termwise scalar products of (6.5.11) and the matrices $A_j$, $1 \le j \le m$. The $ij$th element of the matrix of (*) is

$$\langle A_j, A_i\rangle + \sum_{k=1}^{m} d_k \langle A_k, A_i\rangle \langle A_k, CA_jC\rangle,$$

and the $i$th component of the right-hand side vector is

$$-\sum_{k=1}^{m} d_k \langle A_k, A_i\rangle \langle A_k, CHC\rangle.$$

To form the matrix of (*) and the right-hand side vector of the system, it suffices to compute (i) all the products $\langle A_i, A_j\rangle$ ($O(m^2 n)$ operations); (ii) $m$ matrices $CA_jC$ ($O(n^2 m)$ operations) and all the scalar products of these matrices and matrices $A_i$ ($O(m^2 n)$ operations more); (iii) the matrix $CHC$ ($O(n^3)$ operations) and its scalar products by matrices $A_k$ ($O(n^2\, m)$ operations more).

When evaluating the numbers of operations, we considered the fact that $A_i$ are of rank 1.

After the above quantities are computed, each of the coefficients of system (*) can be computed at the cost $O(m)$. Thus, system (*) can be formed at the cost $O(m^3)$; then it can be solved at the cost $O(m^3)$. After the system is solved, the Newton displacement $X$ can be computed at the cost $O(mn^2)$; see (6.5.10). $\square$

*Initialization of the preliminary stage.* Having found the positive-definite symmetric matrix $C^*$, we compute its factorization $C^* = B_* B_*^T$, where $B_* \in L_n^+$ (the cost is $O(n^3)$). Consider the problem

$$\mathcal{P}_1(B_*, \mathbf{K}): \qquad \text{minimize } P(B) \equiv -2\ln\operatorname{Det} B - \sum_{i=1}^{m} \ln\{b_i^2 - \| BB_*^T a_i \|_2^2\}$$
$$\text{s.t. } B \in S_n^+,\ (B, 0) \in G(B_*).$$

It is not difficult to prove (compare with Lemma 6.5.1) that there exists a one-to-one correspondence between the feasible sets of the problems $\mathcal{P}_1(\mathbf{K})$ and $\mathcal{P}_1(B_*, \mathbf{K})$ with the following property: If $C$ is a feasible point for the first problem and $B$ is the corresponding feasible point for the second problem, then

$R(C) = 2P(B)$+const. Note that, under our correspondence, $C^*$ is transformed into $I_n$.

By the above arguments, (6.5.8) implies the relation $P(I_n) - \min_{G_*} P \leq 0.3\lambda^2$ ($\lambda = 0.01$), where $G_*$ is the feasible domain of $\mathcal{P}_1(B_*, \mathbf{K})$. Let

$$F^*(z) \equiv F^{B_*}(z)$$

be the barrier for the feasible set $G(B_*)$ of problem $\mathcal{P}(B_*, \mathbf{K})$ and let $z^* = (I_n, 0)$. As we have seen,

$$z^* \in \operatorname{int} G(B_*);\ F^*(z^*) - \min\{F^*(B,0) \mid (B,0) \in G(B_*)\} \leq 0.3\lambda^2. \tag{6.5.12}$$

Let $\langle\cdot,\cdot\rangle_*$ denote the scalar product in $E$ defined by the bilinear form $D^2F^*(z^*)[\cdot,\cdot]$ and let $\|\cdot\|_*$ be the corresponding norm. By Theorem 2.2.2(iv), relation (6.5.12) implies that

$$\mid DF^*(z^*)[(H,0)] \mid \leq 0.07 \parallel (H,0) \parallel_* \quad \forall H \in S_n. \tag{6.5.13}$$

Moreover, if $z^{**} = (X^{**}, 0)$ is the minimizer of $F^*$ on the set $G^* = \{(X,u) \in G(B_*) \mid u = 0\}$, then

$$D^2F^*(z^{**})[z^* - z^{**}, z^* - z^{**}] \leq 0.01 \tag{6.5.14}$$

(Theorem 2.2.2(iii)). By (6.5.13), there exists a linear form

$$\psi(w) \equiv \langle\psi^*, w\rangle_* = (\psi^{**}, w)$$

on $E$ (recall that $(\cdot,\cdot)$ is the standard scalar product on $E = S_n \times \mathbf{R}^n$) such that $\parallel \psi^* \parallel_* \leq 0.07$ and the restriction of the form onto $S_n$ coincides with the similar restriction of linear form $DF^*(z^*)[w]$.

Let us compute the form $\psi(w)$. Consider the linear form

$$\phi(w) \equiv -DF^*(z^*)[w] + \psi(w) = \langle\phi^*, w\rangle_* = (\phi^{**}, w).$$

For $\Lambda \in L_n^+$ and $t > 0$, let $F_t^\Lambda(z) = t\phi(z) + 2\mathcal{V}(z) + \Phi^\Lambda(z) \equiv t\phi(z) + F^\Lambda(z) : \operatorname{int} G(\Lambda) \to \mathbf{R}$, so that $F_t^\Lambda \in S_1^+(\operatorname{int} G(\Lambda), S_m \times \mathbf{R}^n)$.

**Proposition 6.5.2** *The form $\phi(w)$ depends only on the $u$-component of $w \in E$, and*

$$\lambda(F_1^{B_*}, z^*) \leq 0.07, \tag{6.5.15}$$

$$\parallel \phi^* \parallel_* \leq O(m^{1/2}). \tag{6.5.16}$$

*Let $z^+$ be the minimizer of $F^*(\cdot)$ over $\operatorname{int} G(B_*)$ and let $\pi_+(z)$ be the Minkowsky function of $G(B_*)$ with the pole at $z^+$. Then*

$$\pi_+(z^*) \leq 1 - \frac{O(1)}{\mathcal{R}m}. \tag{6.5.17}$$

*The vector $\phi^{**} \in E$ can be computed at the cost of $O(m^2n^2 + m^3)$ operations.*

**Proof**. The fact that $\phi$ depends only on the $u$-component of the argument is evident (recall that the restriction of $\psi$ onto $S_n$ coincides with the restriction onto this subspace of the form $DF^*(z^*)[\cdot]$ and $\phi$ is the difference of the above forms).

Furthermore,

$$DF_1^{B_*}(z^*)[w] = DF^*(z^*)[w] + \phi(w) = \psi(w),$$

and (6.5.15) follows from the relation $\| \psi^* \|_* \leq 0.07$ (see the definition of $\psi$). Moreover,

$$\| \phi^* \|_* \leq \| (F^*)'(z^*) \|_* + \| \psi^* \|_* \leq 0.07 + O(m^{1/2}),$$

since $F^*$ is a self-concordant barrier with the parameter $O(m)$; inequality (6.5.16) is proved.

To verify (6.5.17), note that the pair

$$z = (B, u) \in S_n \times \mathbf{R}^n$$

is feasible for $\mathcal{P}(B_*, \mathbf{K})$ if and only if the ellipsoid $W(B_*B, u)$ is contained in $\mathbf{K}$. Let us introduce the norm

$$p(B, u) = \| B_*B \| + \| u \|_2$$

($\| \cdot \|$ is the usual operator norm) on $E$ and let $B_0 = (1/2)(C^*)^{-1/2}$. Since

$$B_*B_*^T = C^*, B_*(C^*)^{-1/2}$$

is an orthogonal matrix, so that the ellipsoid $W(B_*B_0, 0)$ is the Euclidean ball of the radius $\frac{1}{2}$ centered at 0. By (III), the $\frac{1}{4}$-neighbourhood of the point $z_0 = (B_0, 0)$ (in the metric defined by $p$) is contained in $G(B_*)$. At the same time, (III) means that the diameter of $G(B_*)$ in this metric is $\leq O(\mathcal{R})$. Hence,

$$\pi_+(z_0) \leq 1 - \frac{O(1)}{\mathcal{R}}. \tag{6.5.18}$$

Furthermore, the restriction of $F^*$ onto $G^* \equiv \{(B, 0) \in G(B_*)\}$ is an $O(m)$-self-concordant barrier for $G^*$, and the minimizer of this barrier over $G^*$ is $z^{**} = (X^{**}, 0)$. Hence, the set $G^*$ contains the ellipsoid (in $S_n$)

$$U = \{(Y, 0) \mid D^2F^*(z^{**})[Y - X^{**}, Y - X^{**}] \leq 1\}$$

and is contained in the ellipsoid

$$U' = \{(Y, 0) \mid D^2F^*(z^{**})[Y - X^{**}, Y - X^{**}] \leq O(m^2)\}$$

(see Proposition 2.3.2(ii)).

The latter observation, combined with (6.5.14), implies that $z^*$ can be represented as

$$z^* = \alpha z_0 + (1-\alpha)z$$

for certain $z \in G^*$ and $\alpha \in (0,1)$, $\alpha \geq O(1/m)$. In view of convexity of $\pi_+$ and (6.5.18), we have

$$\pi_+(z^*) \leq \alpha\pi_+(z_0) + (1-\alpha)\pi_+(z) \leq 1 - \frac{\alpha O(1)}{\mathcal{R}} \leq 1 - \frac{O(1)}{m\mathcal{R}},$$

which is required in (6.5.17).

It remains to evaluate the cost at which $\phi^{**}$ can be computed. Let $Q$ be the Hessian of $F^*$ at the point $z^*$, let $q$ be the gradient of $F^*$ at this point, and let $\Omega$ be the orthoprojector from $E$ onto $S_n$. Let $x = (X, 0)$ be the solution to the system

$$\Omega(q - Qx) = 0; \qquad x \in S_n \tag{6.5.19}$$

($S_n$ is identified with the subspace $S_n \times \{0\}$ in $E$). It is not difficult to show that $\phi^{**} = Qx - q$. Indeed, for $w \in S_n$, we have

$$\langle x, w\rangle_* = (Qx, w) = (\Omega Qx, w) = (\Omega q, w) = (q, w).$$

If $w \in E$ is $\langle\cdot,\cdot\rangle_*$-orthogonal to $S_n$, then $\langle x, w\rangle_* = 0$ in view of $x \in S_n$. Hence $x$ is $\langle\cdot,\cdot\rangle_*$-orthogonal projection of the gradient of $F^*$ at the point $z^*$ onto $S_n$ (*the gradient is taken with respect to the Euclidean structure* $\langle\cdot,\cdot\rangle_*$), or, which is the same, $x = \psi^*$. Thus, $\psi^{**} = Qx$ and $\phi^{**} = Qx - q$.

Let us write the following expressions for the first- and the second-order differentials of $F^\Lambda$ at the point $(I_n, u)$ :

$$DF^\Lambda(I_n, u)[(H, v)] = -2\langle I_n, H\rangle - \sum_{i=1}^m d_i \left\{c_i a_i^T v - \left\langle H, \Lambda^T A_i \Lambda\right\rangle\right\}, \tag{6.5.20}$$

(6.5.21)

$$D^2F^\Lambda(I_n, u)[(H, v), (H, v)] = 2\langle H, H\rangle + \sum_{i=1}^m r_i \left\{c_i a_i^T v - \left\langle H, \Lambda^T A_i \Lambda\right\rangle\right\}^2 + \sum_{i=1}^m s_i \left\langle \Lambda^T A_i \Lambda H, H\right\rangle,$$

where $A_i = a_i a_i^T$. For given $u$, $\Lambda$, the collection of scalars $d_i$, $c_i$, $r_i$, $s_i$ (which depend on $u$ and $\Lambda$ only) can be computed at the cost of $O(mn^2)$ operations (when referring to the costs of computations, we consider the fact that the matrices $A_i$ are of rank 1). Note that $s_i \geq 0$.

In particular, we see that the computation of $q$, as well as the multiplication of $Q$ by a given vector, can be performed at the cost $O(mn^2)$.

To prove that $\phi^{**}$ can be computed at the cost $O(m^2n^2 + m^3)$, it suffices to show that, at this cost, we can find a symmetric solution $X$ to the matrix equation

$$X + JX + XJ + \sum_{i=1}^m \left\{\alpha_i + \beta_i \left\langle \Lambda^T A_i \Lambda, X\right\rangle\right\} \Lambda^T A_i \Lambda = V. \tag{6.5.22}$$

In this equation, $J = \frac{1}{2}\sum_{i=1}^m s_i \Lambda^T A_i \Lambda$ is a symmetric positive-semidefinite matrix that can be computed at the cost $O(mn^2)$. At the same cost, we can compute the symmetric matrix $V$ and the collection of scalars $\alpha_i$, $\beta_i$.

$1^0$. To solve (6.5.22), we act as follows. Let us reduce the matrix $J$ by an orthogonal transformation $U$ to a three-diagonal form; i.e., let us find (at the cost $O(n^3)$) an orthogonal matrix $U$ and a three-diagonal matrix $P$ such that $UJU^T = P$. The substitution $Y = UXU^T$ transforms (6.5.22) into the equation

$$Y + PY + YP + \sum_{i=1}^{m}\{\alpha_i + \beta_i \left\langle S^T A_i S, Y\right\rangle\} S^T A_i S = L, \tag{6.5.23}$$

where $S = \Lambda U^T$, $L = UVU^T$ are matrices that can be computed at the cost $O(n^3)$. We desire to find a symmetric solution to (6.5.23); this solution (at the cost $O(n^3)$) can be transformed into the required solution to (6.5.22). Thus, we must show that (6.5.23) can be solved at the cost $O(m^2n^2 + m^3)$.

$2^0$. Let us find the solutions to $(m+1)$ matrix equations

$$Y_i + PY_i + Y_iP = S^T A_i S, \qquad 1 \le i \le m,$$

$$Y_{m+1} + PY_{m+1} + Y_{m+1}P = L.$$

Since $P$ is a three-diagonal matrix, each of these equations can be solved at the cost $O(n^3)$. Indeed, the equation (with a $n \times n$ matrix $Z$ as unknown)

$$\mathsf{P}(Z) \equiv Z + PZ + ZP = M \tag{6.5.24}$$

has the unique solution, since the operator $\mathsf{P}$ (regarded as a linear operator in $L_n$) is symmetric and positive-definite (note that $P$ is symmetric positive-semidefinite). The subspace of symmetric matrices is invariant with respect to $\mathsf{P}$. Hence, for a symmetric $M$, the solution $Z$ to (6.5.24) is symmetric, also. In particular, $Y_i$ are symmetric, $1 \le i \le m+1$.

$3^0$. Let us show that (6.5.24) can be solved at the cost $O(n^3)$.

Matrix equation (6.5.24), regarded as a system with $n^2$ unknowns (the elements of $Z$), can be described as follows. The matrix $P$ is symmetric positive-semidefinite and three-diagonal, as follows:

$$Pe_i = \gamma_i e_i + \mu_i e_{i-1} + \mu_{i+1} e_{i+1},$$

where $e_i$, $1 \le i \le n$ are the standard orths in $\mathbf{R}^n$, $e_0 = e_{n+1} = 0$, $\mu_1 = \mu_{n+1} = 0$, $\gamma_i \ge 0$. Let $l_i$ be the columns of $M$ and let $z_i = Ze_i$, $0 \le i \le n+1$. Then (6.5.24) can be written as a system $(\mathcal{U})$ of equations

$$(\mathcal{U}_i): \qquad \mu_i z_i + (\gamma_{i-1}+1) z_{i-1} + P z_{i-1} + \mu_{i-1} z_{i-2} = l_{i-1}, \qquad 2 \le i \le n+1,$$

with unknown vectors $z_i$, which are subject to restrictions $z_0 = z_{n+1} = 0$. To solve the system, let us act as follows. The indices of nonzero elements of the sequence $\mu^* = \{\mu_1, \dots, \mu_{n+1}\}$ can be partitioned into mutually disjoint sequential groups $I_r = \{s_r, s_r + 1, \dots, t_r\}$, $1 \le r \le k$, such that

$$\mu_{s_r - 1} = \mu_{t_r + 1} = 0$$

(note that $\mu_0 = \mu_{n+1} = 0$). Let

$$I_r^- = I_r \bigcup \{s_r - 1\}, \qquad I_r^+ = I_r \bigcup \{t_r + 1\}.$$

Let $i_1, \dots, i_f$ be the elements of $\{1, 2, \dots, n\}$ not belonging to $\bigcup I_r^-$; $I_{r+j}^- = \{i_j\}$ and $I_{r+j}^+ = \{i_j + 1\}$ for $1 \le j \le f$. We have defined the groups $I_j^-$, $I_j^+$, $1 \le j \le k + f$. Let $(\mathcal{U}(r))$ denote the subsystem of system $(\mathcal{U})$ comprised by equations $(\mathcal{U}_i)$ with indices $i$ belonging to $I_r^+$. It is easy to prove that subsystem $(\mathcal{U}(r))$ involves only $z_i$ with $i \in I_r^-$ (so that these subsystems have no common unknowns), and system $(\mathcal{U})$ is a "direct product" of subsystems $(\mathcal{U}(r))$, $r = 1, \dots, k + f$. It suffices to prove that subsystem $(\mathcal{U}(r))$ can be solved at the cost $O(n^2 a(r))$, where $a(r)$ is the number of elements in $I_r^-$.

To avoid cumbersome notation, assume that $(\mathcal{U}(r))$ consists of the equations $(\mathcal{U}_i)$, $i = 2, \dots, p$; thus,

$$\mu_p = 0, \qquad \mu_2, \dots, \mu_{p-1} \neq 0.$$

We desire to solve the subsystem at the cost of $O(n^2 p)$ operations. Let us act as follows. Let $Z_0$ be the zero and $Z_1$ be the unit $n \times n$ matrices and let the matrix $Z_i$ be defined for $2 \le i < p$ as

$$Z_i = -\frac{1}{\mu_i} \left\{ \{(\gamma_{i-1} + 1) I_n + P\} Z_{i-1} + \mu_{i-1} Z_{i-2} \right\}.$$

It is clear that the general solution to the homogeneous system of equations

$$\mu_i z_i + (\gamma_{i-1} + 1) z_{i-1} + P z_{i-1} + \mu_{i-1} z_{i-2} = 0, \qquad 2 \le i < p,$$

(where $z_0 = 0$) is of the form $z_i = Z_i \lambda$, $1 \le i < p$, where $\lambda \in \mathbf{R}^n$. A particular solution $\{z_i^*,\ 1 \le i < p\}$ to the system

$$\mu_i z_i + (\gamma_{i-1} + 1) z_{i-1} + P z_{i-1} + \mu_{i-1} z_{i-2} = l_{i-1}, \qquad 2 \le i < p,$$

can be found recursively, below:

$$z_i^* = \frac{1}{\mu_i} \left\{ l_{i-1} - (\gamma_{i-1} + 1) z_{i-1}^* - P z_{i-1}^* - \mu_{i-1} z_{i-2}^* \right\}, \qquad 2 \le i < p,$$

where $z_0^* = z_1^* = 0$.

Since the matrix $P$ is three-diagonal, to compute all matrices $Z_i$ and vectors $z_i^*$, $1 \le i < p$, it takes totally no more than $O(n^2 p)$ operations (note that the matrices $Z_i$ are $O(p)$-diagonal). It is clear that the solution to the subsystem under consideration is

$$z_i = Z_i \lambda^* + z_i^*, \qquad 1 \le i < p, \tag{6.5.25}$$

where $\lambda^*$ is such that the equation $(\mathcal{U}_p)$ is satisfied by the corresponding $z_i$. In other words, $\lambda^*$ is the solution to the linear system

$$\{(\gamma_{p-1}+1)+P\}(Z_{p-1}\lambda+z^*_{p-1})+\mu_{p-1}(Z_{p-2}\lambda+z^*_{p-2})=l_{p-1}.$$

This system at the cost $O(n^2)$ can be reduced to the standard form, and the matrix of this system is $O(p)$-diagonal; thus, the cost at which the system can be solved by the conjugate gradient method does not exceed $O(n^2p)$. After $\lambda^*$ is computed, it takes no more than $O(np^2)$ operations to compute $z_i$ according to (6.5.25). Thus, the subsystem under consideration can be solved in $O(n^2p)$ operations, as it was announced at the beginning of $3^0$.

$4^0$. Let us come back to (6.5.23). By $3^0$, the total cost at which each of the $Y_i$, $1 \le i \le m+1$ can be computed does not exceed $O(mn^3)$. It is clear that the solution to (6.5.23) can be represented as

$$Y=\sum_{i=1}^{m+1} t_iY_i \tag{6.5.26}$$

with appropriate scalars $t_i$. Substitution of (6.5.26) into (6.5.23) leads, by definition of $Y_i$, to the equation for $t_i$ of the form

$$\sum_{i=1}^{m} t_iS^TA_iS+t_{m+1}L+\sum_{i=1}^{m}\left\{\alpha_i+\beta_i\sum_{j=1}^{m+1} t_j\left\langle S^TA_iS,Y_j\right\rangle\right\}S^TA_iS=L. \tag{6.5.27}$$

This matrix equality is equivalent to a system (let it be (*)) of $m+1$ scalar linear equations with unknowns $t_i$; these equations can be obtained by taking termwise scalar products of (6.5.27) and matrices $S^TA_iS$, $1 \le i \le m$ and $L$. Let us compute all quantities of the form

$$\left\langle S^TA_iS,Y_j\right\rangle,\quad \left\langle S^TA_iS,S^TA_jS\right\rangle,\quad \left\langle S^TA_iS,L\right\rangle,\quad \langle L,L\rangle.$$

Since Rank $A_i=1$, the total cost of this computation is $\le O(m^2n^2)$. After these quantities are computed, it takes no more than $O(m)$ operations to compute each of the coefficients of (*). Thus, (*) can be reduced to the standard form at the total cost $O(m^2n^2+m^3)$. Solving (*) ($O(m^3)$ operations) and then computing $Y$ in accordance with (6.5.26) ($O(mn^2)$ operations more), we find the desired solution to (6.5.23). The proof is complete. □

*Preliminary stage.* At this stage, we compute matrices $B_i \in L_n^+$, vectors $u_i \in \operatorname{int}\mathbf{K}$, and numbers $t_i > 0$, $i \ge 0$ as follows:

$$B_0=B_*,\quad u_0=0,\quad t_0=1;$$

$$z_{i+1}\equiv(B_{i+1},u_{i+1})=(B_iB^{(i+1)},u_{i+1}),$$

where $(B^{(i+1)},u_{i+1})\equiv z^{(i+1)}$ is the Newton iterate of $h^{(i)}\equiv(I_n,u_i)$ (the Newton method is applied to the function $F_{t_i}^{B_i}(\cdot)$), $t_{i+1}=t_i\mathrm{e}^{-\mu}$, where

$$\mu=\frac{0.05}{1+\vartheta^{1/2}},\qquad \vartheta=2(n+m). \tag{6.5.28}$$

The preliminary stage is terminated at the first iteration when the relation

$$\lambda(F^{B_i}, (I_n, u_i)) \leq 0.1 \tag{6.5.29}$$

holds (the number of this iteration is denoted by $i^*$).

The result of the stage is the point

$$z^{\#} = (B^{\#}, u^{\#}) \equiv (B_{i^*}, u_{i^*}).$$

**Proposition 6.5.3** *The preliminary stage is well defined, namely,*

$$B_i \in L_n^+, \quad (I_n, u_i) \in \operatorname{int} G(B_i), \quad 0 \leq i \leq i^*.$$

*For all $i$, $0 \leq i \leq i^*$, the relations*

$$(\mathrm{I}_i): \quad \lambda(F_{t_i}^{B_i}, (I_n, u_i)) \leq 0.1, \tag{6.5.30}$$

$$(\mathrm{J}_i): \quad \lambda(F_{t_i}^{B_i}, z^{(i+1)}) \leq 0.02 \tag{6.5.31}$$

*hold.*

*The number $i^*$ of iterations at the preliminary stage satisfies the inequality*

$$i^* \leq O(m^{1/2} \ln(m\mathcal{R})). \tag{6.5.32}$$

*Each iteration of the preliminary stage can be performed at the arithmetic cost of $O(m^2n^2 + m^3)$ operations (this amount includes the cost of the verification of the termination condition).*

**Proof.** $1^0$. Assume that for some $i$ the following relation holds:

$$\mathcal{M}(i): \begin{cases} \text{for each } j,\ 0 \leq j \leq i, \text{ we have } B_j \in L_n^+, \\ (I_n, u_j) \in \operatorname{int} G(B_j) \text{ and relations } (I_j) \text{ hold}, \\ \text{relations } (J_j) \text{ hold for } 0 \leq j < i. \end{cases}$$

Note that $\mathcal{M}(0)$ is clearly true (see (6.5.15)). Let us show that, if $\mathcal{M}(i)$ holds, then $\mathcal{M}(i+1)$ holds. Indeed, the function $F_{t_i} \equiv F_{t_i}^{B_i}$ is strongly self-concordant on $\operatorname{int} G(B_i)$, so that, by $(\mathrm{I}_i)$ and Theorem 2.2.2(ii), we have $z^{(i+1)} \in G(B_i)$ and

$$\lambda(F_{t_i}, z^{(i+1)}) < \frac{(0.1)^2}{(1-0.1)^2} < 0.013 \tag{6.5.33}$$

(thus $(\mathrm{J}_i)$ holds).

Proposition 3.2.2 as applied to the strongly self-concordant family

$$\mathcal{F} \equiv \{\operatorname{int} G(B_i), F_t(z) = t\phi(z) + F^{B_i}(z), S_n \times \mathbf{R}^n\}_{t>0},$$

combined with the fact that $\phi$ is 0-compatible with the barrier $F^{B_i}$ (the parameter value for this barrier is $\vartheta = 2(n+m)$), implies that

$$\rho_{0.07}(\mathcal{F}; t_i, t_{i+1}) \leq \left\{1 + \frac{\vartheta^{1/2}}{0.07}\right\} \mu \leq \frac{1+\vartheta^{1/2}}{0.07}\mu = \frac{0.05}{0.07},$$

whence, by Theorem 3.1.1 and by (6.5.33),

$$\lambda(F_{t_{i+1}}, z^{(i+1)}) \leq 0.07.$$

By Theorem 2.2.2(iii), the latter relation leads to

$$F_{t_{i+1}}(z^{(i+1)}) - \min_{\operatorname{int} G(B_i)} F_{t_{i+1}}(z) \leq 0.6(0.07)^2. \tag{6.5.34}$$

By Lemma 6.5.2, the left-hand side in (6.5.34) is equal to

$$F^{B_{i+1}}_{t_{i+1}}(I_n, u_{i+1}) - \min_{\operatorname{int} G(B_{i+1})} F^{B_{i+1}}_{t_{i+1}},$$

which, combined with (6.5.34) and Theorem 2.2.2(iv), proves ($\mathrm{I}_{i+1}$). Thereby, the implication $\mathcal{M}(i) \Rightarrow \mathcal{M}(i+1)$ is proved.

$2^0$. Let us prove (6.5.32). Similarly to §6.5.3, let

$$F^*(z) = F^{B_*}(z) : \operatorname{int} G(B_*) \to \mathbf{R}, \qquad F^*_t(z) = t\phi(z) + F^*(z).$$

With the aid of the transformation $G(B_*) \to G(B_i)$ described in Lemma 6.5.2, we set the points $w_i \in G(B_*)$ into correspondence with the points $(I_n, u_i)$. In view of (6.5.34) and Lemma 6.5.2, we have

$$F^*_{t_i}(w_i) - \min_{\operatorname{int} G(B_*)} F^*_{t_i} \leq 0.6(0.07)^2, \tag{6.5.35}$$

whence by Theorem 2.2.2(iv)

$$\lambda(F^*_{t_i}, w_i) \leq 0.08. \tag{6.5.36}$$

Let, as in §6.5.3, $z^+$ be the minimizer of $F^*$ over $\operatorname{int} G(B_*)$ and let

$$W_{1/2} = \left\{w \in E \equiv S_n \times \mathbf{R}^n \mid D^2F^*(z^+)[w - z^+, w - z^+] \leq \tfrac{1}{4}\right\};$$

then $W_{1/2} \subset \operatorname{int} G(B_*)$ (Theorem 2.1.1(ii)). Moreover,

$$D^2F^*(w)[h,h] \geq 0.25 D^2F^*(z^+)[h,h]$$

for $w \in W_{1/2}$ (Theorem 2.1.1), which implies that

$$F^*(w) - F^*(z^+) \geq \tfrac{1}{32} \quad \text{for } w \in \partial W_{1/2}. \tag{6.5.37}$$

Let us provide $E$ with the scalar product $D^2F^*(z^+)[\cdot,\cdot]$ and let $\nabla$ and $\|\cdot\|$ denote the corresponding gradient and norm. In view of (6.5.16), (6.5.17), and Proposition 2.3.2(i.2), we have

$$\|\nabla\phi\| \le O(\mathcal{R}m^{2.5}),$$

which, combined with (6.5.37), implies that

$$F_t^*(w) - F_t^*(z^+) \ge \tfrac{1}{32} - O(t\mathcal{R}m^{2.5}).$$

Hence, under an appropriate choice of *absolute constants* as factors in the below $O(\cdot)$ we have, by virtue of (6.5.35),

$$t_i \le \frac{O(1)}{m^{2.5}\mathcal{R}} \Rightarrow w_i \in W_{1/2}. \tag{6.5.38}$$

If the premise in (6.5.38) holds for a given $i$, then

$$\lambda(F^*, w_i) \le \lambda(F_{t_i}^*, w_i) + 2t_i \|\nabla\phi\|$$

(since the norm induced by the form $D^2F^*(w_i)[\cdot,\cdot]$ is bounded, up to a factor 2, by the norm defined by the form $D^2F^*(z^+)[\cdot,\cdot]$). In view of (6.5.36), we have

$$t_i \le \frac{O(1)}{m^{2.5}\mathcal{R}} \Rightarrow \lambda(F^*, w_i) \le 0.08 + O(t_i m^{2.5}\mathcal{R}). \tag{6.5.39}$$

In particular, the implication

$$t_i \le \frac{O(1)}{m^{2.5}\mathcal{R}} \Rightarrow \lambda(F^*, w_i) \le 0.09$$

holds, whence, by Theorem 2.2.2(iii),

$$t_i \le \frac{O(1)}{m^{2.5}\mathcal{R}} \Rightarrow F^*(w_i) - \min_{\operatorname{int} G(B_*)} F^* \le 0.6(0.09)^2,$$

and, by Lemma 6.5.2,

$$t_i \le \frac{O(1)}{m^{2.5}\mathcal{R}} \Rightarrow F^{B_i}(I_n, u_i) - \min_{\operatorname{int} G(B_i)} F^{B_i} \le 0.6(0.09)^2,$$

so that (see Theorem 2.2.2(iv))

$$t_i \le \frac{O(1)}{m^{2.5}\mathcal{R}} \Rightarrow \lambda(F^{B_i}, (I_n, u_i)) \le 0.1. \tag{6.5.40}$$

This implication combined with the termination rule of the preliminary stage (see (6.5.29)) and the updating formula for $t_i$ leads to (6.5.32).

$3^0$. It remains to verify that an iteration of the preliminary stage can be performed in no more than $O(m^2n^2 + m^3)$ operations. It is clear that the cost of an iteration is

$$O(n^3) + \mathcal{L}_1 + \mathcal{L}_2,$$

where $\mathcal{L}_1$ is the cost of computing $\lambda(F^\Lambda, (I_n, u))$ (this computation is required by the termination rule) and $\mathcal{L}_2$ is the cost of computing the Newton displacement for the function $F_t^\Lambda$ at the point $(I_n, u)$. It is clear (see (6.5.21)) that the gradients of $F^\Lambda$ and $F_t^\Lambda$ at the point $(I_n, u)$ can be computed at the cost $O(mn^2)$. After these gradients are computed, to compute the above quantities, it suffices to to solve the equation (with unknowns $(X, v) \in E$)

$$(6.5.41) \qquad D^2F^\Lambda(I_n, u)[(X, v), (H, h)] = (s, (H, h)) \quad \forall (H, h) \in E,$$

$s \in E$ being given. Thus, the cost of an iteration is

$$O(mn^2 + \mathcal{L}),$$

where $\mathcal{L}$ is the cost at which (6.5.41) can be solved. It suffices to prove that $\mathcal{L} \leq O(m^2n^2 + m^3)$.

In view of (6.5.21), (6.5.41) can be rewritten as a system consisting of the matrix and the vector equations

$$(6.5.42) \qquad \begin{aligned} & X + \sum_{i=1}^m r_i\{c_i a_i^T v - \left\langle X, \Lambda^T A_i \Lambda \right\rangle\}\Lambda^T A_i \Lambda \\ & \quad + \sum_{i=1}^m s_i\{\Lambda^T A_i \Lambda X + X\Lambda^T A_i \Lambda\} \\ & = H, \end{aligned}$$

$$(6.5.43) \qquad \sum_{i=1}^m d_i\{c_i a_i^T v - \left\langle X, \Lambda^T A_i \Lambda \right\rangle\} a_i = h,$$

where $A_i = a_i a_i^T$. The collection of scalars $s_i > 0$, $c_i$, $r_i$, $d_i$, the vector $h \in \mathbf{R}^n$, and the symmetric matrix $H$ (these quantities depend on $u$ and $\Lambda$ only) for given $u$, $\Lambda$ can be computed at the cost $O(mn^2)$.

Let us act as in the proof of Proposition 6.5.2: First, compute the symmetric matrix

$$W = \sum_{i=1}^m s_i \Lambda^T A_i \Lambda$$

($O(mn^2)$ operations), then transform $W$ by an orthogonal transformation to a three-diagonal form ($O(n^3)$ operations) and rewrite (6.5.42), (6.5.43) as

$$(6.5.44) \quad Y + RY + YR + \sum_{i=1}^m r_i \left\{c_i a_i^T v - \left\langle Y, S^T A_i S \right\rangle\right\} S^T A_i S = G,$$

$$(6.5.45) \qquad \sum_{i=1}^m d_i \left\{c_i a_i^T v - \left\langle Y, S^T A_i S \right\rangle\right\} a_i = h$$

where $R$, $S$, and $G$ are certain known matrices (they can be computed at the cost $O(n^3)$; $R$ is symmetric positive-semidefinite and three-diagonal). Note that the solution $(Y, v)$ to system (6.5.44), (6.5.45) can be transformed at the cost $O(n^3)$ into the solution $(X, v)$ to (6.5.42), (6.5.43).

To solve (6.5.44), (6.5.45), we compute the solutions $Y_i$, $1 \le i \le m+1$ to the matrix equations

$$Y_i + RY_i + Y_iR = S^T A_i S, \qquad 1 \le i \le m,$$

$$Y_{m+1} + RY_{m+1} + Y_{m+1}R = G,$$

(see the proof of Proposition 6.5.2; the total cost of these computations is $O(m^2n^2 + m^3)$). Then we represent the $Y$-component of the desired solution as

$$Y = \sum_{i=1}^{m+1} \tau_i Y_i,$$

the scalars $\tau_i$ being our new unknowns. Substituting this representation into (6.5.44) and taking termwise scalar products of the resulting equation and each of the matrices $S^T A_i S$, $1 \le i \le m$, $G$, we obtain a linear system of scalar equations of the form

$$A\tau + Bv = p, \qquad (*)$$

$$C\tau + Dv = q, \qquad (**)$$

where $\tau = (\tau_1, \ldots, \tau_{m+1})^T$ and $v$ are the unknowns and the matrices $A$, $B$, $C$, $D$ are of sizes $(m+1) \times (m+1)$, $(m+1) \times n$, $n \times (m+1)$, $n \times n$, respectively ($(*)$ corresponds to (6.5.44); $(**)$ corresponds to (6.5.45)). By the same arguments as in the proof of Proposition 6.5.2, the quantities $A$, $B$, $C$, $D$, $p$, $q$ can be computed at the cost $O(m^2n^2)$. Thus, it costs no more than $O(m^2n^2 + m^3)$ to form $(*)$–$(**)$, to solve this system, and to transform its solution into the solution to (6.5.44), (6.5.45). $\square$

*Main stage.* At this stage, we produce matrices $C_i \in L_n^+$, vectors $v_i \in \operatorname{int} \mathbf{K}$ and numbers $t_i > 0$, $i \ge 0$, as follows:

$$(C_0, v_0) = (B^{\#}, u^{\#}), \qquad t_0 = 1;$$

$$w_{i+1} \equiv (C_{i+1}, v_{i+1}) = (C_i C^{(i+1)}, v_{i+1}),$$

where $(C^{(i+1)}, v_{i+1}) \equiv w^{(i+1)}$ is the Newton iterate of the point $h^{(i)} \equiv (I_n, v_i)$; the Newton method is applied to

$$\Xi_{t_i}^{C_i}(w) \equiv t_i \mathcal{V}(w) + \mathcal{V}(w) + \Phi^{C_i}(w),$$

where, as above,

$$\Phi^{\Lambda}(w) \equiv -\sum_{i=1}^{m} \ln\{(b_i - a_i^T u)^2 - \| B(\Lambda^T a_i) \|_2^2\} : \operatorname{int} G(\Lambda) \to \mathbf{R},$$

$$t_{i+1} = t_i e^{\mu},$$

$$\mu = \frac{0.05}{1 + 2\vartheta^{1/2}}, \qquad \vartheta = n + 2m.$$

**Proposition 6.5.4** *The main stage is well defined: For all $i \geq 0$, we have*

$$C_i \in L_n^+, \qquad (I_n, v_i) \in \operatorname{int} G(C_i).$$

*For all $i \geq 0$, the relations*

$$\text{(6.5.46)} \qquad (\mathrm{K}_i): \qquad \lambda(\Xi_{t_i}^{C_i}, (I_n, v_i)) \leq 0.1,$$

$$\text{(6.5.47)} \qquad (\mathrm{L}_i): \qquad \lambda(\Xi_{t_i}^{C_i}, w^{(i+1)}) \leq 0.02$$

*hold.*

*Each iteration of the main stage can be performed at the arithmetic cost of $O(m^2n^2 + m^3)$ operations.*

*For each $i$, the ellipsoid $W(C_i, v_i)$ is contained in $\mathbf{K}$ and*

$$\text{(6.5.48)} \qquad \ln \mid W(C_i, v_i) \mid \geq \ln \mid W(C^{\#\#}, v^{\#\#}) \mid -O\left(\frac{m}{t_i}\right),$$

*where $\mid W \mid = \operatorname{mes}_n W$ and $(C^{\#\#}, v^{\#\#})$ is a solution to $\mathcal{P}(\mathbf{K})$.*

**Proof**. The families

$$\{\operatorname{int} G(C), \Xi_t^C, E\}_{t>0}$$

associated with a $C \in L_n^+$ are strongly self-concordant families generated by $(n + 2m)$-self-concordant barriers $\mathcal{V}(w) + \Phi^C(w)$ for the sets $G(C)$ and by the 1-compatible with these barriers function $\mathcal{V}(w)$. By the termination rule for the preliminary stage (see (6.5.29)), we have

$$\lambda(F^{B^\#}, (I_n, u^\#)) \leq 0.01.$$

Clearly,

$$F^{B^\#} \equiv \Xi_1^{C_0}.$$

Thus, in view of $t_0 = 1$, $(\mathrm{K}_0)$ holds. In view of the same arguments as in the proof of Proposition 6.5.3, $(\mathrm{K}_0)$ implies that the iterations of the main stage are well defined and that relations $(\mathrm{K}_i)$, $(\mathrm{L}_i)$ are valid for each $i$. The cost of an iteration can be evaluated in the same manner as in the proof of Proposition 6.5.3. It remains to verify (6.5.48). This inequality, by Lemma 6.5.2, is equivalent to

$$\mathcal{V}(I_n, v_i) - \min \mathcal{V} \leq O\left(\frac{\vartheta}{t_i}\right);$$

the latter inequality follows from $(\mathrm{K}_i)$ by virtue of arguments similar to these used in the proof of Proposition 3.2.4. $\square$

*The main result.* The above propositions can be summarized in the following theorem.

**Theorem 6.5.1** *Assume that conditions* (I)–(III) *are satisfied. Then the described method finds an* $\varepsilon$*-optimal ellipsoid (for all* $\varepsilon \in (0,1)$*) in no more than* $O(m^{1/2}\ln(m\mathcal{R}/\varepsilon))$ *iterations of all three stages. The total arithmetic cost of these iterations does not exceed*

$$O\left(m^{2.5}(n^2+m)\ln\frac{m\mathcal{R}}{\varepsilon}\right).$$

The theorem is a straightforward consequence of Propositions 6.5.1–6.5.4. Recall that, in the case of the IEM, we can take $\mathcal{R} \leq 10n$, $m \leq O(n\ln n)$.

### 6.5.4 The minimum volume ellipsoid that contains a given set

A close to $\mathcal{P}(\mathbf{K})$ problem is to find the minimum volume ellipsoid containing a given finite set. The latter problem can be solved by the above techniques; here, we describe the corresponding results. Let $\Gamma$ be a given $m$-element set in $\mathbf{R}^n$ and $\mathbf{K}$ be the convex hull of $\Gamma$; we wish to find the ellipsoid of minimum volume containing $\Gamma$. This problem is referred to as $\mathcal{T}_n(\Gamma)$.

We use the following traditional trick. Let us regard $\mathbf{R}^n$ as an affine hyperplane $A$ in $\mathbf{R}^{n+1}$ defined by the equation $x_{n+1}=1$; thus, $\Gamma \subset A \subset \mathbf{R}^{n+1}$. Consider the problem

$\mathcal{T}^0_{n+1}(\Gamma)$ : find $(n+1)$-dimensional ellipsoid of minimal volume centered at 0 and containing $\Gamma$.

If $W \supset \Gamma$ is feasible for $\mathcal{T}^0_{n+1}(\Gamma)$, then $W$ defines a $n$-dimensional ellipsoid $W\bigcap A$ feasible for $\mathcal{T}_n(\Gamma)$. It is not difficult to show that this correspondence transforms the solution to $\mathcal{T}_{n+1}(\Gamma)$ into the solution to $\mathcal{T}_n(\Gamma)$. Moreover, if $W$ is an $\varepsilon$-optimal solution to $\mathcal{T}^0_{n+1}(\Gamma)$ (i.e., it is feasible for this problem and $\operatorname{mes}_{n+1} W \leq \exp\{\varepsilon\}V^{**}$, $V^{**}$ being the optimal value of the objective in $\mathcal{T}^0_{n+1}(\Gamma)$), then $W' \equiv W\bigcap A$ is an $\varepsilon$-optimal solution to $\mathcal{T}_n(\Gamma)$. It costs no more than $O(n^3)$ operations to transform the standard description of $W$ into the standard description of $W'$. Thus, we can restrict ourselves to problem $\mathcal{T}^0_{n+1}(\Gamma)$.

The algebraic reformulation of $\mathcal{T}^0_{n+1}(\Gamma)$ is as follows:

$\mathcal{T}^*$ : given a subset $\Gamma = \{x_i \mid 1 \leq i \leq m\} \subset \mathbf{R}^{n+1}$, minimize $\mathcal{V}(B) = -\ln\operatorname{Det} B$ by choice of $B \in L^+_{n+1}$ subject to $\| Bx_i \|_2 \leq 1,\ 1 \leq i \leq m.$

Let $Q$ denote the feasible set of the latter problem. Each $B \in Q$ defines an ellipsoid $W(B^{-1},0)$, which is feasible for $\mathcal{T}^0_{n+1}(\Gamma)$. To find an $\varepsilon$-solution to $\mathcal{T}^0_{n+1}(\Gamma)$, we must find an $\varepsilon$-solution to $\mathcal{T}^*$, i.e., $B \in Q$ such that $\mathcal{V}(B) - \inf_Q \mathcal{V} \leq \varepsilon$.

The optimal value of the objective of $\mathcal{T}^*$ clearly is the same as for the problem $\mathcal{T}^{**}$, which is obtained from $\mathcal{T}^*$ by replacing the restriction $B \in L^+_{n+1}$

by the restriction $B \in S_n$. The substitution $B^2 = C$ transforms $\mathcal{T}^{**}$ into the problem

$$\mathcal{T}^{***}: \qquad \mathcal{V}(C) \equiv -\ln \mathrm{Det}\, C \to \min \mid C \in S_n,\ \langle C, X_i \rangle \le 1,\ 1 \le i \le m,$$
$$\text{where } X_i = x_i x_i^T.$$

If $C$ is an $\varepsilon$-solution to $\mathcal{T}^{***}$ and $C = B^T B$ ($B \in L^+_{n+1}$), then $B$ is an $(\varepsilon/2)$-solution to $\mathcal{T}^*$. Given $C$, we can compute $B$ in $O(n^3)$ operations.

Assume that the following condition holds:

(IV) The convex hull $\mathbf{K}$ of the set $\Gamma$ contains the unit ball centered at 0 and is contained in the concentric ball of a given radius $\mathcal{R}$ (both of the balls are balls in $\mathbf{R}^n$).

It is not difficult to show, that under this assumption, we can add to $\mathcal{T}^{***}$ $n+1$ extra constraints of the type $C_{jj} \le c\mathcal{R}^2 m^4$, $1 \le j \le n+1$, without changing the optimal value of the objective; herein $c \ge \frac{1}{8}$ is an appropriate absolute constant. We come to the problem

$$\mathcal{T}^{\#}: \mathcal{V}(C) \to \min \mid C \in S_{n+1}, \qquad \langle C, X_i \rangle \le a_i,\ 1 \le i \le m+n+1$$

(we have increased the list of matrices $X_i$ to include our new constraints). It remains to find an $\varepsilon$-solution to $\mathcal{T}^{\#}$.

The feasible set $G^{\#}$ of the latter problem admits an $O(m)$-self-concordant barrier

$$F(C) = \mathcal{V}(C) - \sum_{j=1}^{m+n+1} \ln(a_i - \langle C, X_i \rangle),$$

and the objective is 1-compatible with this barrier. The point $C_0 = 0.25\mathcal{R}^{-2} I_{n+1}$ belongs to $\mathrm{int}\, G^{\#}$, and it is easy to show that

$$\ln \frac{1}{1 - \pi_+(C_0)} \le O(\ln(\mathcal{R}m))$$

(see (IV)), where $\pi_+$ is the Minkowsky function of $G^{\#}$ with the pole at the minimizer of $F$ over $\mathrm{int}\, G^{\#}$. Thus, problem $\mathcal{T}^{\#}$ can be solved by the basic barrier method (with $C_0$ taken as the starting point). The total number of iterations required to find an $\varepsilon$-solution to this problem (and hence to the original one) does not exceed $N(\varepsilon) = O(m^{1/2} \ln(\mathcal{R}m/\varepsilon))$. The arithmetic cost of an iteration does not exceed $O(m^3)$ (compare with Proposition 6.5.1). Thus, we can find an $\varepsilon$-solution to $\mathcal{T}_n(\Gamma)$ at the total cost of

$$O\left(m^{3.5} \ln \frac{\mathcal{R}m}{\varepsilon}\right)$$

operations.

**Remark 6.5.1** *The advantage of problem $\mathcal{T}_n$ as compared to $\mathcal{P}$ is that the first of these problems can be reduced to a problem with linear constraints* *(see $\mathcal{T}^{\#}$), which is not the case for $\mathcal{P}$. Recently, Khachiyan and Todd* [KhT 90] *proved that, to find an $\varepsilon$-solution to $\mathcal{P}(\mathbf{K})$, we can form a "small" sequence of auxiliary problems of the same analytical structure as $\mathcal{T}_n$. The sequence is comprised of $O((\ln 1/\varepsilon)(\ln\ln\mathcal{R}))$ problems that should be solved to an accuracy of order of $\varepsilon$, and the basic barrier method solves each of them at the cost of $O(m^{3.5}\ln(m\mathcal{R}/\varepsilon))$ operations. Thus, the total arithmetic cost of finding an $\varepsilon$-solution to $\mathcal{P}(\mathbf{K})$ proves to be $O(m^{3.5}(\ln(m\mathcal{R}/\varepsilon))\ln(1/\varepsilon)\ln\ln\mathcal{R})$, which, for a fixed $\varepsilon$, is approximately $O(m)$ times better than our estimate for $\mathcal{P}$.*

*This page intentionally left blank*

# Chapter 7

# Variational inequalities with monotone operators

In this chapter, we develop interior-point methods for variational inequalities involving monotone operators. This is, in a sense, the most general formulation of an extremum problem of convex structure: Variational inequalities with monotone operators cover not only the usual convex minimization problems, but also saddle-point problems for convex-concave games, Nash equilibriums, and so forth. The order of exposition is as follows. Section 7.1 contains an introduction to the problem and motivate the approach we use. In §7.2 we study self-concordant monotone operators and the related results on the Newton method (the theory is quite similar to that one developed in Chapter 1). In §7.3 we present the path-following method for variational inequalities with monotone operators compatible with a self-concordant barrier for the domain of the inequality, and the concluding §7.4 is devoted to the particular case of inequalities with linear operators.

## 7.1 Preliminary remarks

### 7.1.1 Variational inequalities

Recall the formulation of a variational inequality with a monotone operator. Let $E$ be a finite-dimensional real vector space and let $E^*$ be its conjugate. Let $S$ be a multivalued mapping defined on certain set $\mathrm{Dom}\{S\} \subseteq E$ and taking values in $E^*$; more precisely, $S$ sets into correspondence to a point $x \in \mathrm{Dom}\{S\}$ a nonempty subset $S(x) \subseteq E^*$. The set

$$\mathsf{G}(S) = \{(x, \xi) \in E \times E^* \mid x \in \mathrm{Dom}\{S\},\ \xi \in S(x)\}$$

is called the *graph* of $S$. If $S$ is single-valued (i.e., $S(x)$ consists of a single point for every $x \in \mathrm{Dom}\{S\}$), then the (unique) point of the set $S(x)$, $x \in \mathrm{Dom}\{S\}$ is also denoted by $S(x)$.

The mapping $S$ is called *monotone* if

$$\text{(7.1.1)} \qquad \langle \xi - \eta, x - y \rangle \geq 0 \quad \forall\, (x, \xi), (y, \eta) \in \mathsf{G}(S).$$

A monotone mapping $S$ is called *maximal monotone* if its graph cannot be extended without violation of the monotonicity property: For every $(z, \omega) \notin \mathsf{G}(S)$, there exists $(x, \xi) \in \mathsf{G}(S)$ such that $\langle \omega - \xi, z - x \rangle < 0$.

A pair $(G, S)$ is called a *monotone element* on $E$ if $G$ is a closed convex domain in $E$ and $S$ is a monotone operator with $\operatorname{int} G \subseteq \mathrm{Dom}\{S\} \subseteq G$.

Let $(G, S)$ be a monotone element. The corresponding *variational inequality* $\mathcal{V}(G, S)$ is the following problem:

find $x \in \mathrm{Dom}\{S\}$ such that $\langle \xi, y - x\rangle \geq 0$ or certain $\xi \in S(x)$

and all $y \in G$.

From monotonicity, it follows that every solution $x$ to $\mathcal{V}(G, S)$ satisfies the relation

$$\langle \eta, y - x\rangle \geq 0 \quad \text{for all } (y, \eta) \in \mathsf{G}(S). \tag{7.1.2}$$

A point $x \in G$ satisfying the latter relation is called a *weak solution* to the variational inequality $\mathcal{V}(G, S)$. Thus, every solution to $\mathcal{V}(G, S)$ is a weak solution to the inequality as well. Under mild assumptions, the inverse statement is also valid.

**Proposition 7.1.1** *Let $(G, S)$ be a monotone element on $E$ and let $x$ be a weak solution to $\mathcal{V}(G, S)$.*

(i) *If $S$ is a single-valued continuous mapping defined on $G$, then $x$ is a solution to $\mathcal{V}(G, S)$.*

(ii) *Let $S'$ be a maximal monotone operator and let $S$ be the restriction of $S'$ onto $G$:* $\mathrm{Dom}\{S\} = \mathrm{Dom}\{S'\} \bigcap G$, $S(y) = S'(y)$, $y \in \mathrm{Dom}\{S\}$. *Then $x$ is a solution to $\mathcal{V}(G, S)$.*

**Proof.** First, consider the case when $S$ is single-valued and continuous operator defined on $G$. Since $x^*$ is a weak solution, for $y \in G$ and $0 \leq t \leq 1$, we have

$$\langle S(x^* + t(y - x^*)), y - x^*\rangle \geq 0$$

and, in view of the continuity of $S$, we conclude that $\langle S(x^*), y - x^*\rangle \geq 0$, so that $x^*$ is a solution.

Now let $S$ be the restriction onto $G$ of a maximal monotone operator $S'$ Consider the following operator $g$ with $\mathrm{Dom}\{g\} = G$:

$$g(x) = \{\xi \in E^* \mid \langle \xi, y - x\rangle \leq 0,\ y \in G\}.$$

It is well known that this operator is maximal monotone. Now the interiors of the domains of the maximal monotone operators $S'$ and $g$ have a nonempty intersection, so that the sum of these operators (the operator $\bar{S}(x) = \{\xi + \eta \mid \xi \in S'(x),\ \eta \in g(x)\}$ with the domain $\mathrm{Dom}\{S'\} \bigcap \mathrm{Dom}\{g\}$) is also maximal monotone (Rockafeller's theorem; see, e.g., [GTr 89]). The domain of this operator coincides with $\mathrm{Dom}\{S\}$. Furthermore, $x^*$ is a weak solution to $\mathcal{V}(G, S)$, so that $\langle \eta, y - x^*\rangle \geq 0$, $(y, \eta) \in \mathsf{G}(S)$, whence $\langle \eta + \xi, y - x^*\rangle \geq 0$, $(y, \eta) \in \mathsf{G}(S), (y, \xi) \in \mathsf{G}(g)$, and we conclude that $\langle \chi, y - x^*\rangle \geq 0$ whenever $(y, \chi) \in \mathsf{G}(\bar{S})$. The latter relation means precisely that, adding the pair

$(x^*, 0)$ to the graph of $\bar{S}$, we do not violate the monotonicity; since $\bar{S}$ is maximal monotone, it follows that $(x^*, 0) \in \mathsf{G}(\bar{S})$. Thus, $x^* \in \mathrm{Dom}\{S'\} \bigcap G \equiv \mathrm{Dom}\{\bar{S}\}$ and $0 = \xi^* + \eta^*$ for certain $\xi^* \in S'(x^*) = S(x^*)$, $\eta^* \in g(x^*)$. From the definition of $g$, it immediately follows that $\langle \xi^*, y - x^* \rangle = - \langle \eta^*, y - x^* \rangle \geq 0$, $y \in G$, so that $x^*$ is a solution to $\mathcal{V}(G, S)$. $\square$

An advantage of the notion of a weak solution is that, under minimal assumptions, such a solution does exist.

**Proposition 7.1.2** *Let $(G, S)$ be a monotone element and let $G$ be bounded. Then $\mathcal{V}(G, S)$ admits weak solutions.*

**Proof.** Let us set into correspondence to a finite subset $f \subseteq \mathsf{G}(S)$ the set

$$X(f) = \{x \in G \mid \langle \eta, y - x \rangle \geq 0 \ \forall (y, \eta) \in f\}.$$

Let us prove that $X(f)$ is nonempty. Denote $f = \{(y_i, \eta_i),\ i = 1, \ldots, m\}$ and assume that $X(f)$ is empty. Since $f$ is finite and $G$ is a convex compact set, the emptiness of $X(f)$ implies the existence of scalars $\lambda_i \geq 0, \sum_{i=1}^m \lambda_i = 1$, such that the linear function $g(x) = \sum_{i=1}^m \lambda_i \langle \eta_i, y_i - x \rangle$ is negative on $G$,

$$g(x) \leq -\delta < 0, \qquad x \in G.$$

Let $x^* = \sum_{i=1}^m \lambda_i y_i$; then $x^* \in G$. Let $z \in \mathrm{int}\, G$ and $x(t) = x^* + t(z - x^*)$, $0 < t \leq 1$; then $x(t) \in \mathrm{int}\, G \subset \mathrm{Dom}\{S\}$, $t \in (0, 1]$. Let $\eta_t \in S(x(t))$. We have

$$\begin{aligned} -\delta \geq g(x(t)) &= \sum_{i=1}^m \lambda_i \langle \eta_i, y_i - x(t) \rangle \\ &\geq \sum_{i=1}^m \lambda_i \langle \eta_t, y_i - x(t) \rangle = \langle \eta_t, x^* - x(t) \rangle = -t \langle \eta_t, z - x^* \rangle . \end{aligned}$$

We have established that $t\omega(t) \geq \delta > 0$ for all $t \in (0, 1]$, where $\omega(t) = \langle \eta_t, z - x^* \rangle$. This is impossible, however, since $\omega$ is nondecreasing on $(0, 1]$ due to the monotonicity of $S$. This is the desired contradiction.

Thus, $X(f)$ is nonempty for each $f$. The set $X(f)$ is clearly compact, and the family of sets $X(\cdot)$ is nested, so that their intersection is nonempty; however, the latter intersection evidently is the set of all weak solutions to $\mathcal{V}(G, S)$. $\square$

### 7.1.2 Problems reducible to inequalities with monotone operators

It is well known that many interesting problems arising in convex programming, game theory, and so forth can be reduced to variational inequalities with monotone operators. Let us list the simplest examples.

**1.** *Minimization of a convex function.* Let $(G, f)$ be a functional element on $E$, i.e., a pair comprised of a closed convex domain $H \subset E$ and a convex lower semicontinuous function $f$ mapping $G$ into the extended real axis $\mathbf{R}\bigcup\{+\infty\}$ and such that $f$ is finite on $\operatorname{int} G$. It is well known that the subdifferential

$$\partial f(x) = \{\xi \in E^* \mid \langle \xi, y - x\rangle + f(x) \le f(y)\ \forall y \in G\}$$

(this mapping is defined at the set of all $x$ such that $f(x) < \infty$ and the right-hand side is nonempty) is maximal monotone on its domain, and the latter domain contains $\operatorname{int} G$; recall that the elements of $\partial f(x)$ are called support at $x$ on $G$ functionals to $f$. Thus, if $G' \subset G$ is a closed convex domain, then the pair comprised of $G'$ and of the restriction of $\partial f$ onto $G'$ is a monotone element; the solutions ($\equiv$ weak solutions) of the corresponding variational inequality are precisely the minimizers of $f$ over $G'$.

**2.** *Convex-concave games.* Let $G$ be a closed convex domain in $E$, let $Q$ be a closed convex domain in a finite-dimensional real vector space $H$, and let $f(x,y)$ be a (say, continuous) function defined on $G \times Q$, convex in $x \in G$ for every $y \in Q$ and concave in $y \in Q$ for each $x \in G$. The mapping $(x,y) \to \delta f(x,y) = \partial_x f(x,y) \times (-\partial_y f(x,y))$ putting into correspondence to a point $(x,y) \in G \times Q$ the set of all pairs $(\xi, \eta) \in E^* \times H^*$ such that $\xi$ is a support at $x$ on $G$ functional to $f(\cdot, y)$ and $\eta$ is a support at $y$ on $Q$ functional to $(-f(x,\cdot))$ (the domain of the mapping is comprised of those $(x,y)$ at which the corresponding sets of support functionals both are nonempty). It is well known that the pair $(G\times Q, \delta f)$ is a monotone element, and, if $G' \subseteq G$, $Q' \subseteq Q$ are closed convex domains, then the weak solutions to the variational inequality $\mathcal{V}(G' \times Q', \delta f\,|_{G'\times Q'})$ are precisely the saddle points of $f$ on $G' \times Q'$, i.e., the points $(x^*, y^*) \in G' \times Q'$ such that

$$f(x, y^*) \ge f(x^*, y^*) \ge f(x^*, y), \qquad (x,y) \in G' \times Q'.$$

The set of these saddle points is nonempty if and only if the following convex problems are solvable:

$$(f^*): \qquad \text{minimize } f^*(x) = \sup_{y\in Q'} f(x,y) \quad \text{s.t. } x \in G',$$

$$(f_*): \qquad \text{minimize } f_*(y) = \sup_{x\in G'} f(x,y) \quad \text{s.t. } y \in Q';$$

the set of weak solutions to $\mathcal{V}(G' \times Q', \delta f\,|_{G'\times Q'})$ is simply the direct product of the sets of solutions to these two problems.

**3.** *Nash equilibrium.* Let $G_i$ be closed convex domains in finite-dimensional real vector spaces $E_i$, $1 \le i \le m$ and let $f_i(x_1, \dots, x_m)$ be continuous functions defined on $G = G_1 \times \cdots \times G_m$; we assume that $f_i$ is convex in $x_i$ and concave in $x^i = (x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_m)$, $i = 1, \dots, m$ and that the sum of these functions is convex on $G$. Under this assumption, the operator

$$\delta(x_1, \dots, x_m) = \{(\xi_1, \dots, \xi_m) \mid \xi_i \text{ is a support at } x_i \text{ on } G_i$$
$$\text{functional to the function } f_i(x_1, \dots, x_{i-1}, \cdot, x_{i+1}, \dots, x_m),\ i = 1, \dots, m\}$$

defined on the set of those $x = (x_1, \ldots, x_m)$, where the right-hand side set is nonempty, is monotone, and the pair $(G, \delta)$ is a monotone element. Let $G'_i$ be closed convex domains contained in $G_i$, $i = 1, \ldots, m$. The weak solutions to the variational inequality $\mathcal{V}(G', \delta \mid_{G'})$, $G' = G'_1 \times \cdots \times G'_m$ are precisely the *Nash equilibriums* defined as follows. Consider $m$ players, the $i$th of them choosing $x_i \in G'_i$. The penalty paid by the $i$th player in the situation when the choices of the players comprise a collection of choices $x \in G'$ is $f_i(x)$. The set of choices $x^* \in G'$ is an equilibrium if each of the players cannot decrease his penalty by his separate actions, i.e., if $x_i^*$ is a minimizer of $f_i(x_1^*, \ldots, x_{i-1}^*, \cdot, x_{i+1}^*, \ldots, x_m^*)$ over $G'_i$ for all $i = 1, \ldots, m$. Note that the saddle-point problem (see §7.1.2.2) is a particular case of this situation when $f_1 = f$ and $f_2 = -f$.

**4.** *Complementarity problem.* This problem is closely related to optimality conditions in constrained optimization and is as follows:

$$\text{given a mapping } S : \mathbf{R}^n_+ \to \mathbf{R}^n, \text{ find } x \in \mathbf{R}^n_+ : S(x) \geq 0, \ \langle S(x), x \rangle = 0.$$

If $S$ is continuous, then the latter problem is precisely the problem of finding solutions ($\equiv$ weak solutions) to the variational inequality $\mathcal{V}(\mathbf{R}^n_+, S)$. If $S$ is monotone, then $(\mathbf{R}^n_+, S)$ is a monotone element. Also, when the complementarity problem expresses optimality conditions in a *convex* optimization problem, then $S$ proves to be monotone.

### 7.1.3 Overview of the contents

It is known that polynomial-time methods for general convex problems (like the ellipsoid method, inscribed ellipsoid method, and so forth) can be extended onto variational inequalities with monotone operators (see [Nm 81]). In what follows, we extend the interior-point methods developed for convex programming onto variational inequalities with monotone operators. The order of exposition is as follows.

(i) We introduce a class of *strongly self-concordant* monotone operators and study the convergence of the Newton method at the corresponding variational inequalities. The basic example of a self-concordant monotone mapping is the derivative of a 1-self-concordant function $f$. Such a derivative is a $\mathrm{C}^2$-smooth single-valued mapping $S$ from an open convex domain $Q = \mathrm{Dom}\{f\} \subseteq E$ into $E^*$ satisfying (for all $x \in Q$ and all $h_1$, $h_2$, $h_3 \in E$) the relation

$$(|D^3 f(x)[h_1, h_2, h_3]| =) \ |S''(x)[h_1, h_2, h_3]| \leq 2 \prod_{i=1}^{3} \langle S''(x) h_i, h_i \rangle^{1/2};$$

this is precisely the definition of a self-concordant monotone mapping $S : Q \to E^*$. Note that, in the definition of a self-concordant function, the above inequality was postulated only for the triples $(h_1, h_2, h_3)$ with $h_1 = h_2 = h_3$, and the fact that this inequality then holds for all triples could be proved (done in Appendix 1). The proof is based on symmetry of the three-linear form $S''(x)[\cdot,\cdot,\cdot]$ in the case of $S = f'$. Since now $S''$ should not be symmetric, we are forced to postulate the above inequality for all triples.

It turns out that the Newton method as applied to the equation $S(x) = 0$ (this equation is the "unconstrained" version of the variational inequality associated with $S$) involving a self-concordant monotone operator possesses the same convergence properties as in the case when $S$ is the derivative of a self-concordant function. The corresponding results form the contents of §7.2.

(ii) To develop the barrier-generated path-following method for variational inequalities, we introduce the notion of a (single-valued) monotone mapping *compatible* with a self-concordant barrier for a closed convex domain. This property generalizes the main feature of the first-order derivative of a function compatible with the barrier. It turns out to be possible to solve an inequality involving such a mapping by the path-following method associated with the barrier. This method is presented and studied in §7.3.

Until now, everything appeared similar to the optimization case. In the latter case, the next step was to reduce a given convex optimization problem to a problem with the objective that is definitely compatible with any barrier, i.e., to a problem with a *linear* (or a quadratic) objective. There were no difficulties in this reduction, so that we could claim that each convex programming problem can be solved with the aid of a barrier-generated path-following method, provided that we can find a self-concordant barrier for the feasible domain of the standard reformulation of the initial problem. We know that such a barrier always exists and, in many important cases, we are clever enough to find a "computable" self-concordant barrier; we even possess a kind of calculus for these barriers.

What happens in the case of monotone variational inequalities? In a sense, everything is the same. We have no difficulties with an inequality involving *linear* monotone operator and not too complicated $G$: To such an inequality, the path-following method can be applied directly. As we see, *in principle*, each monotone variational inequality can be reduced to an inequality involving a linear monotone operator, so that, *in principle*, each inequality can be solved with the aid of the interior point machinery. A drawback is that now the "feasible domain of the standard reformulation of the initial inequality" is more complicated than in the optimization case, so that we often cannot find a computable barrier for the latter domain. Nevertheless, we develop a kind of calculus that gives us some tools to find the desired barriers at least in simple situations (e.g., when the initial operator is the sum of a *linear* monotone operator and the derivative of a convex function, provided that we know a covering for the latter function). This calculus is developed in §7.4.

## 7.2 Self-concordant monotone operators and Newton method

Let $E$ be a finite-dimensional real vector space and let $E^*$ be the conjugate space. A linear operator $A : E \to E^*$ generates a bilinear form

$$\langle Ah, e\rangle : E \times E \to \mathbf{R},$$

where, as usual, $\langle \phi, x\rangle$ denotes the value of a functional $\phi \in E^*$ at a vector $x \in E$. The operator $A^*$ conjugate to $A$ acts from $(E^*)^*$ into $E^*$; since $(E^*)^*$ can be canonically identified with $E$, we can regard $A^*$ as an operator from $E$ into $E^*$. In particular, the operator $(A + A^*)/2$ is well defined; the bilinear form

$$\left\langle \tfrac{1}{2}(A + A^*)h, e\right\rangle$$

clearly is symmetric, and the associated quadratic form

$$\left\langle \tfrac{1}{2}(A + A^*)h, h\right\rangle$$

coincides with the quadratic form $\langle Ah, h\rangle$ generated by $A$.

Let $Q \subset E$ be an open nonempty convex set. Assume that $S$ is a $\mathrm{C}^2$-smooth single-valued monotone operator on $Q$. The first derivative $S'(x)$ for every $x$ is a linear mapping from $E$ into $E^*$, and therefore it defines a bilinear form $\langle S'(x)h, e\rangle : E \times E \to \mathbf{R}$. From monotonicity, it clearly follows that

$$\langle S'(x)h, h\rangle \geq 0, \qquad h \in E;$$

thus, the symmetric operator

$$\hat{S}(x) = \tfrac{1}{2}(S'(x) + (S')^*(x)) : E \to E^*$$

is positive semidefinite; i.e., it generates a symmetric bilinear form with nonnegative values on the diagonal of $E \times E$. In particular, $S$ at every $x \in Q$ defines the (possibly, degenerate) scalar product

$$(h, e)_{S,x} = \left\langle \hat{S}(x)h, e\right\rangle$$

on $E$ and a Euclidean seminorm

$$\| h \|_{S,x} = \{(h, h)_{S,x}\}^{1/2}.$$

The second derivative $S''(x)$ can be naturally identified with a three-linear form on $E$ defined as

$$S''(x)[h, e, s] = \frac{d}{dt} |_{t=0} \langle S'(x + th)e, s\rangle .$$

**Definition 7.2.1** *A single-valued monotone operator $S : Q \to E^*$ is called self-concordant (notation: $S \in \mathcal{S}(Q)$) if it is $\mathrm{C}^2$ smooth and if the following relation holds for all $x \in Q$ and all $h_i \in E$, $i = 1, 2, 3$ :*

$$|S''(x)[h_1, h_2, h_3]| \leq 2 \prod_{i=1}^{3} \| h_i \|_{S,x} . \tag{7.2.1}$$

*An operator $S \in \mathcal{S}(Q)$ is called strongly self-concordant on $Q$ (notation: $S \in \mathcal{S}^+(Q)$) if the sequence of operators $\hat{S}(x_i)$ is unbounded whenever $x_i \in Q$ form a sequence converging to a boundary point of $Q$.*

For $S \in \mathcal{S}(Q)$ and $x \in Q$, let

$$\mathcal{E}(S,x) = \{h \in E \mid \| h \|_{S,x} = 0\},$$

$$W_r(S,x) = \{y \in E \mid \| y - x \|_{S,x} \leq r\}, \qquad W_r^0(S,x) = \operatorname{int} W_r(S,x).$$

The following statement is a "finite-difference" reformulation of Definition 7.2.1 (cf. §2.1).

**Proposition 7.2.1** *Let $S$ be a self-concordant on $Q$ monotone operator. Then*
(i) *The space $\mathcal{E}(S,x)$ does not depend on $x \in Q$,*

$$\mathcal{E}(S,x) \equiv \mathcal{E}(S), \qquad x \in Q, \tag{7.2.2}$$

*and $S'$ is constant on every set of the form $(x + \mathcal{E}(S)) \bigcap Q$. If $S$ is strongly self-concordant on $Q$, then also*

$$Q = Q + \mathcal{E}(S). \tag{7.2.3}$$

(ii) *If $x \in Q$, then, for every $y \in Q \bigcap W_1(S,x)$ and all $h, h' \in E$, we have*

$$|\langle (S'(y) - S'(x))h, h' \rangle| \leq \left\{ \frac{1}{(1- \| y - x \|_{S,x})^2} - 1 \right\}, \| h \|_{S,x} \| h' \|_{S,x},$$

$$\| h \|_{S,x} (1- \| y - x \|_{S,x}) \leq \| h \|_{S,y} \leq \frac{\| h \|_{S,x}}{1- \| y - x \|_{S,x})}. \tag{7.2.4}$$

*If $S$ is strongly self-concordant on $Q$, then also*

$$W_1^0(S,x) \subset Q, \qquad x \in Q. \tag{7.2.5}$$

**Proof.** (i) Let $h \in \mathcal{E}(S,x)$, so that the function

$$\phi(y) = \left\langle \hat{S}(y)h, h \right\rangle \equiv \langle S'(y)h, h \rangle : Q \to \mathbf{R}$$

satisfies the relation $\phi(x) = 0$. We have

$$|\langle \phi'(y), e \rangle| = |S''(y)[e,h,h]| \leq 2 \| e \|_{S,y} \| h \|_{S,y}^2 = 2 \| e \|_{S,y} \phi(y).$$

Thus,

$$|\langle \phi'(y), e \rangle| \leq 2 \| e \|_{S,y} \phi(y), \qquad y \in Q;$$

the latter inequality, combined with the relation $\phi(x) = 0$, immediately leads to $\phi \equiv 0$. This identity means that $\mathcal{E}(S,x) \subset \mathcal{E}(S,y), y \in Q$. Since $x$ is an arbitrary point of $Q$, the latter relation implies (7.2.2).

To prove that $S'(y) = S'(x)$ for any pair $x, y \in Q$ satisfying $y - x \in \mathcal{E}(S)$, let us fix these $x$ and $y$; denote $e = y - x$ and let $h_1, h_2 \in E$. Set $f(t) = \langle S'(x + t(y-x))h_1, h_2 \rangle : [0,1] \to \mathbf{R}$. We have $|f'(t)| = |S''(x+te)[e, h_1, h_2]| \leq 2 \| e \|_{S,x+te} \| h_1 \|_{S,x+te} \| h_2 \|_{S,x+te} = 0$ (the latter equality holds, since $e \in$

$\mathcal{E}(S)$). Thus, $\langle S'(x+te)h_1, h_2\rangle$ does not depend on $t$, and therefore $S'(y) = S'(x)$.

To complete the proof of (i), it remains to verify that, in the case of $S \in \mathcal{S}^+(Q)$, we have $Q + \mathcal{E}(S) \subset Q$. Let $x \in Q$ and $y \in x + \mathcal{E}(S)$. If $y \notin Q$, then there exists a bounded sequence $t_i \in \mathbf{R}$ such that $x_i = x + t_i(y-x) \in Q$ and $x_i$ converge to a boundary point of $Q$. Since $S \in \mathcal{S}^+(Q)$, the sequence $\{\hat{S}(x_i)\}$ must be unbounded; but, according to the already-proved part of (i), $S'(x_i) = S'(x)$ and therefore $\hat{S}(x_i) = \hat{S}(x)$, which is impossible. Thus, $y \in Q$. Part (i) is proved.

(ii) Given $x, y \in Q, \| y - x \|_{S,x} < 1$, consider the function

$$\phi(t) = \langle S'(x+te)e, e\rangle : [0,1] \to \mathbf{R},$$

where $e = y - x$. This is a C$^1$ function, and we have

$$|\phi'(t)| = |S''(x+te)[e,e,e]| \leq 2 \| e \|^3_{S,x+te} \equiv 2\phi^{3/2}(t).$$

The inequality

$$|\phi'(t)| \leq 2\phi^{3/2}(t), \qquad 0 \leq t \leq 1$$

implies that either $\phi(t) \equiv 0,\ 0 \leq t \leq 1$ or

$$\left|\frac{d}{dt}\phi^{-1/2}(t)\right| \leq 1, \qquad 0 \leq t \leq 1;$$

since $\phi(0) = \| e \|^2_{S,x}$, in the latter case, we conclude that

$$\text{(7.2.6)} \qquad \frac{\| e \|^2_{S,x}}{(1 + t \| e \|_{S,x})^2} \leq \phi(t) \equiv \| e \|^2_{S,x+te} \leq \frac{\| e \|^2_{S,x}}{(1 - t \| e \|_{S,x})^2}.$$

Of course, (7.2.6) also holds true in the case of $\phi(t) \equiv 0$.

Now let $h_i \in E$, $i = 1, 2$ be such that $\| h_i \|_{S,x} \leq 1$. Let

$$f_i(t) = \langle S'(x+te)h_i, h_i\rangle = \| h_i \|^2_{S,x+te}, \qquad 0 \leq t \leq 1.$$

Then

$$|f_i'(t)| = |S''(x+te)[e, h_i, h_i]| \leq 2 \| e \|_{S,x+te} \| h_i \|^2_{S,x+te} = 2 \| e \|_{S,x+te}\ f_i(t).$$

It follows that either $f_i \equiv 0$, or $f_i > 0$ and

$$\left|\frac{d}{dt}\ln f_i(t)\right| \leq 2 \| e \|_{S,x+te} \leq \frac{2 \| e \|_{S,x}}{1 - t \| e \|_{S,x}}$$

(see (7.2.6)). In the latter case, we have

$$(1 - t \| e \|_{S,x})^2 \leq \frac{f_i(t)}{f_i(0)} \leq \frac{1}{(1 - t \| e \|_{S,x})^2}$$

or

$$(7.2.7) \qquad \| h_i \|_{S,x} (1 - t \| e \|_{S,x}) \leq \| h_i \|_{S,x+te} \leq \| h_i \|_{S,x} (1 - t \| e \|_{S,x})^{-1}.$$

Of course, (7.2.7) also holds in the case of $f_i \equiv 0$. Note that (7.2.7) contains the second relation in (7.2.4).

Now let

$$f(t) = \langle S'(x+te)h_1, h_2 \rangle, \qquad 0 \leq t \leq 1.$$

We have

$$\begin{aligned} |f'(t)| &= |S''(x+te)[e, h_1, h_2]| \leq 2 \| e \|_{S,x+te} \| h_1 \|_{S,x+te} \| h_2 \|_{S,x+te} \\ &\leq 2 \frac{\| e \|_{S,x} \| h_1 \|_{S,x} \| h_2 \|_{S,x}}{(1 - t \| e \|_{S,x})^3}, \end{aligned}$$

so that

$$\begin{aligned} | \langle (S'(y) - S'(x))h_1, h_2 \rangle | &= |f(1) - f(0)| \\ &\leq 2 \| e \|_{S,x} \| h_1 \|_{S,x} \| h_2 \|_{S,x} \int_0^1 \frac{dt}{(1 - t \| e \|_{S,x})^3} \\ &= \left( \frac{1}{(1 - \| y - x \|_{S,x})^2} - 1 \right) \| h_1 \|_{S,x} \| h_2 \|_{S,x}, \end{aligned}$$

as required in the first relation of (7.2.4).

From (7.2.4), it follows that the set of operators $\{\hat{S}(y) \mid y \in W_r^0(S,x) \bigcap Q\}$ is bounded for each $r < 1$. In the case of $S \in \mathcal{S}^+(Q)$, this observation immediately leads to (7.2.5). □

The following simple statement is rather useful.

**Proposition 7.2.2** (i) Stability of self-concordance with respect to affine substitutions of argument. *Let $S$ be self-concordant on $Q$ and let $x = \mathcal{A}(y) \equiv Ay+b$ be an affine mapping of a finite-dimensional space $E^+$ into $E$ with the image intersecting $Q$. Then the set $Q^+ = \mathcal{A}^{-1}(Q)$ is an open nonempty convex subset of $E^+$ and the operator $S^+(y) = A^* S(\mathcal{A}(y)) : Q^+ \to (E^+)^*$ is self-concordant on $Q^+$. If $S$ is strongly self-concordant on $Q$, then $S^+$ is strongly self-concordant on $Q^+$.*

(ii) Stability of self-concordance with respect to summation. *Let $S_i$ be self-concordant on $Q_i \subseteq E$, $i = 1, \ldots, k$ and let $Q = \bigcap_{i=1}^k Q_i \neq \emptyset$. If $\alpha_i \geq 1$, $i = 1, \ldots, k$, then the operator*

$$S(x) = \sum_{i=1}^k \alpha_i S_i(x) : Q \to E^*$$

*is self-concordant on $Q$. If $S_i$ are strongly self-concordant on $Q_i$, $i = 1, \ldots, k$, then also $S$ is strongly self-concordant on $Q$. If $Q_i \equiv Q$, $i = 1, \ldots, k$ and $S_1$ is strongly self-concordant on $Q$, then $S$ also is strongly self-concordant on $Q$.*

(iii) Self-concordance of the derivative of a self-concordant function. *Let $f$ be a 1-self-concordant function on $Q$. Then the operator $S = f'$ is self-concordant on $Q$. If $f$ is strongly 1-self-concordant on $Q$, then $S$ is strongly self-concordant on $Q$.*

**Proof.** (i) The only statement that is not a straightforward consequence of Definition 7.2.1 is the claim that the strict self-concordance of $S$ implies that one of $S^+$. If $S$ is strongly self-concordant on $Q$, then $W_1^0(S,x) \subset Q$ for every $x \in Q$; it immediately follows that $W_1^0(S^+,y) \in Q^+$ for every $y \in Q^+$. Thus, a sequence of points with bounded $\hat{S}^+(\cdot)$ cannot converge to a boundary point of $Q^+$, so that $S^+$ does belong to $\mathcal{S}^+(Q^+)$.

Parts (ii) and (iii) are immediate consequences of Definition 7.2.1. □

Let $S \in \mathcal{S}(Q)$ and let $\mathcal{E}^\perp(S)$ be the annulator of $\mathcal{E}(S)$ as follows:

$$\mathcal{E}^\perp(S) = \{\eta \in E^* \mid \langle \eta, h\rangle = 0 \ \ \forall h \in \mathcal{E}(S)\}.$$

This subspace of $E^*$ can be naturally identified with the space conjugate to $E/\mathcal{E}(S)$. Note that $(\cdot,\cdot)_{S,x}$ is, in fact, a *nondegenerate* scalar product on the latter space, as well as $\| \cdot \|_{S,x}$ is an Euclidean norm on this space. This norm defines the conjugate norm $\| \cdot \|_{S,x}^*$ on the space $(E/\mathcal{E}(S))^*$, i.e., on $\mathcal{E}^\perp(S)$. Note that, in the case of nondegenerate $S'$ (i.e., when $\mathcal{E}(S) = \{0\}$), we have

$$\| \phi \|_{S,x}^* = \left\langle \phi, [\hat{S}(x)]^{-1}\phi \right\rangle^{1/2}, \qquad \phi \in E^*.$$

We call an operator $S \in \mathcal{S}(Q)$ *regular* (notation: $S \in \mathcal{R}(Q)$) if $S(x) \in \mathcal{E}^\perp(S)$ for all $x \in Q$. By $\mathcal{R}^+(Q)$, we denote the subset of $\mathcal{R}(Q)$ formed by strongly self-concordant operators from the above set.

**Proposition 7.2.3** *Let $S \in \mathcal{R}(Q)$, $x \in Q$ and $\eta \in \mathcal{E}^\perp(S)$. Then the linear operators $S'(x)$ and $\hat{S}(x)$ map $E$ onto $\mathcal{E}^\perp(S)$,* Ker $S'(x) =$ Ker $\hat{S}(x) = \mathcal{E}(S)$, *and*

$$\| S'(x)e \|_{S,x}^* \geq \| e \|_{S,x}, \| \hat{S}(x)e \|_{S,x}^* = \| e \|_{S,x} .$$

*In particular, the equation*

$$S'(x)h = \eta \tag{7.2.8}$$

*is solvable, and every solution $h$ to it satisfies the relation*

$$\| h \|_{S,x} \leq \| \eta \|_{S,x}^* . \tag{7.2.9}$$

*Besides this, $S$ is constant on each set of the form $(x + \mathcal{E}(S)) \bigcap Q$.*

**Proof.** Since $\langle S(y), h\rangle \equiv 0$, $y \in Q$, $h \in \mathcal{E}(S)$, we have

$$\langle S'(y)e, h\rangle \equiv 0, \quad y \in Q, \quad e \in E, \quad h \in \mathcal{E}(S),$$

so that Im $S'(x) \subseteq \mathcal{E}^{\perp}(S)$. If $e \in$ Ker $S'(x)$, then $\langle S'(x)e, e\rangle = \left\langle \hat{S}(x)e, e\right\rangle = 0$, so that $e \in \mathcal{E}(S)$. Thus, Ker $S'(x) \subset \mathcal{E}(S)$. The above inclusions lead to the relations

$$\dim E = \dim \mathrm{Ker}\ S'(x) + \dim \mathrm{Im}\ S'(x) \leq \dim \mathcal{E}(S) + \dim \mathcal{E}^{\perp}(S) = \dim E,$$

so that Ker $S'(x) = \mathcal{E}(S)$, Im $S'(x) = \mathcal{E}^{\perp}(S)$. In particular, (7.2.8) is solvable. We also have

$$\| S'(x)e \|^*_{S,x} \| e \|_{S,x} \geq \langle S'(x)e, e\rangle = \| e \|^2_{S,x}$$

(the definitions of $\| \cdot \|_{S,x}$ and $\| \cdot \|^*_{S,x}$), so that $\| S'(x)e \|^*_{S,x} \geq \| e \|_{S,x}$ . The latter relation immediately implies (7.2.9). It remains to prove that Ker $\hat{S}(x) = \mathcal{E}(S)$, Im $\hat{S}(x) = \mathcal{E}^{\perp}(S)$ and that $\| \hat{S}(x)e \|^*_{S,x} = \| e \|_{S,x}$ . We have

$$\mathrm{Ker}\ \hat{S}(x) = \mathcal{E}(S)$$

(Proposition 7.2.1(i)); the latter relation, combined with $(\hat{S}(x))^* = \hat{S}(x)$, implies that

$$\mathrm{Im}\ \hat{S}(x) = \mathcal{E}^{\perp}(S).$$

The relation $\left\langle \hat{S}(x)h, h\right\rangle = \| h \|^2_{S,x}$ means precisely that $\| \hat{S}(x)e \|^*_{S,x} = \| e \|_{S,x}$. Since Ker $S'(\cdot) \equiv \mathcal{E}(S)$, $S$ is constant on the sets $(x + \mathcal{E}(S)) \bigcap Q$. □

The above proposition can be inverted: If $S \in \mathcal{S}(Q)$, Im $S'(x) \subset \mathcal{E}^{\perp}(S)$, $x \in Q$, and $S(x_0) \in \mathcal{E}^{\perp}(S)$ for some $x_0$, then also $S \in \mathcal{R}(Q)$. Indeed, let us fix $h \in \mathcal{E}(S)$ and let $f(x) = \langle S(x), h\rangle$. We have $Df(x)[e] = \langle S'(x)e, h\rangle = 0$, so that $f(x)$ is constant; since $f(x_0) = 0$, we conclude that $f(x) \equiv 0$. Thus, $\langle S(x), h\rangle \equiv 0$, $h \in \mathcal{E}(S)$, so that $S$ is regular.

Given a self-concordant on $Q$ operator $S$, we can associate with the operator the equation

$$S(x) = 0. \tag{7.2.10}$$

In the remaining part of this section, we focus on the properties of the Newton method as applied to the above equation. To describe these properties, it is convenient to measure accuracy of an approximate solution $x$ to (7.2.10) via the quantity ("Newton's decrement of $S$ at $x$ "), as follows:

$$\nu(S, x) = \max\{\langle S(x), h\rangle \mid h \in E, \| h \|_{S,x} \leq 1\} \leq \infty.$$

Note that, in the case of $S = f'$, $f$ being a 1-self-concordant function, $\nu(S, x) = \lambda(f, x)$ is the Newton decrement of $f$ at $x$; see §2.2.

It is clear that $\nu(S, x) \geq 0$ and $x \in Q$ is a solution to (7.2.10) if and only if $\nu(S, x) = 0$. The relation $\nu(S, x) < \infty$ means precisely that $S(x) \in \mathcal{E}^{\perp}(S)$, and, in the latter case, we clearly have

$$\nu(S, x) = \| S(x) \|^*_{S,x} . \tag{7.2.11}$$

**Proposition 7.2.4** *Let $S \in \mathcal{R}(Q)$. Then the function $\nu(S,x)$ is finite continuous function on $Q$.*

It is an immediate corollary of (7.2.11).

**Remark 7.2.1** *If $S \in \mathcal{S}^+(Q)$ and $Q$ is bounded, then $\mathcal{E}(S) = \{0\}$ (see (7.2.3)); it follows that, in the case of $S \in \mathcal{S}^+(Q)$ and bounded $Q$, we have $\mathcal{R}^+(Q) = \mathcal{S}^+(Q)$, and the function $\nu(S,\cdot)$ is finite and continuous on $Q$.*

Let $S \in \mathcal{R}(Q)$. In view of Proposition 7.2.3, the Newton equation

$$S'(x)h = S(x)$$

is solvable for every $x \in Q$, so that there exists a function $e(S,x) : Q \to E$ such that

$$S'(x)e(S,x) \equiv S(x).$$

The value of $e$ at every $x$ is uniquely defined, up to addition of an element from $\mathcal{E}(S)$ (see Proposition 7.2.3), and

$$\| e(S,x) \|_{S,x} \leq \nu(S,x). \tag{7.2.12}$$

Let us introduce the function

$$w(\lambda; s) = \frac{1}{1-s\lambda}\left\{(1-s)\lambda + \int_0^{s\lambda} \frac{t(2-t)}{(1-t)^2}dt\right\}, \tag{7.2.13}$$

defined on the set

$$\{(\lambda, s) \mid \lambda \geq 0,\ 1 \geq s \geq 0,\ s\lambda < 1\}.$$

Let

$$w^*(\lambda) = \min\{w(\lambda; s) \mid 0 \leq s \leq 1,\ s\lambda < 1\}.$$

It is easily seen that $w^*(\lambda)$ is well defined, continuous on the set $\{\lambda \geq 0\}$, and that

$$w^*(0) = 0, w^*(\lambda) < \lambda, \qquad 0 < \lambda < 1. \tag{7.2.14}$$

Furthermore, for $\lambda > 0$, let $\sigma(\lambda)$ be the minimizer of $w(\lambda, s)$ with respect to $s$ running over the set $\{s \mid 0 \leq s \leq 1, s\lambda < 1\}$ and let $\sigma(0) = 1$. We can prove straightforwardly that there exists an *absolute* constant $\kappa > 1$ such that

$$\begin{aligned} &w^*(\lambda) \leq \kappa\lambda^2, \qquad 0 \leq \lambda \leq 1, \\ &\sigma(\lambda) = 1, \quad \lambda \leq \frac{1}{5\kappa}. \end{aligned} \tag{7.2.15}$$

Thus,

$$\sigma(\lambda) \in (0,1],\ w(\lambda, \sigma(\lambda)) = w^*(\lambda) < \lambda, \qquad 0 < \lambda < 1. \tag{7.2.16}$$

The following statement is the basis for further consideration.

**Theorem 7.2.1** *Let $S \in \mathcal{R}^+(Q)$.*

(i) *For $x \in Q$ the Newton-type iterate of $x$, namely, the point*

$$x^+ = x^+(S,x) = x - \sigma(\nu(S,x))e(S,x) \tag{7.2.17}$$

*belongs to $Q$ and*

$$\nu(S,x^+) \le w^*(\nu(S,x)). \tag{7.2.18}$$

(ii) *Equation* (7.2.10) *is solvable if and only if there exists an $x \in Q$ with $\nu(S,x) < 1$. If this is the case, then the sequence of the Newton-type iterates*

$$x_i = x^+(S, x_{i-1})$$

*associated with an arbitrary $x_0 \in Q$ satisfying the relation $\nu(S,x_0) < 1$ converges to the solution of the equation $S(x) = 0$ in the sense that*

$$\lambda_i \equiv \nu(S,x_i) \le w^*(\lambda_{i-1}) \to 0, \qquad i \to \infty.$$

*Besides this, if $x \in Q$ satisfies the relation*

$$\nu(S,x) \le \frac{1}{5\kappa}, \tag{7.2.19}$$

*then*

$$\| x - x^* \|_{S,x^*} \le 5\nu(S,x) \tag{7.2.20}$$

*for every solution $x^*$ to* (7.2.10).

(iii) *Let the image of $Q$ in the factor-space $E/\mathcal{E}(S)$ be bounded. Then* (7.2.10) *is solvable, and the set of solutions to the equation is of the form $x^* + \mathcal{E}(S)$.*

**Proof.** (i) Let us fix $x \in Q$ and denote

$$e = e(S,x), \quad \sigma = \sigma(\nu(S,x)), \quad \lambda = \nu(S,x).$$

We have $\| e \|_{S,x} \le \lambda$ (see (7.2.12)), and, since $\omega \equiv \sigma\lambda < 1$ (the definition of $\sigma(\lambda)$), we obtain $x^+ \equiv x - \sigma e \in Q$ (see Proposition 7.2.1(ii)). Let $h = x - x^+ \equiv \sigma e$, so that $\| h \|_{S,x} \le \omega$, and let $g \in E$ be such that $\| g \|_{S,x} \le 1$. Let

$$f_g(t) = \langle S(x - th), g \rangle, \qquad 0 \le t \le 1.$$

We have

$$f_g'(t) = -\langle S'(x-th)h, g \rangle = -\langle S'(x)h, g \rangle + \langle (S'(x) - S'(x-th))h, g \rangle.$$

Furthermore, $S'(x)h = \sigma S(x)$ (the definition of $h$) and

$$\begin{aligned}|\langle (S'(x) - S'(x-th))h, g \rangle| &\le t \left\{ \frac{1}{(1 - t \| h \|_{S,x})^2} - 1 \right\} \| h \|_{S,x} \| g \|_{S,x} \\ &\le t\omega \left\{ \frac{1}{(1-t\omega)^2} - 1 \right\}\end{aligned}$$

(see (7.2.4)). Thus,

$$|f_g'(t) + \langle \sigma S(x), g\rangle \,| \le t\omega \left\{ \frac{1}{(1-t\omega)^2} - 1 \right\}.$$

It follows that

$$\begin{aligned} |f_g(1) - f_g(0) + \sigma \,\langle S(x), g\rangle \,| &\le \omega \int_0^1 t\{(1-t\omega)^{-2} - 1\} dt \\ &= \frac{1}{\omega} \int_0^\omega \frac{r^2(2-r)}{(1-r)^2} dr \le \int_0^\omega r(2-r)(1-r)^{-2} dr, \end{aligned}$$

or, which is the same,

$$|\,\langle S(x^+) - (1-\sigma) S(x), g\rangle \,| \le \int_0^\omega \frac{r(2-r)}{(1-r)^2} dr.$$

Since $g$ is an arbitrary vector with $\| g \|_{S,x} \le 1$, it follows that

$$\| S(x^+) \|_{S,x}^* \le (1-\sigma) \| S(x) \|_{S,x}^* + \int_0^\omega \frac{r(2-r)}{(1-r)^2} dr. \tag{7.2.21}$$

From the second relation in (7.2.4), it follows that

$$\| q \|_{S,x} \le \frac{\| q \|_{S,x^+}}{1-\omega}, \qquad q \in E.$$

Therefore, for every $\eta \in \mathcal{E}^\perp(S)$, we have

$$\begin{aligned} \| \eta \|_{S,x^+}^* = \max\{\langle \eta, q\rangle \mid \| q \|_{S,x^+} \le 1\} &\le \max\left\{\langle \eta, q\rangle \mid \| q \|_{S,x} \le \frac{1}{1-\omega}\right\} \\ &= \frac{\| \eta \|_{S,x}^*}{1-\omega}. \end{aligned}$$

Thus, (7.2.21) implies that

$$\| S(x^+) \|_{S,x^+}^* \le \frac{1}{1-\omega} \left\{ (1-\sigma) \| S(x) \|_{S,x}^* + \int_0^\omega \frac{r(2-r)}{(1-r)^2} dr \right\}. \tag{7.2.22}$$

To simplify notation, set $\lambda^+ = \| S(x^+) \|_{S,x^+}^*$; since $\lambda \equiv \nu(S,x) = \| S(x) \|_{S,x}^*$, relation (7.2.22) can be rewritten as

$$\lambda^+ \le \frac{1}{1-\sigma\lambda} \left\{ (1-\sigma)\lambda + \int_0^{\sigma\lambda} \frac{r(2-r)}{(1-r)^2} dr \right\}. \tag{7.2.23}$$

The right-hand side of (7.2.23) is precisely $w(\lambda;\sigma)$; since $\sigma = \sigma(\lambda)$, (7.2.23) implies (7.2.18). Part (i) is proved.

(ii) If (7.2.10) is solvable and $x^*$ is a solution to this equation, then $S(x^*) = 0$, and therefore $\nu(S,x^*) = 0$. Now let $x_0 \in Q$ be such that $\nu(S,x_0) < 1$ and let

$$x_i = x^+(S,x_{i-1}), \qquad i \geq 1.$$

According to (i), $\lambda_i \equiv \nu(S,x_i)$ satisfy the relation $\lambda_i \leq w^*(\lambda_{i-1})$. The latter quantity is $< \lambda_{i-1} < 1$ in the case of $\lambda_{i-1} > 0$; in the opposite case, $x_j \equiv x_{i-1}$, $j \geq i-1$ is a solution to (7.2.10). Since $w^*(\cdot)$ is continuous, we have $\lambda_i \to 0$, $i \to \infty$. It remains to prove that (7.2.10) is solvable. Let $\mathcal{E}'(S)$ be a complement to $\mathcal{E}(S)$ in $E$, so that $\mathcal{E}'(S) \bigcap \mathcal{E}(S) = \{0\}$ and $\mathcal{E}'(S) + \mathcal{E}(S) = E$. Recall that the vectors $e(S,x)$ are defined up to the addition of an element of $\mathcal{E}(S)$; in particular, we can choose them to belong to $\mathcal{E}'(S)$. Assume that this is the case. Since $\lambda_i$ converge to 0, there exists $i = i_0$ such that $\mu \equiv \lambda_{i_0} \leq (5\kappa)^{-1}$. Since $w^*(\lambda) \leq \kappa\lambda^2$, we conclude that the quantities $\mu_i \equiv \lambda_{i_0+i}$, $i \geq 0$ satisfy the relation $\mu_i \leq 5^{-i}\mu$. For $i > 0$, we have

$$\| x_{i_0+i} - x_{i_0+i-1} \|_{S,x_{i_0+i-1}} \leq \mu_{i-1}.$$

This inequality combined with (7.2.4) implies that

$$\rho_i \equiv \| x_{i_0+i} - x_{i_0+i-1} \|_{S,x_{i_0}} \leq \frac{\mu_{i-1}}{1 - \| x_{i_0+i-1} - x_{i_0} \|_{S,x_{i_0}}}$$

for all $i > 0$ such that $x_{i_0+i-1} \in W_1^0(S,x_{i_0})$. Thus, if

$$P_i = \| x_{i_0+i} - x_{i_0} \|_{S,x_{i_0}}, \qquad i \geq 0,$$

then

$$P_0 = 0; \quad P_i \leq P_{i-1} + \rho_i; \quad P_{i-1} < 1 \Rightarrow \rho_i \leq \frac{\mu_{i-1}}{1 - P_{i-1}}, \qquad i > 0.$$

Note that $\mu_{i-1} \leq 5^{-i+1}\mu \leq 5^{-i}\kappa^{-1}$. We claim that $P_i \leq \frac{1}{2}$ for all $i$. Indeed, $P_0 = 0 \leq \frac{1}{2}$; if $P_j \leq \frac{1}{2}$, $j < i$, then also $\rho_j \leq \mu_{j-1}(1 - P_{j-1})^{-1} = 2\kappa^{-1}5^{-j}$, $j \leq i$, so that $P_i \leq \sum_{j=1}^{i} 2\kappa^{-1}5^{-j} \leq \frac{1}{2}$. Thus, $P_i \leq \frac{1}{2}$, $i \geq 0$, and therefore

$$P_i \leq P_{i-1} + 2\mu_{i-1} \leq P_{i-1} + 2\mu 5^{1-i},$$

so that, in fact, $P_i \leq 2.5\mu$. We also have proved that $\sum_i \rho_i < \infty$. Since $x_{i_0+i} - x_{i_0} \in \mathcal{E}'(S)$ (our choice of $e(S,\cdot)$) and $\| \cdot \|_{S,x_{i_0}}$ is a norm on $\mathcal{E}'(S)$, the sequence $x_{i_0+i}$ converges to a point $x_\infty$ satisfying the relation $\| x_\infty - x_{i_0} \|_{S,x_{i_0}} \leq 2.5\mu \leq \frac{1}{2}$. In particular, $x_\infty \in W_{1/2}(S,x_{i_0})$. The latter set is contained in $Q$ (Proposition 7.2.1(ii)), so that $x_\infty \in Q$. Since $\nu(S,\cdot)$ is continuous on $Q$ and $\nu(S,x_i) \to 0$, we have $\nu(S,x_\infty) = 0$, so that $x_\infty$ is a solution to (7.2.10). The first statement in (ii) is proved.

To complete the proof of (ii), note that, in the case of $x_0 = x$, where $x$ is the point involved into (7.2.19), in the above considerations, we can set $i_0 = 0$, which implies that there exists a solution to (7.2.10), $x_\infty$, such that

$$\| x_\infty - x \|_{S,x} \leq 2.5\nu(S,x) \leq (2\kappa)^{-1} \leq \tfrac{1}{2}.$$

From (7.2.4), it follows that

$$\| x - x_\infty \|_{S,x_\infty} \leq 2 \| x_\infty - x \|_{S,x} \leq 5\nu(S,x).$$

To prove (7.2.20), it suffices to establish that, if $x^*$ is a solution to (7.2.10), possibly not coinciding with $x_\infty$, then $\| x - x_\infty \|_{S,x_\infty} = \| x - x^* \|_{S,x^*}$. Indeed, since $S$ is monotone and $x^*$, $x_\infty$ are solutions to (7.2.10), $S$ vanishes on the segment with the endpoints $x^*$ and $x_\infty$, so that $S'(x^*)e = 0$, $e = x_\infty - x^*$. The latter relation means that $e \in \mathcal{E}(S)$, so that $\| e \|_{S,x^*} = 0$ and $\| \cdot \|_{S,x_\infty} = \| \cdot \|_{S,x^*}$ (the latter relation follows from Proposition 7.2.1(i)). Thus, $\| x - x_\infty \|_{S,x_\infty} = \| x - x^* \|_{S,x^*}$.

(iii) From Proposition 7.2.1(i), it follows that $Q = Q + \mathcal{E}(S)$, while Proposition 7.2.3 implies that $S$ is constant along translations of $\mathcal{E}(S)$. These observations allow us to immediately reduce the statement under consideration to the case when $\mathcal{E}(S) = \{0\}$. Thus, it is enough to prove (iii) in the case when $Q$ is bounded and $\mathcal{E}(S) = \{0\}$. Let $\| \cdot \|$ be a fixed Euclidean norm on $E$, let $x_0 \in Q$, and let $R > 0$ be such that $Q$ is contained in the $\| \cdot \|$-ball of the radius $R$ centered at $x_0$. Let $B$ be the unit $\| \cdot \|$-sphere and let $r(\xi) : B \to \{t > 0\}$ be defined as

$$r(\xi) = \sup\{t \mid x_0 + t\xi \in Q\}.$$

For $\xi \in B$ and $0 \leq t < 1$, we have

$$\langle S(x_0 + tr(\xi)\xi), \xi \rangle = \int_0^t \langle S'(x_0 + sr(\xi)\xi) r(\xi)\xi, \xi \rangle \, ds + \langle S(x_0), \xi \rangle.$$

Furthermore,

$$\langle S'(x_0 + sr(\xi)\xi)\xi, \xi \rangle = \| \xi \|^2_{S,x_0+sr(\xi)\xi},$$

so that the vector

$$x_0 + sr(\xi)\xi + \tfrac{1}{2}\{\langle S'(x_0 + sr(\xi)\xi)\xi, \xi \rangle\}^{-1/2}\xi$$

belongs to $W_{1/2}(S, x_0 + sr(\xi)\xi)$ and therefore belongs to $Q$. It follows that

$$sr(\xi) + \tfrac{1}{2}\{\langle S'(x_0 + sr(\xi)\xi)\xi, \xi \rangle\}^{-1/2} \leq r(\xi),$$

so that

$$\langle S'(x_0 + sr(\xi)\xi)\xi, \xi \rangle \geq \frac{1}{4r^2(\xi)(1-s)^2}.$$

Thus,

$$\langle S(x_0 + tr(\xi)\xi), \xi\rangle \geq \int_0^t \frac{ds}{4r(\xi)(1-s)^2} + \langle S(x_0), \xi\rangle \equiv g(t, \xi).$$

We clearly have $g(t,\xi) \to \infty$, $t \to 1-0$, uniformly in $\xi \in B$. Thus, $\| S(x_0 + tr(\xi)\xi) \|^* \to \infty$ uniformly in $\xi \in B$ as $t \to 1-0$, where $\| \cdot \|^*$ is the norm dual to $\| \cdot \|$. It follows that $\| S(x_i) \|^* \to \infty$ along each sequence $x_i \in Q$ converging to a boundary point of $Q$. This observation, in turn, means that the set Argmin $\{\| S(x) \|^* \mid x \in Q\}$ is nonempty. Let $x^*$ be a point of the latter set. Note that Ker $S'(x^*) = \mathcal{E}(S)$ (Proposition 7.2.3), and, since $\mathcal{E}(S) = \{0\}$, $S'(x^*)$ is nondegenerate; therefore the only possibility for $x^*$ to be a minimizer of $\| S(x) \|^*$ is to be a solution to (7.2.10). Thus, (7.2.10) is solvable. □

## 7.3 Path-following method

In what follows, we focus on the problem of solving variational inequalities with monotone operators. Let $G$ be a closed convex subset of a finite-dimensional real vector space $E$ with a nonempty interior $G'$, and let $S : G' \to E^*$ be a $C^2$-smooth single-valued monotone operator. We can associate with the monotone element $(G, S)$ the problem of finding a (weak) solution to the variational inequality defined by $(G, S)$ :

$$(7.3.1) \qquad \mathcal{V}(G,S): \quad \text{find } x \in G \quad \text{s.t.} \quad \langle S(u), u - x\rangle \geq 0 \quad \forall u \in G'.$$

Since $G$ is bounded, the problem is solvable (Proposition 7.1.2), and the solution set to it, $\mathcal{V}^*(S, G)$, clearly is closed and convex.

Below, $G'$ denotes the interior of $G$.

### 7.3.1 Operators compatible with a self-concordant barrier

Let $F$ be a $\vartheta$-self-concordant barrier for a closed convex domain $G \subseteq E$. Since the three-linear functional $D^3F(x)[h_1, h_2, h_3]$ is symmetric, from the fact that $F$ is strongly 1-self-concordant on int $G$, it follows that

$$(7.3.2) \qquad |D^3F(x)[h_1, h_2, h_3]| \leq 2\prod_{i=1}^{3}\{\langle F''(x)h_i, h_i\rangle\}^{1/2},$$

$h_i \in E$, $i = 1, 2, 3$, and, since $F$ is a $\vartheta$-self-concordant barrier, we also have

$$(7.3.3) \qquad \langle F'(x), h\rangle \leq \vartheta^{1/2} \quad \text{for all } h \text{ such that } \langle F''(x)h, h\rangle \leq 1.$$

Thus, the monotone operator $\Sigma(x) = F'(x)$ is self-concordant on $G'$. Relation (7.3.3) means that $\nu(\Sigma, x) \leq \vartheta^{1/2}$, $x \in G'$. From the properties of self-concordant barriers established in §2.3, it also follows that $\Sigma$ is strongly self-concordant.

Let us start with the following definition (cf. §3.2).

**Definition 7.3.1** *Let $F$ be a $\vartheta$-self-concordant barrier for $G$ and let $\beta \geq 0$. A $C^2$-smooth monotone operator $S : G' \to E^*$ is called $\beta$-compatible with $F$ (notation: $S \in \mathcal{C}_\beta(G,F)$) if, first, the inequality*

$$(7.3.4) \qquad |S''(x)[h_1,h_2,h_3]| \leq \beta \prod_{i=1}^{3} \{3\left\langle \hat{S}(x)h_i, h_i \right\rangle\}^{1/3} \{3\left\langle F''(x)h_i, h_i\right\rangle\}^{1/6}$$

*holds for every $x \in G'$ and all $h_1, h_2, h_3 \in E$ and, second, the inclusion*

$$(7.3.5) \qquad S(x) \in (E_F)^\perp, \qquad x \in G'$$

*holds (recall that $E_F =$ Ker $F''(x)$ does not depend on $x \in G'$ and $(E_F)^\perp$ is the annulator of $E_F$).*

**Proposition 7.3.1** (i) *$\mathcal{C}_\beta(G,F)$ is a cone: If $S_1$, $S_2 \in \mathcal{C}_\beta(G,F)$, $\alpha_1, \alpha_2 \geq 0$, then also $\alpha_1 S_1 + \alpha_2 S_2 \in \mathcal{C}_\beta(G,F)$; if $\beta' \geq \beta$, then $\mathcal{C}_\beta(G,F) \subseteq \mathcal{C}_{\beta'}(G,F)$.*

*If $S_i \in \mathcal{C}_{\beta_i}(G_i,F_i)$, $i = 1,2$, and* $\mathrm{int}(G_1 \bigcap G_2) \neq \emptyset$, *then*

$$(S_1+S_2)\mid_{G_1 \bigcap G_2} \in \mathcal{C}_{\max\{\beta_1,\beta_2\}}(G_1 \bigcap G_2, (F_1+F_2)\mid_{G_1 \bigcap G_2}).$$

(ii) *If $S$ is a linear monotone mapping taking values in $(E_F)^\perp$, then $S \in \mathcal{C}_0(G,F)$.*

(iii) *If $\mathcal{A}(y) = Ay + b$ is an affine mapping from a finite-dimensional space $E^+$ into $E$ such that $\mathcal{A}(E^+) \bigcap G' \neq \emptyset$, $F^+(y) = F(\mathcal{A}(y)) : \mathrm{int}\,\mathcal{A}^{-1}(G) \to \mathbf{R}$ and $S \in \mathcal{C}_\beta(G,F)$, then $A^* S(\mathcal{A}(\cdot)) \in \mathcal{C}_\beta(\mathcal{A}^{-1}(G), F^+)$.*

**Proof.** (i) To prove that $\mathcal{C}_\beta(G,F)$ is a cone, note first that $\alpha S \in \mathcal{C}_\beta(G,F)$ whenever $\alpha \geq 0$ and $S \in \mathcal{C}_\beta(G,F)$ (evident). It remains to prove that $S_1 + S_2 \in \mathcal{C}_\beta(G,F)$ whenever $S_1$, $S_2 \in \mathcal{C}_\beta(G,F)$. Indeed, $S_1 + S_2$ clearly takes values in $(E_F)^\perp$, and all we must establish is that the sum satisfies (7.3.4). We have

$$\begin{aligned}
|(S_1+S_2)''(x)[h_1,h_2,h_3]| &\leq \beta \left\{ \prod_{i=1}^{3} \{3\left\langle \hat{S}_1(x)h_i, h_i\right\rangle\}^{1/3} \prod_{i=1}^{3} \{3\left\langle F''(x)h_i,h_i\right\rangle\}^{1/6} \right. \\
&\qquad \left. + \prod_{i=1}^{3} \{3\left\langle \hat{S}_2(x)h_i, h_i\right\rangle\}^{1/3} \prod_{i=1}^{3} \{3\left\langle F''(x)h_i,h\right\rangle\}^{1/6} \right\} \\
&= \beta \prod_{i=1}^{3} \{3\left\langle F''(x)h_i,h_i\right\rangle\}^{1/6} \left\{ \prod_{i=1}^{3} \{3\left\langle \hat{S}_1(x)h_i, h_i\right\rangle\}^{1/3} \right. \\
&\qquad \left. + \prod_{i=1}^{3} \{3\left\langle \hat{S}_2(x)h_i, h_i\right\rangle\}^{1/3} \right\}.
\end{aligned}$$

In view of the latter relation, to prove that $S_1 + S_2 \in \mathcal{C}_\beta(G,F)$, it suffices to verify that

$$\begin{aligned}
&\prod_{i=1}^{3} \left\{ \left\langle \hat{S}_1(x)h_i, h_i\right\rangle \right\}^{1/3} + \prod_{i=1}^{3} \left\{ \left\langle \hat{S}_2(x)h_i, h_i\right\rangle \right\}^{1/3} \\
&\leq \prod_{i=1}^{3} \left\{ \left( \left\langle \hat{S}_1(x)h_i, h_i\right\rangle + \left\langle \hat{S}_2(x)h_i, h_i\right\rangle \right) \right\}^{1/3}.
\end{aligned}$$

Let $a_i = \left\langle \hat{S}_1(x)h_i, h_i \right\rangle^{1/3}$ and $b_i = \left\langle \hat{S}_2(x)h_i, h_i \right\rangle^{1/3}$ (these quantities are non-negative). We must prove that

$$\prod_{i=1}^{3} a_i + \prod_{i=1}^{3} b_i \le \prod_{i=1}^{3} (a_i^3 + b_i^3)^{1/3}.$$

We can restrict ourselves to the case of $a_i^3 + b_i^3 = 1$, $i = 1, 2, 3$ (it suffices to perform the updating $(a_i, b_i) \to (a_i(a_i^3 + b_i^3)^{-1/3},\ b_i(a_i^3 + b_i^3)^{-1/3})$). We have

$$\prod_{i=1}^{3} a_i \le \frac{1}{3}(a_1^3 + a_2^3 + a_3^3), \qquad \prod_{i=1}^{3} b_i \le \frac{1}{3}(b_1^3 + b_2^3 + b_3^3),$$

so that

$$\prod_{i=1}^{3} a_i + \prod_{i=1}^{3} b_i \le \frac{1}{3}(a_1^3 + a_2^3 + a_3^3 + b_1^3 + b_2^3 + b_3^3) = 1.$$

Thus, $\mathcal{C}_\beta(G, F)$ is a cone. The inclusion $\mathcal{C}_\beta(G, F) \subseteq \mathcal{C}_{\beta'}(G, F)$ and the last statement in (i) are evident.

Items (ii) and (iii) are straightforward consequences of definitions. □

**Proposition 7.3.2** *Let $S \in \mathcal{C}_\beta(G, F)$ and let $t > 0$. Set*

$$S_t(x) = (1 + \beta)^2 \{tS(x) + F'(x)\} : G' \to E^*.$$

*Then $S_t \in \mathcal{R}^+(G')$ and $\mathcal{E}(S_t) = E_F$.*

**Proof.** Let us verify first that $S_t \in \mathcal{S}(G')$. Let $x \in G'$ and $h_i \in E$, $i = 1, 2, 3$. We have (see (7.3.2))

$$|S_t''(x)[h_1, h_2, h_3]| \le (1 + \beta)^2 \left\{ t\beta \prod_{i=1}^{3} \{3 \left\langle \hat{S}(x)h_i, h_i \right\rangle\}^{1/3} \prod_{i=1}^{3} \{3 \langle F''(x)h_i, h_i \rangle\}^{1/6} + 2 \prod_{i=1}^{3} \langle F''(x)h_i, h_i \rangle^{1/2} \right\}.$$

We must prove that the right-hand side of the latter inequality is not greater than

$$2 \prod_{i=1}^{3} \left\{ \left\langle (1 + \beta)^2 \hat{S}_t(x)h_i, h_i \right\rangle \right\}^{1/2} = 2(1+\beta)^3 \prod_{i=1}^{3} \{t \langle S'(x)h_i, h_i \rangle + \langle F''(x)h_i, h_i \rangle\}^{1/2}.$$

Due to homogeneity with respect to $h_i$, we clearly can restrict ourselves to the case of $\langle F''(x)h_i, h_i \rangle = 1$. Thus, it suffices to verify that

$$3^{3/2} t\beta \prod_{i=1}^{3} p_i^{1/3} + 2 \le 2(1 + \beta) \prod_{i=1}^{3} \{tp_i + 1\}^{1/2},$$

where $p_i = \{\langle S'(x)h_i, h_i\rangle\}$. The inequality is evident in the case of $p_1p_2p_3 = 0$. Now let $p_i > 0$, $i = 1, 2, 3$. Substitution $p_i = \exp\{s_i\}$ transforms the inequality under consideration into

$$3^{3/2}t\beta\exp\{(s_1 + s_2 + s_3)/3\} + 2 \le 2(1+\beta)\prod_{i=1}^{3}(t\exp\{s_i\} + 1)^{1/2};$$

since $\prod_{i=1}^{3}\{t\exp\{s_i\} + 1\}$ is convex and symmetric in $s = (s_1, s_2, s_3)$, we have

$$2(1+\beta)\prod_{i=1}^{3}(t\exp\{s_i\} + 1)^{1/2} \ge 2(1+\beta)(1 + t\exp\{(s_1 + s_2 + s_3)/3\})^{3/2}.$$

Thus, proving our inequality is the same as proving that $2(1+\beta)(1+tp)^{3/2} \ge 3^{3/2}t\beta p+2$ for all $t, p > 0$. The latter inequality can be proved straightforwardly.

Thus, $S_t \in \mathcal{S}(G')$. Since $F' \in \mathcal{S}^+(G')$ and $\langle S'(\cdot)h, h\rangle \ge 0$, the inclusion $S_t \in \mathcal{S}(G')$ implies $S_t \in \mathcal{S}^+(G')$. It remains to prove that $S_t$ is regular and that $\mathcal{E}(S_t) = E_F$.

Since $F$ is a self-concordant barrier, $F'(x) \in (E_F)^\perp$; since $S(x) \in (E_F)^\perp$ (the definition of compatibility), we have $S_t(x) \in (E_F)^\perp$. If $h \in E_F$, then $\langle S_t(x), h\rangle \equiv 0, x \in G'$, so that $\langle S_t'(x)e, h\rangle = 0, e \in E$. We conclude that $\left\langle \hat{S}_t(x)h, h\right\rangle = 0, h \in E_F$, so that $\mathcal{E}(S_t) \supset E_F$; at the same time, $\hat{S}_t(x) - (1+\beta)^2F''(x)$ is symmetric positive semidefinite, whence $E_F \supset \mathcal{E}(S_t)$. Thus, $S_t$ is regular and $\mathcal{E}(S_t) = E_F$. □

### 7.3.2 Barrier-generated family

Let $F$ be a $\vartheta$-self-concordant barrier for $G$ and let $S$ be a $\beta$-compatible with the barrier monotone operator. We henceforth assume that $G$ is bounded. To solve the variational inequality generated by $S$, we can act as follows: Consider the family

$$\{S_t(x) = (1+\beta)^2(tS(x) + F'(x)) : G' \to E^*\}_{t>0};$$

according to Proposition 7.3.2, this family is comprised of strongly self-concordant monotone mappings with the property

$$\mathcal{E}(S_t) \equiv E_F = \{0\} \tag{7.3.6}$$

(the latter equality follows from the fact that $F$ is a self-concordant barrier for *bounded* $G$; see Proposition 2.3.2).

From Theorem 7.2.1, it follows that the equations

$$(S_t): \qquad S_t(x) = 0$$

have uniquely defined solutions $x^*(t) \in G'$ for all $t > 0$. It is clear that, for large $t$, equation $(S_t)$ "is close" to the variational inequality $\mathcal{V}(G, S)$, so that we can expect that the trajectory $x^*(t)$ converges, in a natural sense, to the set of solutions to $\mathcal{V}(G, S)$. The following statement demonstrates what type of convergence we can ensure.

**Proposition 7.3.3** *Let $S \in \mathcal{C}_\beta(G,F)$ and let $G$ be bounded. Then*

$$(7.3.7) \qquad (\forall y \in G): \qquad \langle S(x^*(t)), x^*(t) - y\rangle \le \frac{\vartheta}{t}.$$

**Proof.** We have $S_t(x^*(t)) = 0$ or $S(x^*(t)) = -t^{-1}F'(x^*(t))$. Since $F$ is a $\vartheta$-self-concordant barrier, we have $\langle F'(x), y - x\rangle \le \vartheta$, $x \in G'$, $y \in G$ (see (2.3.2)), which immediately implies (7.3.7). □

Note that (weak) solutions to $\mathcal{V}(G,S)$ are precisely the points of $G$ at which the (clearly nonnegative) function

$$\varepsilon(S,x) \equiv \sup\{\langle S(y), x - y\rangle \mid y \in G\} : G' \to \mathbf{R}\bigcup\{+\infty\}$$

equals zero; thus, $\varepsilon(S,x)$ can be taken as a natural accuracy measure for the points of $G$ regarded as approximate solutions to the variational inequality. Note that the monotonicity of $S$ implies that if

$$\varepsilon^+(S,x) = \sup\{\langle S(x), x - y\rangle \mid y \in G\} : G' \to \mathbf{R}\bigcup\{+\infty\},$$

then

$$(7.3.8) \qquad \varepsilon(S,x) \le \varepsilon^+(S,x), \qquad x \in G';$$

thus, $\varepsilon^+(S,\cdot)$ is even "more strong" accuracy measure than $\varepsilon(S,\cdot)$.

Now note that (7.3.7) means precisely that

$$(7.3.9) \qquad \varepsilon^+(S, x^*(t)) \le \frac{\vartheta}{t},$$

so that Proposition 7.3.3 gives us an upper bound for the naturally measured error of $x^*(t)$ regarded as an approximate solution to $\mathcal{V}(G,S)$.

### 7.3.3 Updating rule

The method we describe (cf. §3.2) forms a sequence of approximations $x(t)$ to the points $x^*(t)$ along an increasing in certain ratio sequence $t = t_i$ of values of $t$. The approximations are $\lambda$-*tight*; i.e., the predicate

$$\mathsf{P}(S,t,\lambda;x) = \{x \in G'\} \quad \text{and} \quad \{\nu(S_t,x) \le \lambda\}$$

holds at the pairs $(t_i, x(t_i))$, where $\lambda$ is an arbitrary constant in the interval $(0, 1/(25\kappa))$ (the choice of $\lambda$ determines the rate at which $t_i$ are varied). The method is based on the following updating rule:

$\mathsf{N}(\lambda)$ : given $(t,x)$ with $t>0$ and $x$ satisfying the predicate $\mathsf{P}(S,t,\lambda;x)$, choose

$$t^+ = t(1+\gamma(\lambda,\vartheta,\beta)), \qquad \gamma(\lambda,\vartheta,\beta) = \frac{3\lambda}{2\lambda\left(\lambda+(1+\beta)\vartheta^{1/2}\right)}$$

and set

$$x^+ = x - e(S_{t^+},x),$$

where $e(S_\tau,x)$ is the Newton direction of $S_\tau$ at $x$(i.e., the solution to the *Newton equation* $S'_\tau(x)u = S_\tau(x)$).

The following statement demonstrates that this updating rule maintains the predicate $\mathsf{P}(S,t,\lambda;x)$; note that now we do not assume that $G$ is bounded.

**Proposition 7.3.4** *Let $S \in \mathcal{C}_\beta(G,F)$ and let $t>0$, $\lambda \in (0,1/(25\kappa))$. Assume that $x$ satisfies $\mathsf{P}(S,t,\lambda;x)$ and that $t^+>0$ is such that*

$$|t^+/t - 1| \le \gamma(\lambda,\vartheta,\beta). \tag{7.3.10}$$

*Then $x$ satisfies the predicate $\mathsf{P}(S,t^+,5\lambda;x)$, and the Newton iterate $x^+ = x - e(S_{t^+},x)$ of $x$ is well defined and satisfies the predicate $\mathsf{P}(S,t^+,\lambda;x^+)$.*

**Proof**. We have

$$\| h \|_{S_{t^+},x} \ge \frac{\| h \|_{S_t,x}}{1+|t^+/t-1|}, \qquad h \in E$$

(evident), and $\mathcal{E}(S_t) = \mathcal{E}(S_{t^+}) = E_F$ (see Proposition 7.3.2), so that

$$\| \eta \|^*_{S_{t^+},x} \le (1+|t^+/t-1|) \| \eta \|^*_{S_t,x}, \qquad \eta \in (E_F)^\perp. \tag{7.3.11}$$

Let $x$ satisfy $\mathsf{P}(S,t,\lambda;x)$, so that $\| tS(x)+F'(x) \|^*_{S_t,x} \le (1+\beta)^{-2}\lambda$. As we have seen (Proposition 7.3.2), $S(x)$, $F'(x) \in (E_F)^\perp$, and the above inequality implies that

$$\| tS(x) \|^*_{S_t,x} \le \| F'(x) \|^*_{S_t,x} + \frac{\lambda}{(1+\beta)^2}.$$

At the same time,

$$\| F'(x) \|^*_{S_t,x} \le \frac{1}{1+\beta} \| F'(x) \|^*_{F'(\cdot),x}$$

(evident), and $\| F'(x) \|^*_{F'(\cdot),x} \le \vartheta^{1/2}$ (the definition of a $\vartheta$-self-concordant barrier). We conclude that $\| tS(x) \|^*_{S_t,x} \le (1+\beta)^{-2}\lambda + (1+\beta)^{-1}\vartheta^{1/2}$. Therefore

$$\| (t^+-t)S(x) \|^*_{S_t,x} = |t^+/t-1| \, \| tS(x) \|^*_{S_t,x} \le |t^+/t-1| \left( \frac{\lambda}{(1+\beta)^2} + \frac{\vartheta^{1/2}}{1+\beta} \right).$$

It follows that

$$\| t^+S(x) + F'(x) \|^*_{S_t,x} \leq \| tS(x) + F'(x) \|^*_{S_t,x} + \| (t^+ - t)S(x) \|^*_{S_t,x}$$
$$\leq \frac{1}{(1+\beta)^2}\left\{\lambda + |t^+/t - 1|\left(\lambda + (1+\beta)\vartheta^{1/2}\right)\right\},$$

which, combined with (7.3.11), leads to

$$\| t^+S(x) + F'(x) \|^*_{S_{t^+},x}$$
$$\leq \frac{1}{(1+\beta^2}\left\{\lambda + |t^+/t - 1|\left(\lambda + (1+\beta)\vartheta^{1/2}\right)\right\}(1 + |t^+/t - 1|).$$

The right-hand side in the latter relation is $\leq 5\lambda(1+\beta)^{-2}$ by virtue of (7.3.10), so that $\mathsf{P}(S, t^+, 5\lambda; x)$ is true. Since $\mathsf{P}(S, t^+, 5\lambda; x)$ is true and $S_{t^+} \in \mathcal{R}^+(G')$ (see Proposition 7.3.2), from Theorem 7.2.1, it follows that $x^+$ is well defined, belongs to $G'$, and $\nu(S_{t^+}, x^+) \leq \kappa(5\lambda)^2 \leq \lambda$ (we considered that $\kappa\lambda \leq \frac{1}{25}$). □

### 7.3.4 Initialization

To follow the path $x^*(t)$ with the aid of the above updating rule $\mathsf{N}(\lambda)$, we must initialize the procedure, i.e., to find an initial pair $(t_0, x_0)$, $t_0 > 0$, such that the predicate $\mathsf{P}(S, t, \lambda; x)$ holds true at $t = t_0$, $x = x_0$. To find such a pair, we can use the same approach as in the case of the barrier-generated path-following method from §3.2, namely, we can approximate the minimizer $x(F)$ of the barrier $F$. Recall that $G$ is assumed to be bounded. We know (see Proposition 2.3.2(ii)) that, under the latter assumption, $x(F)$ is uniquely defined; we clearly have $\nu(F', x(F)) = 0$. Now let $x_0$ be close enough to $x(F)$, namely, let

$$\nu(F', x_0) \leq \frac{\lambda}{2(1+\beta)}. \tag{7.3.12}$$

Define $t_0$ as

$$t_0 = \frac{\lambda}{2(1+\beta) \| S(x_0) \|^*_{F',x_0}}. \tag{7.3.13}$$

Then clearly

$$\nu(S_{t_0}, x_0) \leq \lambda,$$

so that $(t_0, x_0)$ is the desired pair satisfying $\mathsf{P}(S, t_0, \lambda; x_0)$.

Thus, to initialize the above process, it suffices to find an approximation $x_0$ to $x(F)$ satisfying (7.3.12). In §3.2.3, it was shown that such an approximation can be found with the aid of the same path-following technique, provided that we are given a *starting point* $w \in \operatorname{int} G$. It suffices to introduce the *constant* monotone operator $T(x) \equiv -F'(w)$ (which, of course, is 0-compatible with $F$) and to consider the family

$$T_t(\cdot) = tT(\cdot) + F'(\cdot).$$

The corresponding trajectory $x_*(t)$ is defined by the equation

$$F'(x_*(t)) = -tT(x) \equiv tF'(w)$$

and passes through $w$: $x_*(1) = w$. We can follow it as $t \to 0$ with the aid of the updating rule

$$t^{i+1} = \left(1 - \gamma\left(\frac{\lambda}{3(1+\beta)}, \vartheta, 0\right)\right) t^i, \qquad x^{i+1} = x^i - e(T_{t^{i+1}}, x^i),$$

starting with $(t^0, x^0) = (1, w)$; in view of Proposition 7.3.4, this procedure is well defined and the predicate $\mathsf{P}(T, t, \lambda/(3(1+\beta)); x)$ holds true for $(t, x) = (t^i, x^i)$. We terminate the above process at the moment when the relation $\nu(F', x^i) \leq (1+\beta)^{-1}\lambda/2$ is satisfied; the resulting $x^i$ can be chosen as the above $x_0$. According to Proposition 3.2.3, the number $M$ of steps of the described procedure satisfies the relation

$$M \leq O(1)(1+\beta)\frac{\vartheta^{1/2}}{\lambda} \ln \frac{\vartheta}{\alpha(G:w)}, \tag{7.3.14}$$

where $O(1)$ is an absolute constant and

$$\alpha(G:w) = \max\{t \mid w + t(w - G) \subset G\} \tag{7.3.15}$$

is the asymmetry coefficient of $G$ with respect to $w$.

**Remark 7.3.1** *The quantity $t_0$ defined by* (7.3.13) *admits the lower estimate*

$$t_0 \geq \frac{\lambda}{2(1+\beta)\max\{\langle S(x_0), y - x_0\rangle \mid y \in G\}}.$$

Indeed,

$$\begin{aligned} \| S(x_0) \|^*_{F',x_0} &= \sup\{\langle S(x_0), y - x_0\rangle \mid y \in W_1(F', x_0)\} \\ &\leq \max\{\langle S(x_0), y - x_0\rangle \mid y \in G\}, \end{aligned}$$

since $W_1(F', x_0) \subset G$; see Proposition 7.2.1(ii).

### 7.3.5 Accuracy of approximations

We described a rule that allows us to form $\lambda$-tight approximations to the trajectory $x^*(t)$ along the sequence of values of $t$ increasing in the ratio $(1 + O(1)(\lambda(1+\beta)^{-1}\vartheta^{-1/2}))$. We also know that $x^*(t)$ converges in a proper sense to the set of (weak) solutions to $\mathcal{V}(G, S)$ as $t \to \infty$. We are now interested in what can be said about the convergence of $\lambda$-tight approximations to $x^*(t)$ to the solution set as $t \to \infty$. We give two possible answers to the problem.

**A.** Let us call a monotone single-valued operator $S : G' \to E^*$ *semibounded* if the quantity

$$\| S \|_G = \sup\{\langle S(x), y - x\rangle \mid x \in G',\ y \in G\}$$

is finite.

**Proposition 7.3.5** *Let $S \in \mathcal{C}_\beta(G,F)$ be semibounded and let $t > 0$, $\mu \in [0, 1/(5\kappa)]$. Assume that $x$ satisfies the predicate* $\mathsf{P}(S,t,\mu;x)$. *Then*

$$\varepsilon(S,x) \leq 5\mu \parallel S \parallel_G + \frac{\vartheta}{t}. \tag{7.3.16}$$

**Proof.** Assume that $x$ satisfies $\mathsf{P}(S,t,\mu;x)$. By virtue of Theorem 7.2.1(ii), there exists a solution $x^*$ to the equation $S_t(\cdot) = 0$, such that

$$\parallel x - x^* \parallel_{S_t,x^*} \leq 5\mu. \tag{7.3.17}$$

Let $y \in G'$. We have $\langle S(y), x^* - y\rangle \leq \langle S(x^*), x^* - y\rangle$ (since $S$ is monotone). As in the proof of Proposition 7.3.3, we have $\langle S'(x^*), x^* - y\rangle \leq \vartheta t^{-1}$, so that

$$\langle S(y), x^* - y\rangle \leq \frac{\vartheta}{t}. \tag{7.3.18}$$

Now, since $\parallel x - x^* \parallel_{S_t,x^*} \leq 5\mu \leq 1$ (see (7.3.17)), there exists $u \in W_1(S_t, x^*)$ such that $x = 5\mu u + (1-5\mu)x^*$. By virtue of Proposition 7.2.1(ii), we have $u \in G$; therefore

$$\begin{aligned}\langle S(y), x - y\rangle &= 5\mu \langle S(y), u - y\rangle + (1-5\mu)\langle S(y), x^* - y\rangle \\ &\leq 5\mu \parallel S \parallel_G + (1-5\mu)\frac{\vartheta}{t}\end{aligned}$$

(we used (7.3.18)), which immediately leads to (7.3.16). □

The above statement means that, if $t$ is *large*, if $\mu$ is *small*, and if $x$ is a $\mu$-tight approximation of the point $x^*(t)$, then $x$ is a high-quality approximation to the solution of $\mathcal{V}(G,S)$ with respect to the accuracy measure $\varepsilon(S,\cdot)$. A drawback of this result is that the estimate involves $\mu$ as well as $t$. This is not too dangerous: We can maintain $\lambda$-tightness to the trajectory $x^*(t)$ with a "large" $\lambda$, say, with $\lambda = 1/(25\kappa)$, until $t$ becomes large enough and then apply to $S_t$ the Newton method described in Theorem 7.2.1 to update the $\lambda$-tight approximation into a $\mu$-tight one with small $\mu$. By virtue of Theorem 7.2.1, the latter updating requires $O(\ln\ln \lambda/\mu)$ Newton steps. It is interesting that we have no difficulties with necessity to find too-tight approximations to the central path when using another accuracy measure.

**B.** Let us introduce a new accuracy measure $\varepsilon^*(S,x)$, $x \in G'$, as follows. Let us associate with $x \in G'$ the convex quadratic form

$$S_x(y) = \langle S(x), y - x\rangle + \tfrac{1}{2}\left\langle \hat{S}(x)(y-x), y - x\right\rangle$$

of $y \in E$. The quantity

$$\varepsilon^*(S,x) = S_x(x) - \min\{S_x(y) \mid y \in G\}$$

is nonnegative; under mild restrictions on $S$, it is closely related to the accuracy measure $\varepsilon^+(S,x)$, as demonstrated by the following statement.

**Proposition 7.3.6** *Let $G \subset E$ be bounded and let $S$ be a Lipschitz continuous single-valued monotone operator defined on $G' \equiv \operatorname{int} G$. Let $\|\cdot\|$ be a Euclidean norm on $E$ and let $\|\cdot\|^*$ be the conjugate norm on $E^*$. Let*

$$d_{\|\cdot\|}(G) = \max\{\| x - y \| \mid x, y \in G\}$$

*and*

$$L_{\|\cdot\|}(S) = \sup\left\{\frac{\| S(x) - S(y) \|^*}{\| x - y \|} \mid x \neq y,\ x, y \in G'\right\}.$$

*Then, for all $x \in G'$, we have*

$$\varepsilon^+(S,x) \leq \max\{2\varepsilon^*(S,x); (2\Omega_{\|\cdot\|}(S)\varepsilon^*(S,x))^{1/2}\},$$

*where*

$$\Omega_{\|\cdot\|}(S) = L_{\|\cdot\|}(S) d^2_{\|\cdot\|}(G).$$

**Proof.** By definition, $\varepsilon^+(S,x) = \max\{\langle S(x), x - y\rangle \mid y \in G\}$. For a given $x \in G'$, choose $y \in G$ such that $\varepsilon \equiv \varepsilon^+(S,x) = \langle S(x), x - y\rangle$ and consider the function $f(t) = S_x(x + t(y - x))$, $0 \leq t \leq 1$. This is a quadratic function with the properties

$$f'(0) = \langle S(x), y - x\rangle = -\varepsilon, \qquad f''(0) = \left\langle \hat{S}(x)(y - x), (y - x)\right\rangle \leq \Omega \equiv \Omega_{\|\cdot\|}(S).$$

It follows that

$$\varepsilon^*(S,x) \geq f(0) - \min\{f(t) \mid 0 \leq t \leq 1\} \geq \min\left\{\frac{\varepsilon^2}{2\Omega}; \frac{\varepsilon}{2}\right\}$$

or

$$\varepsilon = \varepsilon^+(S,x) \leq \max\{2\varepsilon^*(S,x); (2\Omega\varepsilon^*(S,x))^{1/2}\}. \qquad \square$$

It turns out that, if $x \in G'$ is a $\lambda$-tight approximation to $x^*(t)$, then $\varepsilon^*(S,x)$ is of order of $t^{-1}$.

**Proposition 7.3.7** *Let $G$ be bounded, let $S \in \mathcal{C}_\beta(G,F)$, and let $x \in G'$, $t > 0$ and $\lambda \in [0, 1/(25\kappa)]$ be such that $\mathsf{P}(S,t,\lambda;x)$ holds. Then*

$$\varepsilon^*(S,x) \leq 2\frac{\vartheta}{t}. \tag{7.3.19}$$

**Proof.** Let $T(y) = S(x) + \hat{S}(x)(y - x)$ be the derivative of $S_x(y)$ with respect to $y$. Consider the mapping

$$T_t(\cdot) = (1+\beta)^2(tT(\cdot) + F'(\cdot)) : G' \to E^*.$$

We clearly have $T \in \mathcal{C}_\beta(G,F)$ and $\| T_t(x) \|^*_{T_t,x} = \| S_t(x) \|^*_{S_t,x}$, so that $\| T_t(x) \|^*_{T_t,x} \leq \lambda$. The latter relation means that the derivative $f'$ of the function

$$f(y) = tS_x(y) + F(y)$$

satisfies the inequality

$$\nu(f', x) \le \frac{\lambda}{1+\beta}. \tag{7.3.20}$$

Since the derivative $T(\cdot)$ of $S_x(\cdot)$ clearly is 0-compatible with $F$, the mapping $f' = tT + F'$ is strongly self-concordant and regular on $G'$ (Proposition 7.3.2). In view of (7.3.20), Theorem 7.2.1, as applied to $f'$, means that there exists $x^* \in G'$ with the properties

$$f'(x^*) = 0, \tag{7.3.21}$$

$$\| x - x^* \|_{f',x^*} \le \frac{5\lambda}{1+\beta} \le \frac{1}{5\kappa}. \tag{7.3.22}$$

We have

$$\begin{aligned} \varepsilon^*(S,x) &= S_x(x) - \min_G S_x(\cdot) \\ &= \{S_x(x) - S_x(x^*)\}_1 + \{S_x(x^*) - \min_G S_x(\cdot)\}_2. \end{aligned} \tag{7.3.23}$$

The derivative $T(x^*)$ of $S_x(\cdot)$ at the point $x^*$ equals to $-t^{-1}F'(x^*)$; since $S_x(\cdot)$ is convex, we have

$$\{\cdot\}_2 \le \max\{\langle T(x^*), x^* - y\rangle \mid y \in G\} = \frac{1}{t}\max\{\langle F'(x^*), y - x^*\rangle \mid y \in G\};$$

as we already mentioned (see the proof of Proposition 7.3.3), the latter quantity is $\le t^{-1}\vartheta$. Thus,

$$S_x(x^*) - \min_G S_x(\cdot) \le \frac{\vartheta}{t}. \tag{7.3.24}$$

It remains to estimate $\{\cdot\}_1$. We have

$$t(S_x(x) - S_x(x^*)) = \{f(x) - f(x^*)\} + \{F(x^*) - F(x)\}.$$

The function $f$ clearly is strongly self-concordant on $G'$, and (7.3.20) means precisely that $\lambda(f,x) \le (1+\beta)^{-1}\lambda < 1/(25\kappa)$; from Theorem 2.2.2(iii), it follows that $f(x) - f(x^*) \le (1+\beta)^{-2}\lambda^2$. We evidently have

$$\| x - x^* \|_{F',x^*} \le \| x - x^* \|_{f',x^*},$$

so that (7.3.22) implies

$$\| x - x^* \|_{F',x^*} \le 5(1+\beta)^{-1}\lambda.$$

Since $F$ is convex and $\| F'(x^*) \|^*_{F',x^*} \le \vartheta^{1/2}$ (the definition of a $\vartheta$-self-concordant barrier), we have

$$F(x^*) - F(x) \le \| F'(x^*) \|^*_{F',x^*} \| x - x^* \|_{F',x^*} \le \frac{5\vartheta^{1/2}\lambda}{1+\beta}.$$

Thus,

$$S_x(x) - S_x(x^*) \le \frac{1}{t}\left\{\frac{\lambda^2}{(1+\beta)^2} + \frac{5\vartheta^{1/2}\lambda}{1+\beta}\right\},$$

which, combined with (7.3.24), leads to (7.3.19). □

### 7.3.6 Summary

Combining the above considerations, we obtain the following method for solving variational equations. Let $G$ be a bounded closed convex set with a nonempty interior in $E$, let $F$ be a $\vartheta$-self-concordant barrier for $G$, and let $S : G \to E^*$ be a Lipschitz continuous single-valued monotone operator $\beta$-compatible with $F$. Assume that we are given a starting point $w$ belonging to $\operatorname{int} G$. Let

$$\lambda = \frac{1}{25\kappa},$$

$\kappa$ being the absolute constant from §7.2 (see (7.2.15)).

To solve $\mathcal{V}(G, S)$, we can use the following two-stage method:

**PRELIMINARY_STAGE** (input: $w$; output: $x_{\text{out}}$)
$(t, x) := (1, w)$;
**Step:**
$t := (1 - (3/2)\lambda(\lambda + 3(1+\beta)\vartheta^{1/2})^{-1})t$;
$x := x - (F''(x))^{-1}(F'(x) - tF'(w))$;
**IF**
$\langle F'(x), [F''(x)]^{-1}F'(x)\rangle^{1/2} \equiv \| F'(x) \|^*_{F',x} > \frac{\lambda}{2(1+\beta)}$
**THEN**
goto Step
**ELSE**
$x_{\text{out}} := x$;
Stop;
**ENDIF**
**MAIN_STAGE** (input: the output $x_{\text{out}}$ of the preliminary stage)
$x := x_{\text{out}}$;
$t := 0.5(1+\beta)^{-1} \langle S(x), [F''(x)]^{-1}S(x)\rangle^{-1/2} \lambda$;
**Step**:
$t := (1 + (3/2)\lambda(\lambda + (1+\beta)\vartheta^{1/2})^{-1})t$;
$x := x - (tS'(x) + F''(x))^{-1}(tS(x) + F'(x))$;
goto **Step**;

**Theorem 7.3.1** *Let $G$ be a bounded closed convex subset of $E$, $w \in \operatorname{int} G$, let $F$ be a $\vartheta$-self-concordant barrier for $G$, and let $S \in \mathcal{C}_\beta(G, F)$ be Lipschitz continuous. Then the above procedure is well defined, and, for each $\varepsilon > 0$, the total amount of steps of the preliminary and the main stages after which all values of $x$ generated at the main stage satisfy the inequality*

$$\varepsilon^*(S, x) \le \varepsilon \tag{7.3.25}$$

*does not exceed the quantity*

$$O(1)(1+\beta)\vartheta^{1/2} \ln(2 + \vartheta\alpha^{-1}(G : w) + \vartheta(1+\beta) \| S \|_G \varepsilon^{-1}),$$

*where $O(1)$ is an absolute constant,*

$$\alpha(G : w) = \max\{t \mid w + t(w - G) \subset G\}$$

*is the asymmetry coefficient of $G$ with respect to $w$, and*

$$\| S \|_G = \sup\{\langle S(x), y - x\rangle \mid x, y \in \operatorname{int} G\}.$$

*Besides this, if $\| \cdot \|$ is an arbitrary Euclidean norm on $E$, then* (7.3.25) *implies that*

$$\varepsilon(S, x) \leq \varepsilon^+(S, x) \leq \max\{2\varepsilon; (2\Omega_{\|\cdot\|}(S)\varepsilon)^{1/2}\},$$

*where*

$$\Omega_{\|\cdot\|}(S) = L_{\|\cdot\|}(S) d^2_{\|\cdot\|}(G),$$

*$L_{\|\cdot\|}(S)$ being the Lipschitz constant of $S : G \to E^*$ with respect to the norm $\| \cdot \|$ on $E$ and the conjugate norm $\| \cdot \|^*$ on $E^*$ and $d_{\|\cdot\|}(G)$ being the diameter of $G$ with respect to $\| \cdot \|$ .*

### 7.3.7 Application example

As an application example for the above approach, consider the following problem (it is a particular case of the pure exchange model of Arrow–Debreu). There are $m$ customers and $n$ kinds of goods. The total amount of the $k$th good that can be distributed among the customers is $x_k^* > 0$. The $i$th customer possesses $w_i$ money units and is characterized by the utility function $f_i(x_1, \dots, x_n)$ depending on the vector of goods bought by the customer. The distribution of goods is organized as follows. The goods are sold at certain prices, which form the vector of prices $p = (p_1, \dots, p_n) > 0$. For a given $p$, the $i$th customer buys the amount of goods $x^{(i)}(p) = (x_1^{(i)}(p), \dots, x_n^{(i)}(p))$, which is the optimal solution to the program

$$\text{maximize } f_i(x) \quad \text{s.t. } x \geq 0,\ p^T x \leq w_i.$$

The problem is to find the equilibrium prices, i.e., to find a positive solution of the equation

$$\sum_{i=1}^{m} x^{(i)}(p) = x^*,$$

or, which is the same, to solve the variational inequality generated by the operator

$$S(p) = x^* - \sum_{i=1}^{m} x^{(i)}(p)$$

on the interior $\operatorname{int} \mathbf{R}_+^n$ of the nonnegative $n$-dimensional orthant.

Let us restrict ourselves to utility functions of the form

$$f_i(x) = \sum_{k=1}^{n} c_{ik} x_k^{a_{ik}},$$

where $c_{ik} > 0$ and $a_{ik} \in (0,1)$.

It is known (see [Po 73], [PM 78]) that, for the above utility function, the mapping $p \to (-x^{(i)}(p)) : \operatorname{int} \mathbf{R}^n_+ \to \mathbf{R}^n_+$ is monotone. As we prove in Appendix 2, it is also compatible with the standard $n$-self-concordant barrier

$$F(p) = -\sum_{k=1}^{n} \ln p_k$$

for $\mathbf{R}^n_+$:

$$-x^{(i)}(\cdot) \in \mathcal{C}_{\beta_i}(\mathbf{R}^n_+, F),$$

where

$$\beta_i = \frac{3(a^{(i)}_{\max}/a^{(i)}_{\min})^{1/3}}{(1 - a^{(i)}_{\max})^{5/2}} \sqrt{2n},$$

$$a^{(i)}_{\max} = \max\{a_{i1}, \dots, a_{in}\}, \qquad a^{(i)}_{\min} = \min\{a_{i1}, \dots, a_{in}\}.$$

It follows that the mapping $S(\cdot)$ is $(\max_i \beta_i)$-compatible with $F$ (Proposition 7.3.1), so that, to find equilibrium prices, we can use the above path-following method. Recall that the method requires $G$ to be bounded, while now it is not the case. This difficulty can be easily avoided. Indeed, the solution $p^*$ to our problem clearly satisfies the inequality $(p^*)^T x^* \le \mu \equiv \sum_{i=1}^m w_i$, so that $p^*$ belongs to the interior of the simplex $G = \{p \ge 0 \mid p^T x^* \le 2\mu\}$. The function $F^+(x) = F(x) - \ln(2\mu - p^T x^*)$ is an $(n+1)$-self-concordant barrier for $G$, and $S$ is $(\max_i \beta_i)$-compatible with this barrier (Proposition 7.3.1(i)). Thus, it suffices to solve the variational inequality generated by $S$ on $G$, and now there is no difficulty with the unboundedness of the domain.

## 7.4 Inequalities with linear operators. Reducibility to linear case

### 7.4.1 Variational inequalities with linear monotone operators

One of the simplest classes of monotone operators is formed by linear operators

$$S(x) = Ax + a,$$

where $A : E \to E^*$ satisfies the relation

$$\langle Ah, h \rangle \ge 0, \qquad h \in E,$$

and $a \in E^*$.

For every closed convex domain $G \in E$, the pair $(G, S|_G)$ is, of course, a monotone element, and solutions to the corresponding variational inequality are exactly the same as weak solutions to it (see Proposition 7.1.1). The operator involved into the inequality is 0-compatible with any self-concordant barrier for $G$, provided that $G$ does not contain straight lines. Thus, in the case of bounded $G$ the only difficulty when solving a variational inequality with a linear monotone operator by a path-following method is to find a self-concordant barrier for $G$.

It turns out that, in fact, a variational inequality with a linear monotone operator can be more or less straightforwardly reduced to a convex programming problem.

To this purpose, consider the intersection $G^+$ of the cone

$$K = \mathrm{cl}\left\{(t,x) \in \mathbf{R} \times E \mid t > 0,\ \frac{x}{t} \in G\right\}$$

(this is the conic hull of $G$) and the affine hyperplane $\{t = 1\}$ in $\mathbf{R} \times E$. Note that $K$ is a closed convex cone with a nonempty interior in $\mathbf{R} \times E$ (this cone is pointed, if $G$ does not contain straight lines). Let

$$S^+(t,x) : \mathbf{R} \times E \to (\mathbf{R} \times E)^*$$

be the linear homogeneous mapping defined as

$$S^+(t,x) = (-\langle a, x\rangle, Ax + ta).$$

We have $\langle S^+(t,x), (t,x)\rangle = \langle Ax, x\rangle \geq 0$, so that the mapping $S^+$ is monotone. If $x^*$ is a solution to $\mathcal{V}(G,S)$, then the point $\bar{x}^* = (1, x^*)$ belongs to $G^+$ and clearly satisfies the relation

$$\langle S^+(\bar{x}^*), z - \bar{x}^*\rangle \geq 0, \qquad z \in G^+.$$

Conversely, if $\bar{x}^* = (1, x^*)$ satisfies the latter relation, then, as it is easily seen, $x^*$ is a solution to $\mathcal{V}(G,S)$.

Thus, we have proved that a variational inequality with linear monotone operator can be straightforwardly reformulated as the following *conic variational inequality*:

$$(\mathcal{P}): \qquad \text{find } x \in K \bigcap (L+b) \quad \text{s.t. } \langle \alpha x, y - x\rangle \geq 0 \quad \text{for all } y \in K \bigcap (L+b),$$

where $K$ is a closed convex cone with a nonempty interior in a finite-dimensional vector space $H$; $L$ is a linear subspace in $H$; $b \in H$; and $\alpha : H \to H^*$ is a linear homogeneous monotone mapping

$$\langle \alpha h, h\rangle \geq 0, \qquad h \in H.$$

Note that our reduction scheme ensures *regularity* of the conic variational inequality, i.e., the property

$$(L+b) \bigcap \mathrm{int}\, K \neq \emptyset.$$

Note also that the above scheme is not the only way to reduce a variational inequality $\mathcal{V}(G, S)$ to a conic variational inequality. In fact, it is easily seen that any conic representation of $G$ induces (at least in the case when $G$ does not contain straight lines) a conic reformulation of $\mathcal{V}(G, S)$.

### 7.4.2 Reduction to convex program

Now let us demonstrate that a *regular* conic variational inequality $(\mathcal{P})$ can be easily reduced to an equivalent conic optimization problem, namely, to the problem

$$(\mathcal{P}^*): \quad \begin{array}{l} \text{minimize } f(x, s) = \langle Ax, x\rangle - \langle Ax - s, b\rangle \\ \text{s.t. } (x, s) \in (K \times K^*) \bigcap \{\mathcal{L} + \mathsf{b}\}, \end{array}$$

where $K^*$, as always, is the cone dual to $K$ and $\mathcal{L}$, $\mathsf{b}$ are the linear subspace and a vector in $H \times H^*$ defined by the relation

$$\mathcal{L} + \mathsf{b} = \{(x, s) \mid x \in L + b,\ Ax - s \in L^{\perp}\},$$

where $L^{\perp}$ is the annulator of $L$.

Note that the objective in $(\mathcal{P}^*)$ is a convex quadratic form.

**Proposition 7.4.1** *Let $(\mathcal{P})$ be regular; i.e., let $L+b$ intersect* $\operatorname{int} K$. *Then $(\mathcal{P})$ and $(\mathcal{P}^*)$ are equivalent in the sense that, if $(\mathcal{P})$ is solvable, then $(\mathcal{P}^*)$ also is solvable, the optimal value in $(\mathcal{P}^*)$ is* 0, *and the solutions to $(\mathcal{P})$ are precisely the $x$-components of the solutions to $(\mathcal{P}^*)$. Conversely, if $(\mathcal{P}^*)$ is solvable and the optimal value in this problem equals* 0, *then $(\mathcal{P})$ is solvable.*

**Proof**. Assume that $(\mathcal{P})$ is solvable and $x^*$ is a solution to $(\mathcal{P})$. Let

$$Q = \operatorname{cl}\{h \mid x^* + th \in K \text{ for all small enough } t > 0\}$$

be the tangent cone of $K$ at $x^*$. Then the functional $Ax^*$ is nonnegative at the intersection of the closed convex cone $Q$ and the linear subspace $L$, and the latter intersection contains points interior to $Q$ (recall that $(\mathcal{P})$ is regular). From the Dubovitskii–Milutin lemma, it follows that there exists a representation $Ax^* = s^* + y^*$ with $s^*$ being nonnegative on $Q$ and $y^*$ being nonnegative on $L$. The latter means that $y^* \in L^{\perp}$, while $s^* \in K^*$ and $\langle s^*, x^*\rangle \le 0$ (since $-x^* \in Q$). In view of $x^* \in K$, relation $\langle s^*, x^*\rangle \le 0$ implies that $\langle s^*, x^*\rangle = 0$. Thus, the pair $(x^*, s^*)$ is feasible for $(\mathcal{P}^*)$ and $f(x^*, s^*) = \langle Ax^*, x^*\rangle - \langle Ax^* - s^*, b\rangle$. Since $Ax^* - s^* = y^* \in L^{\perp}$ and $b - x^* \in L$, we have

$$\langle Ax^*, x^*\rangle - \langle Ax^* - s^*, b\rangle = \langle Ax^*, x^*\rangle - \langle Ax^* - s^*, x^*\rangle = \langle s^*, x^*\rangle = 0.$$

Thus, $(x^*, s^*)$ is feasible for $(\mathcal{P}^*)$ and $f(x^*, s^*) = 0$.

On the other hand, if $(x, s)$ is a feasible solution to $(\mathcal{P}^*)$, then

$$f(x,s) = \langle Ax, x\rangle - \langle Ax - s, b\rangle = \langle Ax, x\rangle - \langle Ax - s, x\rangle$$

(we considered that $Ax - s \in L^\perp$ and $x - b \in L$). It follows that $f(x, s) = \langle s, x\rangle \geq 0$ (recall that $(x, s) \in K \times K^*$). We see that the objective in $(\mathcal{P}^*)$ is nonnegative at the feasible set of the problem, so that the above $(x^*, s^*)$ is an optimal solution to $(\mathcal{P}^*)$, and the optimal value in $(\mathcal{P}^*)$ is 0.

We have proved that, if $(\mathcal{P}^*)$ is solvable, then $(\mathcal{P}^*)$ is solvable with zero optimal value, and each optimal solution to $(\mathcal{P}^*)$ can be represented as $x$-component of an optimal solution to $(\mathcal{P}^*)$. To complete the proof, it suffices to verify that, if $(\mathcal{P}^*)$ is solvable with zero optimal value and $(x^*, s^*)$ is an optimal solution to $(\mathcal{P}^*)$, then $x^*$ is a solution to $(\mathcal{P})$. Indeed, we have $0 = \langle Ax^*, x^*\rangle - \langle Ax^* - s^*, b\rangle = \langle Ax^*, x^*\rangle - \langle Ax^* - s^*, x^*\rangle$ (since $(x^*, s^*)$ is feasible, $Ax^* - s^* \in L^\perp$ is orthogonal to the vector $x^* - b \in L$), so that $\langle s^*, x^*\rangle = 0$. Let $y \in K \bigcap \{L + b\}$. Then $\langle Ax^*, y - x^*\rangle = \langle s^*, y - x^*\rangle$ (since $Ax^* - s^* \in L^\perp$, $y - x^* \in L$), so that $\langle Ax^*, y - x^*\rangle = \langle s^*, y - x^*\rangle = \langle s^*, y\rangle \geq 0$ (we considered that $\langle s^*, x^*\rangle = 0$ and $\langle s^*, y\rangle \geq 0$ in view of $s^* \in K^*$, $y \in K$). Thus, $x^*$ is a solution to $(\mathcal{P})$. $\square$

We have reduced a variational inequality with a linear monotone operator to $(\mathcal{P}^*)$; the latter problem is a conic optimization problem, and therefore it can be solved by polynomial-time interior-point methods, in particular, by potential reduction ones. To apply these methods, it suffices to know a $\vartheta$-logarithmically homogeneous self-concordant barrier $F$ for the cone $K$ and its Legendre transformation $F^*$; if it is the case, we can take the function

$$\Phi(x, s) = F(x) + F^*(-s)$$

as a $2\vartheta$-logarithmically homogeneous self-concordant barrier for the cone $K \times K^*$ involved into $(\mathcal{P}^*)$. Note that the Legendre transformation of this barrier is $\Phi(-x, -s)$, so that $F$ and $F^*$ induce both the primal and the dual barrier for $(\mathcal{P}^*)$.

### 7.4.3 Convex representation of a monotone operator

Recall that we can now apply the interior-point machinery to a rather restricted class of variational inequalities with monotone operators, since the operator involved should be compatible with a barrier for the corresponding domain $G$. The same difficulties occurred in convex optimization: To apply a barrier-generated path-following method to a convex optimization problem *directly*, we should require compatibility of the objective and the barrier for the feasible domain, which is a severe restriction on the objective. In the optimization case, however, it was easy to overcome this difficulty: We always could reduce the problem to an equivalent one with a linear objective, so that, in fact, the possibility of solving convex programming problems was limited only by our abilities to point out "computable" self-concordant barriers for convex domains. In a sense, the situation with monotone variational inequalities is similar: As we see, such an inequality admits a "convex representation," i.e., can be reduced to an inequality with a linear operator on a modified domain. Below, we describe the corresponding reduction and develop a kind of calculus for convex representations (cf. §§5.1, 5.2).

**Definition 7.4.1** *A convex representation (c.r.) for a monotone element $(G, S)$ is, by definition, a closed convex set $G^+$ contained in the set*

$$G^\# = G \times E^* \times \mathbf{R} \subset E^\# = E \times E^* \times \mathbf{R}$$

*such that the following two conditions hold:*
(a) $(x, \xi) \in \mathsf{G}(S) \;\Rightarrow\; (x, \xi, \langle \xi, x\rangle) \in G^+$;
(b) $(x, \xi, s) \in G^+ \;\Rightarrow\; s - \langle \xi, y\rangle \geq \langle \eta, x - y\rangle$ *for all* $(y, \eta) \in \mathsf{G}(S)$.

Let us establish some relations between the introduced notions.

**Proposition 7.4.2** *Let $(G, S)$ be a monotone element.* (i) *Convex representations for $(G, S)$ do exist; the minimal (with respect to inclusion) among them is*

$$G^*(G, S) = \mathrm{cl}\,\{\mathrm{conv}\,\{(x, \xi, \langle \xi, x\rangle) \mid (x, \xi) \in \mathsf{G}(S)\}\}.$$

(ii) *Let $G^+$ be a convex representation of $(G, S)$ and let $\mathcal{V}(G, S)$ admit a solution. Consider the bilinear game with the cost function*

$$\gamma(w, y) = \{s - \langle \xi, y\rangle\},\ w = (x, \xi, s) \in G^+, \qquad y \in G$$

*(the player choosing $w$ tries to minimize, and that one choosing $y$ tries to maximize the function) and let*

$$\bar{\gamma}(w) = \sup_{y \in G} \gamma(w, y)$$

*be the cost function of the first player. Then the saddle set of $\gamma$ is nonempty, and, consequently, the set of solutions to the problem of the first player*

$$\text{minimize } \bar{\gamma}(w) \quad \text{s.t. } w \in G^+$$

*also is nonempty; the natural projection of the latter set onto $E$ contains all solutions to $\mathcal{V}(G, S)$ and is contained in the set of all weak solutions to $\mathcal{V}(G, S)$. In particular, in the case when the sets of solutions and weak solutions to $\mathcal{V}(G, S)$ coincide and are nonempty, the optimal set of the first player coincides with the set of solutions to the variational inequality.*

**Proof.** (i) It is clear that the set $G^*(G, S)$ is contained in any convex representation for $(G, S)$, so that it suffices to prove that this set is a convex representation of the element. It evidently contains all the triples

$$(x, \xi, s) \equiv (x, \xi, \langle \xi, x\rangle), \quad x \in \mathrm{Dom}\{S\}, \quad \xi \in S(x),$$

and, for such a triple, we have

$$s - \langle \xi, y \rangle = \langle \xi, x - y \rangle \geq \langle \eta, x - y \rangle$$

for all $y \in \mathrm{Dom}\{S\}$ and all $\eta \in S(y)$ (recall that $S$ is monotone). Thus, the inequality $s - \langle \xi, y \rangle \geq \langle \eta, x - y \rangle$ holds for all of the above triples and all $y \in \mathrm{Dom}\{S\}$, $\eta \in S(y)$. Since both sides of the inequality are linear in $(x, \xi, s)$, it also holds for all triples from the set

$$\mathrm{conv}\,\{(x, \xi, \langle \xi, x \rangle) \mid x \in \mathrm{Dom}\{S\},\ \xi \in S(x)\},$$

and consequently for all triples from $G^*(G, S)$. Thus, $G^*(G, S)$ is a conic representation for $(G, S)$.

(ii) Let

$$\mu(x; y, \eta) = \langle \eta, x - y \rangle\,, x \in G, (y, \eta) \in \mathsf{G}(S) \equiv \{(y, \eta) \mid y \in \mathrm{Dom}\{S\},\ \eta \in S(y)\}.$$

By definition of a convex representation, we have $\gamma(w, y) \geq \mu(x; y, \eta)$ for all $w = (x, \xi, s) \in G^+$ and all $(y, \eta) \in \mathsf{G}(S)$, so that

$$\bar{\gamma}(w) \geq \sup_{(y,\eta) \in \mathsf{G}(S)} \mu(x; y, \eta) \equiv a(x).$$

Let us prove that $a(x) \geq 0$. Indeed, let $y \in \mathrm{int}\, G$; then the points $y_t = x + t(y - x)$ belong to $\mathrm{int}\, G \subseteq \mathrm{Dom}\{S\}$ for all $t \in (0, 1]$, so that there exist $\eta_t \in S(y_t)$. From the monotonicity of $S$, it follows that the function

$$\sigma(t) = \langle \eta_t, y - x \rangle$$

is nondecreasing on $(0, 1]$, and, of course,

$$a(x) \geq \sup_{0 < t \leq 1} \langle \eta_t, x - y_t \rangle = \sup_{0 < t \leq 1} \{-t\sigma(t)\} \geq 0$$

(we considered that $\sigma(\cdot)$ is nondecreasing). Thus, $\bar{\gamma}(w) \geq 0$, $w \in G^+$.

Now let $x^*$ be a solution to $\mathcal{V}(G, S)$. Then there exists $\xi^* \in S(x^*)$ such that

$$\langle \xi^*, y - x^* \rangle \geq 0, \qquad y \in G.$$

We have $w^* = (x^*, \xi^*, \langle \xi^*, x^* \rangle) \in G^+$ (the definition of a convex representation) and

$$\bar{\gamma}(w^*) = \sup_{y \in G} \{\langle \xi^*, x^* \rangle - \langle \xi^*, y \rangle\} \leq 0.$$

Thus, $\bar{\gamma}(w^*) = 0$, and $w^*$ is the minimizer of $\bar{\gamma}$ over $G^+$. Moreover, $(w^*, x^*)$ is a saddle point of $\gamma$, since, for $w = (x, \xi, s) \in G^+$ and $y \in G$, we have

$$\gamma(w, x^*) = s - \langle \xi, x^* \rangle \geq \langle \xi^*, x^* - x^* \rangle = 0 \geq \langle \xi^*, x^* - y \rangle = \gamma(w^*, y)$$

(the first inequality follows from $w \in G^+$ and Definition 7.4.1 (b)).

We have proved that the solutions to $\mathcal{V}(G, S)$ are contained in the projection onto $E$ of the optimal set of $\bar{\gamma}$. In turn, let $w^* = (x^*, \xi^* s^*)$ be a minimizer of $\bar{\gamma}$ or, which is the same, let $\bar{\gamma}(x^*, \xi^*, s^*) = 0$. It follows that

$$\gamma(x^*, \xi^*, s^*; y) = s^* - \langle \xi^*, y \rangle \leq 0, \qquad y \in G,$$

and also, since

$$s^* - \langle \xi^*, y \rangle \geq \langle \eta, x^* - y \rangle \quad \text{for all } (y, \eta) \in \mathsf{G}(S)$$

(due to $(x^*, \xi^*, s^*) \in G^+$), we conclude that $\langle \eta, x^* - y \rangle \leq 0$, so that $x^*$ is a weak solution to $\mathcal{V}(G, S)$. □

The advantage of a convex representation for a monotone element $(G, S)$ is that, given such a convex representation $G^+$, we can reduce the problem of solving variational inequality $\mathcal{V}(G, S)$ associated with the element to a bilinear game with the cost function

$$\gamma(w; y) = s - \langle \xi, y \rangle, \quad w = (x, \xi, s) \in G^+, \quad y \in G$$

(see Proposition 7.4.2(ii)). Assume that $\mathcal{V}(G, S)$ has a solution; then, as we have already proved, $\gamma$ possesses a nonempty saddle set on $G^+ \times G$, and the $x$-component of the projection onto $E$ of any saddle point of $\gamma$ is a weak solution to $\mathcal{V}(G, S)$. Assume that we would be satisfied by any weak solution to $\mathcal{V}(G, S)$; then it suffices to find a saddle point of $\gamma$ (and we know that such a point exists). Since $\gamma$ is a bilinear function, finding a saddle point of it is the same as solving a variational inequality with an *affine* monotone operator

$$S(w, y) : \langle S(w, y), (\delta w, \delta y) \rangle = \langle -\delta\xi, y \rangle + \langle \xi, \delta y \rangle + \delta s,$$

where $w = (x, \xi, s) \in E \times E^* \times \mathbf{R}$, $y \in E$ and $\delta w = (\delta x, \delta\xi, \delta s) \in E \times E^* \times \mathbf{R}$, $\delta y \in E$. We know how to solve the latter problem by path-following and potential reduction methods, provided that we know coverings (respectively, conic representations) of $G$ and $G^+$.

Thus, it is important for us to develop a kind of calculus that allows us to point out "explicit" convex representations of monotone elements. Let us start with "raw materials" for this calculus.

The following statement is a straightforward consequence of Proposition 7.4.2(i).

**Proposition 7.4.3** *Let $G = E$ and let $S(x) = \alpha x + b$ be an affine monotone operator. Then the set*

$$G^+ = \{(x, \xi, s) \mid x \in E,\ \xi = \alpha x + b,\ s \geq \langle \alpha x, x \rangle + \langle b, x \rangle\}$$

*is a convex representation of $(G, S)$.*

Note that $G^+$ is a closed convex domain in the affine space

$$E^+ = \{(x, \xi, s) \mid s \in \mathbf{R},\ \xi = \alpha x + b\},$$

and, in fact, it is the epigraph of a convex quadratic form, so that there is no problem with finding coverings and conic representations for $G^+$.

Our next statement demonstrates how to find a convex representation for a *potential* monotone element. Let $(G, f)$ be a functional element on $E$; then the operator

$$S_f(x) = \partial f(x)$$

(see §7.1.2, item 1) with the domain comprised of all those points $x$ of $G$ where $f$ is finite and the set $\partial f(x)$ of the support (on $G$) functionals to $f$ is nonempty, is monotone on its domain, and the latter contains at least the interior of $G$; thus, $(G, S_f)$ is a monotone element. Consider the Legendre transformation $(G^*, f^*)$ of the functional element $(G, f)$. Recall that the latter pair is defined as follows. $G^*$ is the closure of all those $\xi \in E^*$ for which the function $\langle \xi, x\rangle - f(x)$ is above bounded in $x \in G$, and

$$f^*(\xi) = \sup\{\langle \xi, x\rangle - f(x) \mid x \in G\}, \qquad \xi \in G^*.$$

**Proposition 7.4.4** *The set*

$$G^+ = \{(x, \xi, s) \in E \times E^* \times \mathbf{R} : x \in \mathrm{Dom}\{f\},\ \xi \in \mathrm{Dom}\{f^*\},\ s \geq f(x) + f^*(\xi)\}$$

*is a convex representation for the monotone element* $(G, S_f)$.

**Proof.** First, as we know (see §5.2.3), $(G^*, f^*)$ is a functional element. It follows that the pairs $(G \times G^*, \phi(x, \xi) = f(x))$ and $(G \times G^*, \psi(x, \xi) = f^*(\xi))$ both are functional elements, so that their sum also is a functional element. Since $G^+$ is precisely the epigraph of the latter element, $G^+$ is closed and convex.

Now let us prove that $G^+$ contains each triple $(x, \xi, s)$ with $x \in \mathrm{Dom}\{S_f\}$, $\xi \in S_f(x)$ and $s = \langle \xi, x\rangle$. Indeed, for the above triple, we clearly have $x \in \mathrm{Dom}\{f\}$, $\xi \in \mathrm{Dom}\{f^*\}$ and $s = f^*(\xi) + f(x)$, so that $(x, \xi, s) \in G^+$. It remains to verify that $(x, \xi, s) \in G^+$ implies that

$$s - \langle \xi, y\rangle \geq \langle \eta, x - y\rangle, \qquad (y, \eta) \in \mathsf{G}(S_f).$$

Indeed, we have $x \in \mathrm{Dom}\{f\}$, $\xi \in \mathrm{Dom}\{f^*\}$ and $s \geq f(x) + f^*(\xi)$, so that

$$\begin{aligned} s - \langle \xi, y\rangle &\geq f(x) + f^*(\xi) - \langle \xi, y\rangle \geq f(x) - f(y) \\ &\geq \{f(y) + \langle \eta, x - y\rangle\} - f(y) \geq \langle \eta, x - y\rangle. \qquad \square \end{aligned}$$

Note that a covering for the f.e. $(G, f)$ induces in "almost explicit" manner a covering for its Legendre transformation $(G^*, f^*)$ (§5.2.3), and a simple combination of these coverings (see Corollary 5.2.3) is a covering for $G^+$.

Now let us demonstrate how to obtain a convex representation of "a combined" monotone element on the basis of convex representations of the elements involved into the combination. There are five standard operations preserving monotonicity: multiplying by a positive constant factor, restriction onto a smaller domain, affine substitution of argument, summation, and taking the inverse. The induced transformations of convex representations are given by the statements that follow.

Several transformation rules are evident.

**Proposition 7.4.5** *Let $(G, S)$ be a monotone element and let $G^+$ be its convex representation.*

(i) *Let $\alpha > 0$. Then the set*

$$G^{\#} = \{(x, \xi, s) \mid (x, \alpha^{-1}\xi, \alpha^{-1}s) \in G^+\}$$

*is a convex representation of the element $(G, \alpha S)$.*

*(Note that $G^{\#}$ is the image of $G^+$ under an invertible affine mapping, so that a covering (a conic representation) of $G^+$ naturally induces a covering (respectively, a conic representation) of $G^{\#}$.)*

(ii) *Let $H$ be a closed convex domain contained in $G$ and let $S\mid_H$ be the multivalued mapping defined by the relations*

$$\mathrm{Dom}\{S\mid_H\} = H \bigcap \mathrm{Dom}\{S\}, \ (S\mid_H)(x) = S(x), \ x \in \mathrm{Dom}\{S\mid_H\}.$$

*Then $(H, S\mid_H)$ is a monotone element, and the set*

$$H^+ = \{(x, \xi, s) \in G^+ \mid x \in H\}$$

*is a convex representation of this element.*

*(Note that $H^+$ is the intersection of $G^+$ and the direct product of $H$ and a linear space; therefore a pair of coverings (a pair of conic representations) of $G^+$ and $H$ induces a covering (respectively, a conic representation) of $H^+$.)*

(iii) *Let $\mathcal{B}(y) = By + b$ be an affine mapping from a finite-dimensional real vector space $E'$ into $E$ such that* Im $\mathcal{B}$ *intersects* int $G$. *Set $G' = \mathcal{B}^{-1}(G)$ and let $S'$ be the operator defined by the relations*

$$\mathrm{Dom}\{S'\} = \mathcal{B}^{-1}(\mathrm{Dom}\{S\}), \quad S'(y) = \{B^*\xi \mid \xi \in S(\mathcal{B}(y))\}, \quad y \in \mathrm{Dom}\{S'\},$$

*where $B^*$ is the operator conjugate to $B$. Then $(G', S')$ is a monotone element, and the set*

$$G' = \mathrm{cl}\,\{(y, \eta, s) \mid \exists (x, \xi, s) \in G^+ : \ x = \mathcal{B}(y), \ \eta = B^*\xi\}$$

*is a convex representation of $(G', S')$.*

*(Note that $G' = \mathrm{cl}\ \beta(G'')$, $\beta(x, \xi, s) = (x, B^*\xi, s)$, where $G'' = (\bar{\mathcal{B}}^{-1}(G^+)$ and $\bar{\mathcal{B}}$ is the affine mapping $(x, \xi, s) \to (\mathcal{B}(x), \xi, s)$ (which image clearly intersects the relative interior of $G^+$). Therefore a covering (a conic representation) of $G^+$ induces a covering (respectively, a conic representation) of $G''$. In turn,* in the case when $\beta(G'')$ is closed, *the latter covering/conic representation induces a covering/conic representation of $G'$.)*

Now consider summation.

**Proposition 7.4.6** *Let $(G, S_1)$ and $(G, S_2)$ be monotone elements and let the operator $S_1 + S_2$ be defined by the relations*

$$\mathrm{Dom}\{S_1 + S_2\} = \mathrm{Dom}\{S_1\} \bigcap \mathrm{Dom}\{S_2\},$$

$$(S_1 + S_2)(x) = \{\xi + \eta \mid \xi \in S_1(x),\ \eta \in S_2(x)\},\ x \in \mathrm{Dom}\{S_1 + S_2\}.$$

*Also, let $G_1^+$ and $G_2^+$ be convex representations of $(G, S_1)$ and $(G, S_2)$, respectively. Then the set*

$$G^+ = \mathrm{cl}\,\{(x, \xi + \eta, s + t) \mid (x, \xi, s) \in G_1^+, (x, \eta, t)\} \in G_2^+\}$$

*is a convex representation for $(G, S_1 + S_2)$.*

The proof is quite straightforward.

Note that $G^+ = \mathrm{cl}\,\{\beta(G')\}$, where $G' = \mathcal{B}^{-1}(G_1^+ \times G_2^+)$ and the linear mappings $\beta$ and $\mathcal{B}$ are defined by the relations

$$\beta(x, \xi, s, \eta, t) = (x, \xi + \eta, s + t),$$

$$\mathcal{B}(x, \xi, s, \eta, t) = ((x, \xi, s), (x, \eta, t)).$$

The image of $\mathcal{B}$ clearly intersects the relative interior of $G_1^+ \times G_2^+$, so that a pair of coverings (conic representations) of $G_1^+$, $G_2^+$ induces a covering (respectively, a conic representation) of $G'$. This covering/conic representation, in turn, induces a covering/conic representation of $G^+$, *provided that $\beta(G')$ is closed* (and therefore coincides with $G^+$). The latter requirement is satisfied if at least one of the sets $G_1^+$, $G_2^+$ is bounded.

It remains to consider inversion. Let $(G, S)$ be a monotone element. The projection $\mathrm{Im}\ S$ of the graph $\mathsf{G}(G, S) \subseteq E \times E^*$ onto $E^*$ is formed precisely by those $\xi$ for which the inclusion $\xi \in S(x)$ is solvable; let $S^{-1}(\xi)$ be set of solutions to the inclusion. Thus, we have defined a multivalued mapping $S^{-1}$ that maps a point $\xi$ from $\mathrm{Dom}\{S^{-1}\} \equiv \mathrm{Im}\ S$ into a nonempty subset of $E = (E^*)^*$. It is easily seen that this mapping is monotone, namely,

$$\langle \xi - \eta, x - y \rangle \geq 0, \quad \xi, \eta \in \mathrm{Im}\ S, \quad x \in S^{-1}(\xi), \quad y \in S^{-1}(\eta).$$

If $S$ is defined unproperly, then the above mapping $S^{-1}$ can be bad, e.g., it can be defined in a neighbourhood of a point, but not at the point itself. Let us call a monotone element *normal* if the set $G' = \mathrm{cl}\ \mathrm{Im}\ S$ is convex and possesses a nonempty interior, and the latter interior is contained in $\mathrm{Im}\ S$. In this case, $(G', S^{-1})$ clearly is a monotone element on $E^*$; we call it the element inverse to $(G, S)$. It is well known that, if $S$ is maximal monotone and the affine span of $\mathrm{Im}\ S$ coincides with $E^*$, then $(G, S)$ is normal.

The following statement demonstrates that a convex representation of a normal monotone element is, in fact, a convex representation of the inverse element.

**Proposition 7.4.7** *Let $(G,S)$ be a normal monotone element, $(G',S^{-1})$ be its inverse and let $G^+$ be a convex representation for $(G,S)$. Let*

$$H = \{(\xi, x, s) \in E^* \times (E^*)^* \times \mathbf{R} \equiv E^* \times E \times \mathbf{R} \mid (x, \xi, s) \in G^+\},$$

$$R = \{(\xi, x, s) \mid \xi \in G'\}.$$

*Then $G^{++} \equiv H \bigcap R$ is a convex representation of $(G', S^{-1})$. If $G^+ = G^*(G,S)$ is the minimal convex representation of $(G,S)$, then $G^{++} = H$.*

**Proof**. Let us prove that $G^{++}$ is a convex representation of $(G', S^{-1})$. We should verify that

(a) $G^{++}$ is closed convex set contained in $G' \times (E^*)^* \times \mathbf{R}$;

(b) $G^{++}$ contains all triples $(\xi, x, \langle\xi, x\rangle)$ associated with $(\xi, x) \in \mathsf{G}(S^{-1})$;

(c) For every $(\xi, x, s) \in G^{++}$, we have

$$s - \langle\eta, x\rangle \geq \langle\xi - \eta, y\rangle$$

for all $(\eta, y) \in \mathsf{G}(S^{-1})$. The verification is as follows:

(a) This part is evident;

(b) If $(\xi, x) \in \mathsf{G}(S^{-1})$, then $(x, \xi) \in \mathsf{G}(S)$ and $\xi \in \operatorname{Im} S \subseteq G'$. Since $(x, \xi) \in \mathsf{G}(S)$, we have $(x, \xi, \langle\xi, x\rangle) \in G^+$ and therefore $(\xi, x, \langle\xi, x\rangle) \in H$; since $\xi \in G'$, we also have $(\xi, x, \langle\xi, x\rangle) \in R$, so that $(\xi, x, \langle\xi, x\rangle) \in G^{++}$, and (b) is proved;

(c) If $(\xi, x, s) \in G^{++}$ and $(\eta, y) \in \mathsf{G}(S^{-1})$, then $(x, \xi, s) \in G^+$ and $(y, \eta) \in \mathsf{G}(S)$; these inclusions, combined with the fact that $G^+$ is a convex representation of $(G,S)$, imply that

$$s - \langle\xi, y\rangle \geq \langle\eta, x - y\rangle,$$

which is equivalent to the inequality required by (c).

It remains to prove that, if $G^+$ is the minimal convex representation of $(G,S)$, then $G^{++} = H$ or, which is the same, to verify that $H \subseteq R$. The minimal convex representation is, by definition, the closure of the convex hull of the set $\{(x, \xi, \langle\xi, x\rangle) \mid (x, \xi) \in \mathsf{G}(S)\}$, so that in the case under consideration

$$H = \operatorname{cl}\{\operatorname{conv}\{(\xi, x, \langle\xi, x\rangle) \mid (\xi, x) \in \mathsf{G}(S^{-1})\}\},$$

and, of course, the projection of this set onto $E^*$ is contained in $G'$ (the definition of the inverse to a normal monotone element). $\square$

*This page intentionally left blank*

# Chapter 8

# Acceleration for linear and linearly constrained quadratic problems

## 8.1 Introduction and preliminary results

### 8.1.1 Motivation

In this chapter, we consider the problem

$$(\mathcal{LQ}): \qquad \begin{array}{l} \text{minimize } \psi(x) \equiv \frac{1}{2}x^T Ax - a^T x \\ \text{s.t. } x \in \mathbf{R}^n, \quad f_i(x) \equiv -a_i^T x + b_i \geq 0, \quad 1 \leq i \leq m, \end{array}$$

where $A$ is a positive semidefinite symmetric $n \times n$ matrix, $a, a_1, \ldots, a_m \in \mathbf{R}^n$, $b_1, \ldots, b_m \in \mathbf{R}$. In other words, we deal with a linearly constrained convex quadratic programming problem.

Henceforth, set

$$G = \{x \in \mathbf{R}^n \mid f_i(x) \geq 0,\ 1 \leq i \leq m\}. \tag{8.1.1}$$

We assume that $G$ is a bounded set with a nonempty interior (hence $m > n$). Without loss of generality, we suppose that all $a_i$ are nonzeros. Then

$$G' \equiv \operatorname{int} G = \{x \in \mathbf{R}^n \mid f_i(x) > 0,\ 1 \leq i \leq m\}. \tag{8.1.2}$$

To solve the problem, we can use, say, the barrier-generated path-following method (see §3.2; the method is referred to as the *basic* one) associated with the standard logarithmic barrier

$$F(x) = -\sum_{i=1}^{m} \ln(b_i - a_i^T x) : G' \to \mathbf{R}, \tag{8.1.3}$$

which is an $m$-self-concordant barrier for $G$; of course, the objective is 0-compatible with the barrier. As we know, to improve the accuracy of current approximate solution by an absolute constant factor, the method requires $O(1)m^{1/2}$ Newton-type steps, and each of these steps costs $O(1)mn^2$ arithmetic operations. It follows that the arithmetic cost of finding an $\varepsilon$-solution to the problem, i.e., a point $x_\varepsilon$ satisfying the relation

$$x_\varepsilon \in G', \quad \psi(x_\varepsilon) - \min_G \psi \leq \varepsilon\{\max_G \psi - \min_G \psi\} \tag{8.1.4}$$

is

$$M^*(\varepsilon) = O(1)m^{1.5}n^2 \ln\left(\frac{m}{\alpha(G:w)\varepsilon}\right) \tag{8.1.5}$$

operations, where $\alpha(G:w)$ is the asymmetry coefficient of $G$ with respect to the starting point $w \in G'$ used by the method.

In what follows, $M^*(\varepsilon)$ always denotes the arithmetic cost (corresponding to the method under consideration) of finding an $\varepsilon$-solution to $(\mathcal{LQ})$.

It is well known that the cost given by (8.1.5) can be reduced. To this purpose, we can use different strategies and tactics.

### Strategy

In a path-following method, we are interested in approximating the path

$$x^*(t) = \operatorname{argmin}\{F_t(x) = t\phi(x) + F(x) \mid x \in G'\}$$

($\phi$ is either a linear form (the preliminary stage), or the objective involved into $(\mathcal{LQ})$ (the main stage)). This path should be approximated along a sequence of values of $t$ tending to 0 (at the preliminary stage) or to $\infty$ (at the main stage), and the computational effort is dominated by the arithmetic cost at which we can solve the *basic subproblem*, i.e., to update a given "good" approximation to $x^*(t)$ into an approximation of the same quality to $x^*(2t)$ (or $x^*(t/2)$); of course, the above constant 2 could be replaced by an arbitrary absolute constant greater than 1. To simplify our explanation, restrict ourselves to the basic subproblem associated with the main stage, where $t$ should be increased. A general-type path-following method solves this subproblem by a number of updatings of the type $t \to t^+ = \kappa t$, where $\kappa > 1$ is the rate of increasing $t$. Each step $t \mapsto t^+$ is accompanied by an updating $x \mapsto x^+$, and the only requirement on $x^+$ is that this point should be good (in a proper sense) approximate minimizer of $F_{t^+}(\cdot)$ over $G'$. The most natural way to form $x^+$ is to minimize $F_{t^+}$ by certain standard procedure for smooth convex minimization. Let us restrict ourselves for a moment to one of the standard first-order procedures, e.g., the gradient descent method. An obstacle for fast convergence of this method is that, although our objective $F_{t^+}$ is very smooth, it may happen to be very ill conditioned with respect to our initial (standard) Euclidean structure on the space of variables, and the gradient descent method implemented in a straightforward way may be very time-consuming. To overcome this difficulty, it is reasonable to precede the gradient descent by an appropriate scaling of variables, which ensures a reasonable condition number of the Hessian of $F_{t^+}$ (at least in a neighbourhood of the point $x$ where the procedure starts). This scaling is equivalent to introducing a new Euclidean structure on $\mathbf{R}^n$ with the aid of a scalar product $u^TPv$, where $P$ should be "compatible" with the Hessian of $F_{t^+}$ at $x$; to apply the associated "prescaled" gradient descent method, we should know the matrix $P^{-1}$. Note that the simplest way to implement this scheme leads precisely to the basic method, where we, in fact, perform a single step of the prescaled gradient descent method, the scaling being defined by the Hessian of $F_{t^+}$ at $x$ (indeed, the step of prescaled (in this way) gradient descent method is, up to the stepsize, the Newton step). The rate of varying $t$ in the basic method is chosen according to the requirement that the prescaled gradient descent method could minimize $F_{t^+}$ to the desired accuracy in a single step.

Now it is clear that the strategy underlying the basic method is an "extreme point" of a whole family of strategies. There is neither a priori reason to restrict ourselves to the prescaled gradient descent method (there exist more efficient first-order procedures for smooth convex minimization), nor any necessity to update $x$ by a single step of the procedure. In other words, we could try to vary $t$ at a higher rate than that one used in the basic method and accompany every step $t \mapsto t^+$ by a number of steps of a prescaled first-order procedure for minimizing $F_{t^+}(\cdot)$.

## Tactics

As we have already seen, the above strategies require appropriate scaling of the space of variables: Each step $t \mapsto t^+$ should be accompanied by computing a matrix "compatible" with the inverse Hessian of $F_{t^+}$ at $x$. In the basic method, this is done quite straightforwardly: We simply compute the Hessian and its inverse. Note that this is, in a sense, superfluous: All we need is not the exact inverse Hessian, but a matrix compatible with this inverse Hessian within a factor of order of 1. It turns out that appropriate approximate inverse Hessians can be recursively computed at the lower average (over iterations) arithmetic cost than the exact ones; this possibility for acceleration was discovered in the same paper where the first polynomial-time interior-point method for LP was developed (we mean the landmark paper of Karmarkar [Ka 84]). As applied to interior-point methods for $(\mathcal{LQ})$ (and LP), this *Karmarkar acceleration* was used by most of researchers. Since it is used in all our strategies, it is reasonable to recall the idea of the Karmarkar acceleration.

Assume that we are generating a sequence of points $(t_1, x_1), \ldots, (t_i, x_i), \ldots$, $(t_i > 0,\ x_i \in G')$ and should compute symmetric positive definite matrices $P_i$ and $P_i^{-1}$ in such a way that

$$(*) \qquad \rho^{-1}\left(\Phi + t_i^{-1}F''(x_i)\right) \le P_i \le \rho\left(\Phi + t_i^{-1}F''(x_i)\right),$$

where $\Phi$ is a given symmetric positive semidefinite matrix and $\rho > 1$ is a given constant. Note that the matrices $T_i \equiv t_i^{-1}F''(x_i)$ are of the form $A^T D(i) A$, where $A$ is a constant $m \times n$ matrix and $D(i)$ is a diagonal $m \times m$ matrix with positive diagonal entries.

Karmarkar's scheme of computing $P_i$ and $P_i^{-1}$ is as follows. We set

$$P_i = \Phi + A^T D'(i) A$$

and update the new diagonal matrices $D'(i)$ according to the rules

$$D'(1) = D(1);$$

$$[D'(i+1)]_{jj} = \begin{cases} [D'(i+1)]_{jj}, & \text{if } \rho^{-1}[D(i+1)]_{jj} \le [D'(i)]_{jj} \le \rho[D(i+1)]_{jj}, \\ [D(i+1)]_{jj}, & \text{otherwise} \end{cases}$$

(note that this rule clearly maintains $(*)$). Now, for $i = 1$, we compute $P_i$ and then $P_i^{-1}$ straightforwardly. To transform $(P_i, P_i^{-1})$ into $(P_{i+1}, P_{i+1}^{-1})$, we act as follows. Let $\Delta(i) = D'(i+1) - D'(i)$ and let $k = k_i$ denote the number of nonzero diagonal entries in $\Delta(i)$. Consider the $k \times k$ diagonal matrix $\delta = \delta(i)$ with diagonal entries being nonzero diagonal entries of $\Delta(i)$ and let $\alpha = \alpha(i)$ be the $k \times n$ matrix formed by the rows of $A$ corresponding to the nonzero diagonal entries of $\Delta(i)$. Then

$$P_{i+1} = P_i + \alpha^T \delta \alpha,$$

and the above computation costs $O(kn^2)$ operations of the standard linear algebra routines.

Now, in the case of $k \le n$, to update $P_i^{-1}$ into $P_{i+1}^{-1}$, we can use the Sherman–Morrison formula

$$P_{i+1}^{-1} = P_i^{-1} - (\alpha P_i^{-1})^T \delta \{I_k + (\alpha P_i^{-1})\alpha^T \delta\}^{-1} (\alpha P_i^{-1}),$$

which takes $O(kn^2)$ operations; in the case of $k > n$, we can invert $P_{i+1}$ directly, which takes $O(n^3)$ operations. Thus, the cost of updating the pair $(P_i, P_i^{-1})$ always does not exceed $O(kn^2)$, while the straightforward computation of $(P_{i+1}, P_{i+1}^{-1})$ takes $O(mn^2)$ operations.

As already mentioned, in the case of the basic barrier method, we can replace exact Newton steps

$$x_{i+1} = x_i - [\nabla^2 F_{t_{i+1}}(x_i)]^{-1} \nabla F_{t_{i+1}}(x_i)$$

by approximate steps

$$x_{i+1} = x_i - t_i^{-1} P_i^{-1} \nabla F_{t_{i+1}}(x_i),$$

where $P_i$ satisfies $(*)$ with an appropriate *absolute* constant $\rho$ close to 1 (the matrix $\Phi$ involved into $(*)$ is either zero (at the preliminary stage) or the Hessian of the objective (at the main stage)). Thus, we can use the Karmarkar scheme for updating $P_i$ in the basic method (the resulting procedure is called the Karmarkar acceleration of the basic method). It turns out that the average (over iterations) value of $k$ in the latter method is of order of $m^{1/2}$, so that the average arithmetic cost of an iteration is $O(m^{1/2}n^2 + mn)$ instead of $O(mn^2)$ in the initial version of the basic method. Thus, the total arithmetic cost of finding an $\varepsilon$-solution by the accelerated basic method is

$$M^*(\varepsilon) = O(mn^2 + m^{3/2}n) \ln\left(\frac{m}{\alpha(G:w)\varepsilon}\right). \tag{8.1.6}$$

Note that the arithmetic cost of an $\varepsilon$-solution to $(\mathcal{LQ})$ depends, in particular, on the linear algebra tools used for scaling (i.e., for forming and exact or approximate inverting of the corresponding Hessians). Bounds (8.1.5) and (8.1.6) correspond to the traditional linear algebra routines, where the cost of inverting a $k \times k$ matrix is $O(k^3)$. However, there are theoretically less costly ways to perform the above computation, which require no more than $O(k^{2+\gamma})$ operations for certain $\gamma < 1$ (the best of the currently known values of $\gamma$ is 0.376... [CW 86]). Of course, the "advanced" linear algebra techniques reduce the cost of finding an $\varepsilon$-solution to $(\mathcal{LQ})$. In what follows, we fix $\gamma \in (0,1]$ and assume that we can invert a $k \times k$ matrix at the cost of $O(k^{2+\gamma})$ arithmetic operations.

### Overview

The aim of this chapter is to combine all aforementioned possibilities for accelerating the basic barrier method as applied to $(\mathcal{LQ})$, i.e., "large steps in $t$ with multistep prescaled first-order procedures for updating $x$," recursive computation of (approximate) inverse Hessians used for scaling, and "advanced" linear algebra routines.

Let us outline the contents of the chapter. In §8.2 we establish an important inequality that, in a sense, bounds from above the rate at which the path $x^*(\cdot)$ associated with problem $(\mathcal{LQ})$ can vary "at large"; this inequality forms the basis for all our further considerations. In §8.3 we describe three strategies with "large" steps in $t$ and multistep updating $x$ based on three (prescaled) first-order procedures, namely, the gradient descent method, the "optimal" method for smooth convex minimization, and the standard iterative method for solving the equation $\nabla F_{t^+} = 0$. In §8.4 we develop the strategy with the same stepsizes in $t$ as in the basic barrier method and the (approximate) Newton steps in $x$ computed by a preconditioned (in certain specific way) conjugate gradient method.

The results are as follows (for simplicity, we now restrict ourselves to the case of $n = O(m)$). The arithmetic cost of finding an $\varepsilon$-solution to $(\mathcal{LQ})$ for each of our procedures is of the type

$$O_\gamma(1)q(\gamma, m)\ln\left(\frac{m}{\alpha(G:w)\varepsilon}\right),$$

where $\alpha(G : w)$ and $\varepsilon$ are the same as in (8.1.5) and $q(\cdot,\cdot)$ is specific for the procedure under consideration. Our "reference point" is the Karmarkar acceleration of the basic barrier method

$$q_{\mathrm{Karm}}(\gamma, m) = m^{(5+\gamma)/2};$$

for the three strategies described in §8.3, we have

$$q_{\mathrm{GrDsc}}(\gamma, m) = m^{(20+\gamma)/(8-\gamma)}(\ln m)^{(1-\gamma)/(8-\gamma)},$$

$$q_{\text{Opt}}(\gamma, m) = m^{(10-\gamma)/(4-\gamma)}(\ln m)^{(1-\gamma)/(4-\gamma)},$$

$$q_{\text{It}}(\gamma) = m^{(5+\gamma)/2},$$

and, for the conjugate-gradient-based strategy from §8.4, we have

$$q_{\text{CG}}(\gamma, m) = m^{5/2+2\gamma^2/(2+3\gamma-\gamma^2)}.$$

We see that, in the case where $\gamma = 1$, i.e., for the traditional linear algebra, all strategies are of the same cost: $q.(1, m) = m^3$ (note that the strategies, however, differ from each other even in the case of $\gamma = 1$). The "ideal" case $\gamma = 0$ is similar: All $q.(0, \gamma)$ are (sometimes within logarithmic factors) equal to $m^{5/2}$. In the case of $0 < \gamma < 1$, the quality of the third strategy from §8.3 is the same as of the Karmarkar acceleration, and the remaining are better in order, the conjugate-gradient-based being the best one (for all $\gamma \in (0, 1)$). For example, for the best-known value of $\gamma$, the reference strategy yields

$$q_{\text{Karm}}(0.376..., m) = m^{2.688...},$$

while, for the conjugate-gradient-based strategy, we have

$$q_{\text{CG}}(0.376..., m) = m^{2.594...}.$$

As we have mentioned, in the case of $\gamma = 1$, all our strategies lead to the same complexity estimates as the widely used Karmarkar acceleration of the basic method. Nevertheless, we suppose that these new strategies are of practical interest even in the context of the standard linear algebra. Indeed, the basic method seems to work in accordance with its theoretical worst-case efficiency estimate, while the other strategies appear to be more flexible. We mean that the performance of the first-order procedures for smooth convex optimization, as well as the performance of the conjugate gradient method for solving linear systems, usually is better than that one prescribed by the theoretical worst-case analysis, so that we may hope that the proposed acceleration strategies typically are more efficient that the basic method with Karmarkar's acceleration.

### 8.1.2 Preliminary results

The remaining part of this section contains some preliminary results about advanced linear algebra.

In what follows, we fix $\gamma \in (0, 1]$ such that, for all positive integers $k$, a nonsingular $k \times k$ matrix can be inverted at the arithmetic cost $c_\gamma k^{2+\gamma}$. It is known that, in this case, the product of two $k \times k$ matrices can be computed at the cost of $O_\gamma(k^{2+\gamma})$ arithmetic operations (henceforth, the constant factors in $O_\gamma(\cdot)$ depend on $\gamma$ only, while the constant factors in $O(\cdot)$ are, as always, absolute constants). The following statement is a straightforward consequence of these assumptions.

**Lemma 8.1.1** *Let*

$$\sigma(l,k,r) = lkr(\min\{l,k,r\})^{\gamma-1}, \qquad l,k,r \in \mathbf{N}.$$

*The product of an $l \times k$ matrix $A$ and a $k \times r$ matrix $B$ can be computed at the cost of $O_\gamma(\sigma(l,k,r))$ arithmetic operations.*

**Proof.** Let $s = \min\{l,k,r\}$. Without loss of generality, we can assume that $l,k,r$ are divisible by $s$. After partitioning the matrices $A$ and $B$ into square $s \times s$ submatrices, we obtain $(l/s)\times(k/s)$ and $(k/s)\times(r/s)$ matrices $A'$, $B'$ with elements from the ring $\mathcal{L}$ of real $s \times s$ matrices. The multiplication of $A'$ and $B'$ in the traditional manner costs $O(lkrs^{-3})$ multiplications and additions of pairs of elements of $\mathcal{L}$. Each of these $\mathcal{L}$-operations costs no more than $O_\gamma(s^{2+\gamma})$ arithmetic operations, which proves the statement. □

Assume that the data in $(\mathcal{LQ})$ are represented in the natural way (by listing the entries of the matrix and the vectors) and let $\phi$ be a similarly represented convex quadratic form. Let

$$F_t^\phi(x) = t\phi(x) + F(x). \tag{8.1.7}$$

Henceforth, $f'$ and $f''$ denote the gradient and the Hessian of a function $f : G' \to \mathbf{R}$ with respect to the standard Euclidean structure on $\mathbf{R}^n$.

Let, for $x \in G'$, $t > 0$,

$$\begin{gathered}\delta(t,x) = (t^{-1/2}f_1^{-1}(x),\dots,t^{-1/2}f_m^{-1}(x))^T \in \mathbf{R}^m,\\ d(t,x) = (t^{-1}f_1^{-2}(x),\dots,t^{-1}f_m^{-2}(x))^T \in \mathbf{R}^m,\\ D(t,x) = D_t(x) = \operatorname{diag}\{d(t,x)\} \in \mathcal{D},\\ \Phi_t(x) = (F_t^\phi)'(x), \qquad \Psi_t(x) = (F_t^\phi)''(x),\end{gathered}$$

where $\mathcal{D}$ is the set of diagonal $m \times m$ matrices with positive diagonal entries. Let $Z$ be the $n \times m$ matrix with the columns $a_1,\dots,a_m$ and let

$$M(\phi,D) = \phi'' + ZDZ^T, \qquad D \in \mathcal{D}.$$

The $n \times n$ matrix $M(\phi,D)$ is symmetric and positive definite (the latter statement holds in view of the boundedness of $G'$). We use the notation $M_t^\phi(x)$ for the matrix $M(\phi,D_t(x))$; note that, in view of (8.1.7),

$$(F_t^\phi)''(x) = tM_t^\phi(x), \qquad x \in G'. \tag{8.1.8}$$

For a pair $h$, $s$ of positive $m$-dimensional vectors, let

$$\nu(h,s) = \max\{h_1/s_1, s_1/h_1,\dots,h_m/s_m, s_m/h_m\} - 1. \tag{8.1.9}$$

Henceforth,

$$\mathcal{S} \equiv \{d \in \mathbf{R}^m \mid d > 0\}.$$

The following lemma holds.

**Lemma 8.1.2** (i) *Given* $x \in G'$, $t > 0$, *and* $D \in \mathcal{D}$, *we can compute* $(F_t^\phi)'(x)$ *at the cost of* $O(mn)$ *operations; compute* $D_t(x)$ *at the cost of* $O(mn)$ *operations; compute the product of* $M(\phi, D)$ *and a given vector* $h \in \mathbf{R}^n$ *at the cost of* $O(mn)$ *operations; compute* $M(\phi, D)$ *at the cost of* $O_\gamma(mn^{1+\gamma})$ *operations.*

(ii) *Assume that we have computed* $D$, $D' \in \mathcal{D}$, *and the matrix* $L = [M(\phi, D)]^{-1}$ *and let* $k$ *be the number of diagonal positions in which the entries of* $D$ *and* $D'$ *do not coincide. Then the matrix* $[M(\phi, D')]^{-1}$ *can be computed at the cost of* $O_\gamma(m + l[n,k])$ *operations, where*

$$l[n,k] = \begin{cases} n^2 k^\gamma, & k \le n, \\ kn^{1+\gamma}, & k > n. \end{cases}$$

**Proof.** (i) The first and the second statements are evident; the third follows from the relation

$$M(\phi, D)h = \phi'' h + Z[D(Z^T h)].$$

The fourth statement follows from Lemma 8.1.1, since the computation of the $n \times m$ matrix $DZ^T$ costs $O(mn)$ operations, the multiplication of $Z$ and this matrix costs $O_\gamma(mn^{1+\gamma})$ operations, and it takes $O(n^2)$ operations to add $\phi''$ to the result.

(ii) If $k = 0$, then the statement is evident. Let $k$ be a positive integer. It is clear that

$$M'_{n,n} \equiv M(\phi, D') = M(\phi, D) + V_{n,k} S_{k,n} \equiv M_{n,n} + V_{n,k} S_{k,n}$$

(subscripts denote the numbers of rows and columns), where $V_{n,k}$ and $S_{k,n}$ can be computed at the cost $O(m + nk)$ operations.

Let $k \le n$. We have

$$[M'_{n,n}]^{-1} = M_{n,n}^{-1} - [M_{n,n}]^{-1} V_{n,k} (I_k + S_{k,n} V_{n,k})^{-1} S_{k,n} [M_{n,n}]^{-1},$$

where $I_k$ means the $k \times k$ identity matrix. By Lemma 8.1.1, the matrix $(I_k + S_{k,n} V_{n,k})$ can be computed at the cost $O_\gamma(nk^{1+\gamma})$; the resulting matrix can be inverted at the cost $O_\gamma(k^{2+\gamma})$; each of the remaining matrix multiplications costs no more than $O_\gamma(n^2 k^\gamma)$, thus $(M'_{n,n})^{-1}$ can be computed in no more than $O_\gamma(n^2 k^\gamma)$ operations.

Now let $k > n$. We have

$$(M'_{n,n})^{-1} = (I_n + (M_{n,n})^{-1} V_{n,k} S_{k,n})^{-1} (M_{n,n})^{-1}.$$

The matrix $I_n + (M_{n,n})^{-1} V_{n,k} S_{k,n}$ can be computed at the cost $O(kn^{1+\gamma})$ (Lemma 8.1.1), the resulting matrix can be inverted at the cost $O(n^{2+\gamma})$, and the result can be multiplied by $(M_{n,n})^{-1}$ at the cost $O(n^{2+\gamma})$. Thus, $(M'_{n,n})^{-1}$ can be computed at the cost of $O(kn^{1+\gamma})$ operations. □

## 8.2 The main inequality

Let us fix a convex quadratic form $\phi$ on $\mathbf{R}^n$. For $t > 0$, denote

$$x^*(t) = \operatorname{argmin}\{F_t^\phi(x) \mid x \in G'\}, \tag{8.2.1}$$

$$\xi^*(t) = \left\{t^{-1/2}f_1^{-1}(x^*(t)), \ldots, t^{-1/2}f_m^{-1}(x^*(t))\right\}^T \tag{8.2.2}$$

($F_t^\phi$ is defined by (8.1.7)).

The following lemma states an important regularity property of the trajectory $\xi^*$.

**Lemma 8.2.1** *Let $t_1$, $t_2 > 0$. Then*

$$\begin{aligned} &\{t_1t_2\}^{1/2}(x^*(t_1) - x^*(t_2))^T\phi''(x^*(t_1) - x^*(t_2)) \\ &\quad + \sum_{i=1}^m \{\xi_i^*(t_1) - \xi_i^*(t_2)\}^2\{\xi_i^*(t_1)\xi_i^*(t_2)\}^{-1} \\ &= m\{t_1^{1/2} - t_2^{1/2}\}^2\{t_1t_2\}^{-1/2}. \end{aligned} \tag{8.2.3}$$

**Proof**. We have

$$\phi'(x^*(t)) - t^{-1}\sum_{i=1}^m f_i'f_i^{-1}(x^*(t)) = 0$$

(note that $f_i'$ does not depend on $x$). Subtracting such an equality corresponding to $t = t_2$ from that one corresponding to $t = t_1$ and multiplying the resulting equality by $(x^*(t_1) - x^*(t_2))$, we obtain

$$\begin{aligned} &(x^*(t_1) - x^*(t_2))^T\phi''(x^*(t_1) - x^*(t_2)) \\ &= \sum_{i=1}^m \{t_1^{-1}[f_i(x^*(t_1)) - f_i(x^*(t_2)]f_i^{-1}(x^*(t_1)) \\ &\qquad -t_2^{-1}[f_i(x^*(t_1)) - f_i(x^*(t_2))]f_i^{-1}(x^*(t_2))\}, \end{aligned}$$

whence

$$\begin{aligned} &\sum_{i=1}^m t_1^{-1/2}t_2^{-1/2}\{\xi_i^*(t_1)/\xi_i^*(t_2) + \xi_i^*(t_2)/\xi_i^*(t_1)\} \\ &\quad +(x^*(t_1) - x^*(t_2))^T\phi''(x^*(t_1) - x^*(t_2)) = m(t_1^{-1} + t_2^{-1}), \end{aligned}$$

which immediately leads to (8.2.3). □

**Corollary 8.2.1** *Let $t_1$, $t_2 > 0$. Assume that $x(t_1)$, $x(t_2) \in G'$ are such that*

$$\lambda(F_{t_j}^\phi, x(t_j)) \le \lambda \le 0.1, \qquad j = 1, 2. \tag{8.2.4}$$

*Then*

$$
\begin{aligned}
&\{t_1t_2\}^{1/2}(x(t_1)-x(t_2))^T\phi''(x(t_1)-x(t_2)) \\
(8.2.5)\qquad &+\sum_{i=1}^{m}\{\delta_i(t_1,x(t_1))-\delta_i(t_2,x(t_2))\}^2\{\delta_i(t_1,x(t_1))\delta_i(t_2,x(t_2))\}^{-1} \\
&\le \mu_0^2\left\{m\{t_1^{1/2}-t_2^{1/2}\}^2+\omega^2(\lambda)\{t_1^{1/2}+t_2^{1/2}\}^2\right\}\{t_1t_2\}^{-1/2}
\end{aligned}
$$

*(recall that* $\omega(\lambda)=1-(1-3\lambda)^{1/3}$*) with an absolute constant* $\mu_0>0$.

**Proof.** Let us define $m$-dimensional vectors and $m\times m$ matrices as follows (below $t=t_1$ or $t=t_2$):

$$
\begin{aligned}
h(t) &= \{f_i(x^*(t))/f_i(x(t)) \mid 1\le i\le m\}^T, \\
h_-(t) &= \left(h_1^{-1}(t),\dots,h_m^{-1}(t)\right)^T, \\
H(t) &= \mathrm{diag}\{h(t)\}, \qquad e=(1,\dots,1)^T \\
g &= \{(\xi_i^*(t_2)/\xi_i^*(t_1))^{1/2} \mid 1\le i\le m\}^T, \\
g_- &= (g_1^{-1},\dots,g_m^{-1})^T, \\
\eta(t) &= \left(\sqrt{h_1(t)},\dots,\sqrt{h_m(t)}\right)^T, \\
\eta_-(t) &= (\eta_1^{-1}(t),\dots,\eta_m^{-1}(t))^T.
\end{aligned}
$$

We have

$$
\begin{aligned}
&D^2F_t^\phi(x(t))[x^*(t)-x(t),x^*(t)-x(t)] \\
&\quad= t(x^*(t)-x(t))^T\phi''(x^*(t)-x(t)) \\
&\qquad+\sum_{i=1}^{m}\{(x^*(t)-x(t))^Tf_i'\}^2f_i^{-2}(x(t)) \\
(8.2.6)\qquad&\quad= t\,(x^*(t)-x(t))^T\,\phi''\,(x^*(t)-x(t)) \\
&\qquad+\sum_{i=1}^{m}\left\{f_i^{-1}(x(t))\,[f_i(x^*(t))-f_i(x(t))]\right\}^2 \\
&\quad= t\,(x^*(t)-x(t))^T\,\phi''\,(x^*(t)-x(t)) \\
&\qquad+\parallel h(t)-e\parallel_2^2
\end{aligned}
$$

and similarly

$$
\begin{aligned}
(8.2.7)\qquad &D^2F_t^\phi(x^*(t))[x^*(t)-x(t),x^*(t)-x(t)] \\
&= t\,(x^*(t)-x(t))^T\,\phi''\,(x^*(t)-x(t))+\parallel h_-(t)-e\parallel_2^2\,.
\end{aligned}
$$

Thus, Theorem 2.2.2(iii) and (8.2.4) imply that

$$t\,(x^*(t)-x(t))^T\,\phi''\,(x^*(t)-x(t))+\parallel h(t)-e\parallel_2^2\leq\omega^2(\lambda) \tag{8.2.8}$$

and

$$t\,(x^*(t)-x(t))^T\,\phi''\,(x^*(t)-x(t))+\parallel h_-(t)-e\parallel_2^2\leq\frac{\omega^2(\lambda)}{(1-\omega(\lambda))^2}. \tag{8.2.9}$$

Furthermore, (8.2.3) implies that

$$\begin{aligned}&\{t_1t_2\}^{1/2}\,(x^*(t_1)-x^*(t_2))^T\,\phi''\,(x^*(t_1)-x^*(t_2))+\parallel g-g_-\parallel_2^2\\&=m\{t_1^{1/2}-t_2^{1/2}\}^2\{t_1t_2\}^{-1/2}\equiv\theta^2.\end{aligned} \tag{8.2.10}$$

Since for positive $s$ we have $(s-s^{-1})^2\geq(s-1)^2+(s^{-1}-1)^2$, (8.2.10) leads to

$$\begin{aligned}&\{t_1t_2\}^{1/2}\,(x^*(t_1)-x^*(t_2))^T\,\phi''\,(x^*(t_1)-x^*(t_2))\\&\quad+\parallel g-e\parallel_2^2+\parallel g_--e\parallel_2^2\leq\theta^2.\end{aligned} \tag{8.2.11}$$

By (8.2.8), (8.2.9), we have

$$t\,(x^*(t)-x(t))^T\,\phi''\,(x^*(t)-x(t))+\parallel \eta(t)-e\parallel_2^2\leq\omega^2(\lambda), \tag{8.2.12}$$

$$t(x^*(t)-x(t))^T\phi''(x^*(t)-x(t))+\parallel \eta_-(t)-e\parallel_2^2\leq\frac{\omega^2(\lambda)}{(1-\omega(\lambda))^2}. \tag{8.2.13}$$

Therefore,

(8.2.14)

$$\begin{aligned}\zeta&\equiv\{t_1t_2\}^{1/2}\,(x(t_1)-x(t_2))^T\,\phi''\,(x(t_1)-x(t_2))\\&\quad+\sum_{i=1}^m\{\delta_i(t_1,x(t_1))-\delta_i(t_2,x(t_2))\}^2\,\{\delta_i(t_1,x(t_1))\delta_i(t_2,x(t_2))\}^{-1}\\&=\{t_1t_2\}^{1/2}\,(x(t_1)-x(t_2))^T\,\phi''\,(x(t_1)-x(t_2))\\&\quad+\parallel H^{-1/2}(t_1)H^{1/2}(t_2)g-H^{1/2}(t_1)H^{-1/2}(t_2)g_-\parallel_2^2\\&\leq\{t_1t_2\}^{1/2}\,(x(t_1)-x(t_2))^T\,\phi''\,(x(t_1)-x(t_2))\\&\quad+2\parallel H^{-1/2}(t_1)H^{1/2}(t_2)g-e\parallel_2^2\\&\quad+2\parallel H^{-1/2}(t_2)H^{1/2}(t_1)g_--e\parallel_2^2.\end{aligned}$$

Furthermore,

$$\begin{aligned}\| H^{-1/2}(t_1)H^{1/2}(t_2)g - e \|_2 &\le \| H^{-1/2}(t_1)H^{1/2}(t_2)(g-e) \|_2 \\ &\quad + \| H^{-1/2}(t_1)\{(H^{1/2}(t_2)e - e\} \|_2 \\ &\quad + \| H^{-1/2}(t_1)e - e \|_2 \\ &\le \| \eta_-(t_1) \|_\infty \| \eta(t_2) \|_\infty \| g - e \|_2 \\ &\quad + \| \eta_-(t_1) \|_\infty \| \eta(t_2) - e \|_2 + \| \eta_-(t_1) - e \|_2 \\ &\le \frac{1+\omega(\lambda)}{1-\omega(\lambda)}\theta + \frac{\omega(\lambda)}{1-\omega(\lambda)} + \frac{\omega(\lambda)}{1-\omega(\lambda)}\end{aligned}$$

(we considered (8.2.12), (8.2.13)). The same estimate holds for the quantity $\| H^{1/2}(t_2)^{1/2}H^{-1/2}(t_1)g_- - e \|_2$.

Thus, (8.2) implies that

$$\zeta \le \{t_1t_2\}^{1/2}\left(x(t_1) - x(t_2)\right)^T \phi''\left(x(t_1) - x(t_2)\right) + 24(\theta + \omega(\lambda))^2. \tag{8.2.15}$$

We also have

$$\begin{aligned}&\{t_1t_2\}^{1/2}\left(x(t_1) - x(t_2)\right)^T \phi''\left(x(t_1) - x(t_2)\right) \\ &\le 3\{t_1t_2\}^{1/2}\{(x^*(t_1) - x^*(t_2))^T\phi''(x^*(t_1) - x^*(t_2)) \\ &\qquad +(x^*(t_1) - x(t_1))^T\phi''(x^*(t_1) - x(t_1)) \\ &\qquad +(x^*(t_2) - x(t_2))^T\phi''(x^*(t_2) - x(t_2))\} \\ &\le 3\{t_1t_2\}^{1/2}\left\{\{t_1t_2\}^{-1/2}\theta^2 + t_1^{-1}\omega^2(\lambda) + t_2^{-1}\omega^2(\lambda)\right\}\end{aligned}$$

(the latter inequality holds in view of (8.2.11), (8.2.12)). The obtained inequality combined with (8.2.15) proves (8.2.5). □

## 8.3 Multistep barrier methods

### 8.3.1 Notation

Recall that the barrier method, when applied to $(\mathcal{LQ})$, follows the trajectory (8.2.1), where $\phi$ is a linear form (at the preliminary stage) and $\phi = \psi$ (at the main stage). The method computes approximations $x(t)$ to $x^*(t)$ along certain sequence $\{t_i \mid i \ge 0\}$ and maintains the inequality

$$\lambda(F_{t_i}^{\phi}, x(t_i)) \le \lambda,$$

where $\lambda$ is an appropriate absolute constant. In fact, this inequality is sufficient for ensuring the efficiency estimate, no matter by which strategy this inequality is provided. An example of such a strategy was described in §3.2.2. In what follows, three more strategies are presented.

Let us introduce the following subproblem $\mathcal{P} = \mathcal{P}(\tau, y)$ : Given $\tau > 0$ and $y \in G'$ such that

$$\lambda(F_\tau^\phi, y) \le \lambda, \tag{8.3.1}$$

find $y' \in G'$ and $\tau'$ satisfying the relations

$$\lambda(F_{\tau'}^\phi, y') \le \lambda \tag{8.3.2}$$

and

$$\tau' \begin{cases} \le \tau/2, & \text{at the preliminary stage,} \\ \ge 2\tau, & \text{at the main stage.} \end{cases} \tag{8.3.3}$$

Assume that $\mathcal{M}$ is a procedure that solves this subproblem (for each $\tau$, $y$ satisfying (8.3.1)) at the cost of $M$ arithmetic operations. Then we can find an $\varepsilon$-solution of $(\mathcal{LQ})$ in no more than

$$M^*(\varepsilon) = O\left(M \ln\left(\frac{m}{\alpha(G:w)\varepsilon}\right)\right)$$

arithmetic operations. Indeed, it suffices to follow the trajectory (8.2.1) in accordance with the scheme

$$\{\tau(=t_i),\ y(=x(t_i))\} \overset{\mathcal{M}}{\mapsto} \{t_{i+1} = \tau',\ x(t_{i+1}) = y'\}$$

(compare with the proofs of Propositions 3.2.3 and 3.2.4).

Thus, we can restrict ourselves to the subproblem $\mathcal{P}$.

**Sets $K_\alpha(x)$**

For $x \in G'$ and $\alpha > 0$, let

$$K_\alpha(x) = \{y \in G' \mid (1+\alpha)^{-1} f_i(x) < f_i(y) < (1+\alpha) f_i(x),\ 1 \le i \le m\}. \tag{8.3.4}$$

Let $\mu_1 > 0$ be an *absolute constant* such that

$$|\ln(r/s)|^2 \le \mu_1^2 \frac{\{r-s\}^2}{rs} \tag{8.3.5}$$

for all $s$, $r > 0$ (we can take $\mu_1 = 2$).

**Lemma 8.3.1** (i) *Let $x \in G'$, $\alpha > 0$, $t$, $s > 0$. Then*

$$\begin{aligned} &\forall y \in K_\alpha(x) : \ \| (\Psi_t(x))^{-1/2} \Psi_s(y) (\Psi_t(x))^{-1/2} - I_n \| \\ &\le \max\{2\alpha + \alpha^2; |1 - s/t|\}. \end{aligned} \tag{8.3.6}$$

(ii) *There exists an absolute constant* $\mu_2 > 0$ *such that, for each* $t,\ t' > 0,\ x,\ x' \in G'$ *and* $\lambda \in (0, 0.1)$ *satisfying the conditions*

$$\lambda(F_t^\phi, x) \le \lambda,\quad \lambda(F_{t'}^\phi, x') \le \lambda,\quad \tfrac{1}{2} \le t/t' \le 2, \tag{8.3.7}$$

*the implications*

$$\alpha \ge \mu_2\{m \ln^2(t/t') + \omega^2(\lambda) + 1\} \;\Rightarrow\; x' \in K_\alpha(x), \tag{8.3.8}$$

$$\alpha \ge \mu_2\{m \ln^2(t/t') + \omega^2(\lambda)\}^{1/2} \;\Rightarrow\; x' \in K_\alpha(x) \tag{8.3.9}$$

*hold.*

**Proof.** (i) We have

$$\Psi_r(u) = r\phi'' + Z\,\mathrm{Diag}\,\{f_i^{-2}(u)\}Z^T,$$

whence

$$\begin{aligned}&\min\{s/t, f_1^2(x)/f_1^2(y), \ldots, f_m^2(x)/f_m^2(y)\}\Psi_t(x)\\ &\le \Psi_s(y) \le \max\{s/t, f_1^2(x)/f_1^2(y), \ldots, f_m^2(x)/f_m^2(y)\}\Psi_t(x)\end{aligned}$$

or

$$\begin{aligned}&\min\{s/t, f_1^2(x)/f_1^2(y), \ldots, f_m^2(x)/f_m^2(y)\}I_n \le \Psi_t^{-1/2}(x)\Psi_s(y)\Psi_t^{-1/2}(x)\\ &\le \max\{s/t, f_1^2(x)/f_1^2(y), \ldots, f_m^2(x)/f_m^2(y)\}I_n.\end{aligned}$$

The latter relation immediately implies (8.3.6).

(ii) Let

$$\theta = (t/t')^{1/2}, \quad v_i = \delta_i(t,x)/\delta_i(t',x').$$

Then, by virtue of (8.2.5), we have

$$(1-v_i)^2/v_i \le \mu_0^2\{m(1-\theta)^2\theta^{-1} + \omega^2(\lambda)(1+\theta)^2\theta^{-1}\};$$

moreover, $\frac{1}{2} \le \theta^2 \le 2$. Therefore

$$\max\{v_i, 1/v_i\} \le O(m \ln^2\theta + \omega^2(\lambda) + 1),$$

$$\max\{v_i, 1/v_i\} \le O(1)(m \ln^2\theta + \omega^2(\lambda))^{1/2}.$$

It remains to note that $f_i(x')/f_i(x) = \{t'/t\}^{-1/2}v_i$. $\square$

For $a > 0$ and $q \in (0,1)$, let the function $g(t) \equiv g(a,q,t)$ be defined as

$$g(t) = \begin{cases} -\ln(a/q) - (q/a)(t - a/q) + (q/a)^2(t-a/q)^2, & \text{if } t \ge a/q;\\ -\ln(t), & \text{if } qa \le t \le a/q;\\ -\ln(qa) - (qa)^{-1}(t-qa) + (qa)^{-2}(t-qa)^2, & \text{if } t \le qa.\end{cases}$$

Also, let

$$\alpha(q) = \frac{1-q}{q}.$$

It is clear that the function $g(t)$ is a $C^2$ smooth extension of the function $(-\ln t)$ from the segment $[qa, a/q]$ onto $\mathbf{R}$ (the second derivative of the extension is constant for $t > a/q$ and for $t < qa$).

For $u \in G'$, let

$$F_{u,q}(x) = \sum_{i=1}^{m} g(f_i(u), q, f_i(x)) : \mathbf{R}^n \to \mathbf{R}.$$

The following statement is evident.

**Lemma 8.3.2** *For each $u \in G'$, we have*

$$x \in K_{\alpha(q)}(u) \;\Rightarrow\; F_{u,q}(x) = F(x); \tag{8.3.10}$$

*moreover, for each $x \in \mathbf{R}^n$, we have*

$$q^2 F''(u) \le F''_{u,q}(x) \le q^{-2} F''(u). \tag{8.3.11}$$

### 8.3.2 Gradient-descent-based acceleration

Let $2 \ge \rho > 1$, $2 > \eta > 1$, $\lambda \in (0, 0.1]$. We will describe a procedure $\mathcal{M}(\rho, \eta, \lambda)$, which solves the subproblem $\mathcal{P}$. This procedure performs a "large" step in $t$. The stepsize depends on $\gamma$, $m$, $n$; in the case of $\gamma = 1$ and $n = O(m)$, the optimal step is

$$t \to t' = (1 \pm O(m^{-3/7}))t$$

instead of the usual step

$$t \to (1 \pm O(m^{-1/2}))t.$$

The return to the neighbourhood of the trajectory $x^*(\cdot)$ is performed by minimizing $F^{\phi}_{t'}$ by the gradient descent method. To avoid some difficulties (coming from the restrictions $x \in G'$), it is convenient to apply the gradient descent method not to $F^{\phi}_{t'}$, but rather to the function $F_{u,q}$, which is defined (and strongly convex) on the whole space and at the same time coincides with $F^{\phi}_{t'}$ in a neighbourhood of $x^*(t')$. The latter property is provided by an appropriate choice of $u$ and $q$. Of course, the gradient descent method corresponds not to the initial Euclidean structure on $\mathbf{R}^n$ but to the structure defined by a matrix close to $(F^{\phi}_{t'})''$.

The procedure is as follows.

*Initialization.* Let $\beta$ be the positive root of the equation

$$\beta^2/(1-\beta) = \lambda^2,$$

$$q = \left(1 + \mu_2[m \ln^2 \eta + \omega^2(\lambda) + 1]\right)^{-1},$$
$$h = (q^{-2}\rho\eta)^{-1},$$
$$\omega = q^3\beta(\rho\eta)^{-3/2},$$
$$x_0 = y, \qquad t_0 = \tau.$$

Compute $d^0 = d(t_0, x_0)$ and the matrices

$$M^{\phi}_{t_0}(x_0), \quad Q_0 = (M^{\phi}_{t_0})^{-1}.$$

*The kth step,* $k \geq 0$. Assume that, after $k-1$ steps of the procedure, we have formed a point $x_k \in G'$, a vector $d^k \in \mathcal{S}$, a number $t_k > 0$ and the matrix

(8.3.12) $$(\mathrm{I}_k): \qquad Q_k = (M(\phi, D_k))^{-1}, \qquad D_k \equiv \mathrm{Diag}\,\{d^k\},$$

such that

(8.3.13) $$(\mathrm{J}_k): \qquad \nu(d^k, d(t_k, x_k)) \leq \rho$$

(for the definition of $\nu$, see (8.1.9)). Note that the initialization rules ensure $(\mathrm{I}_0)$, $(\mathrm{J}_0)$.

At the $k$th step, we

(a) Set

$$t_{k+1} = \begin{cases} t_k\eta, & \text{at the main stage,} \\ t_k/\eta, & \text{at the preliminary stage,} \end{cases}$$
$$p_{k+1}(x) = t_{k+1}\phi(x) + F_{x_k,q}(x);$$

(b) Minimize the function $p_{k+1}$ over $x \in \mathbf{R}^n$ by the gradient descent method corresponding to the metric defined by the matrix $Q_k^{-1}$, i.e., set $x_{k,0} = x_k$ and compute

$$x_{k,i+1} = x_{k,i} - t_{k+1}^{-1} h Q_k p'_{k+1}(x_{k,i}).$$

The process is terminated at the earliest step $l$ when the condition

$$(p'_{k+1}(x_{k,l}))^T t_{k+1}^{-1} Q_k p'_{k+1}(x_{k,l}) \leq \omega^2$$

holds;

(c) Set $x_{k+1} = x_{k,l}$;

(d) Compute $d(t_{k+1}, x_{k+1})$ and $d^{k+1} \in \mathcal{S}$:

$$d_i^{k+1} = \begin{cases} d_i^k, & d_i^k/\rho \leq d_i(t_{k+1}, x_{k+1}) \leq d_i^k\rho, \\ d_i(t_{k+1}, x_{k+1}), & \text{otherwise} \end{cases}$$

(note that this updating ensures $(\mathrm{J}_{k+1})$) and use $Q_k$ to compute $Q_{k+1}$ in accordance with $(\mathrm{I}_{k+1})$.

This completes the $k$th step of $\mathcal{M}(\rho, \eta, \lambda)$.

If $t_{k+1}/t_0 \geq 2$ (at the main stage) or $t_{k+1}/t_0 \leq 1/2$ (at the preliminary stage), then set

$$k^* = k, \quad \tau' = t_{k+1}, \quad y' = x_{k+1}$$

and terminate; otherwise, perform the $(k+1)$th step.

**Theorem 8.3.1** *Each of the points $x_k$, $0 \le k \le k^*+1$, produced by $\mathcal{M}(\rho,\eta,\lambda)$ belongs to $G'$ and satisfies the relation*

$$(\mathrm{K}_k): \qquad \lambda(F^{\phi}_{t_k}, x_k) \le \lambda, \tag{8.3.14}$$

*and the procedure solves $\mathcal{P}$. Moreover,*

$$k^* \le O\left(\frac{1}{\eta-1}\right). \tag{8.3.15}$$

*Let $\eta$ satisfy the condition*

$$m^{1/2}(\eta-1) \ge 1. \tag{8.3.16}$$

*Then the arithmetic cost of $\mathcal{M}(\rho,\eta,\lambda)$ does not exceed the quantity*

$$\begin{aligned} M^{(1)} = O_\gamma(q^{-4}(\eta-1)^{-1}mn\ln(mq^{-1}\lambda^{-1}) \\ +mn^{1+\gamma} + (\eta-1)^{-1+\gamma}(\rho-1)^{-1}m^\gamma n^2). \end{aligned} \tag{8.3.17}$$

*In particular, under the choice*

$$\eta = 1 + m^{-(5-\gamma)/(8-\gamma)}\left(\frac{n^*}{\ln m}\right)^{1/(8-\gamma)}, \qquad \rho = 1.5, \qquad \lambda = 0.1, \tag{8.3.18}$$

*where*

$$n^* = \max\{n, m^{(2-\gamma)/2}\ln m\}, \tag{8.3.19}$$

*we have*

$$\begin{aligned} M^{(1)} &\le O_\gamma(m^{(5+2\gamma)/(8-\gamma)}(n^*)^{(15-\gamma)/(8-\gamma)}(\ln m)^{(1-\gamma)/(8-\gamma)}) \\ &\le O_\gamma(m^{(20+\gamma)/(8-\gamma)}(\ln m)^{(1-\gamma)/(8-\gamma)}). \end{aligned} \tag{8.3.20}$$

*Hence, under the choice (8.3.18), we have*

$$\begin{aligned} M^*(\varepsilon) &\le O_\gamma\left(M^{(1)}\ln\left(\frac{m}{\alpha(G:w)\varepsilon}\right)\right) \\ &\le O_\gamma\left(m^{(20+\gamma)/(8-\gamma)}\{\ln m\}^{(1-\gamma)/(8-\gamma)}\ln\left(\frac{m}{\alpha(G:w)\varepsilon}\right)\right). \end{aligned} \tag{8.3.21}$$

**Proof.** $1^0$. (8.3.15) immediately follows from (a).

$2^0$. Let us verify $(\mathrm{K}_k)$. For $k=0$, this relation holds by virtue of the initialization rules and (8.3.1). Assume that $(\mathrm{K}_k)$ holds for some $k \le k^*$ and let us prove that $(\mathrm{K}_{k+1})$ also holds. Denote $C_k = p''_{k+1}(x_k)$. Then, by virtue of Lemma 8.3.2, we have

$$q^2 C_k \le p''_{k+1}(x) \le q^{-2}C_k. \tag{8.3.22}$$

Furthermore, ($\mathrm{I}_k$), ($\mathrm{J}_k$), and the relation $|\ln(t_k/t_{k+1})| = \ln\eta$ imply that

$$\rho^{-1}\eta^{-1}t_kQ_k^{-1} \le C_k \le \rho\eta t_kQ_k^{-1}. \tag{8.3.23}$$

Consider $\mathbf{R}^n$, as provided with the scalar product

$$\langle u, v\rangle = u^T t_kQ_k^{-1}v,$$

and let $\|\cdot\|$ be the corresponding norm. Inequalities (8.3.22) and (8.3.23) mean that $p_{k+1}$ is a strongly convex function with respect to our Euclidean structure, and the spectrum of its Hessian belongs to the segment

$$[r, R] \equiv [q^2\rho^{-1}\eta^{-1}, q^{-2}\rho\eta].$$

By definition of $h$, process (b) describes the usual gradient descent method as applied to $p_{k+1}$. Thus, by the standard arguments, we have, for all $i$,

$$(p'_{k+1}(x_{k,i}))^T t_k^{-1}Q_kp'_{k+1}(x_{k,i}) \le (1-r/R)^iR^2 \| x_k - x_k^* \|^2, \tag{8.3.24}$$

where $x_k^*$ is the minimizer of $p_{k+1}$.

By definition of $q$ and in view of ($\mathrm{K}_k$) and Lemma 8.3.1, we have $x^*(t_{k+1}) \in K_{\alpha(q)}(x_k)$. Hence $F^{\phi}_{t_{k+1}}$ coincides with $p_{k+1}$ in a neighbourhood of $x^*(t_{k+1})$, which means that $x_k^* = x^*(t_{k+1})$. By virtue of Corollary 8.2.1, relation ($\mathrm{K}_k$) leads to

(8.3.25)
$$\| x_k - x_k^* \|^2 = (x_k - x_k^*)^T t_kQ_k^{-1}(x_k - x_k^*) \le \rho(x_k - x_k^*)^T\Psi_{t_k}(x_k)(x_k - x_k^*)$$

$$= \rho t_k(x_k - x_k^*)^T\phi''(x_k - x_k^*) + \rho\sum_{i=1}^{m}\{1 - f_i^{-1}(x_k)f_i(x_k^*)\}^2$$

$$\le \rho(t_k/t_{k+1})^{1/2}\mu_0^2\left\{m\{t^{1/2} - t_{k+1}^{1/2}\}^2 + \omega^2(\lambda)\{t_k^{1/2} + t_{k+1}^{1/2}\}^2\right\}\{t_kt_{k+1}\}^{-1/2}$$

$$+\rho m(1-q^{-1})^2 \le O(mq^{-2}).$$

Therefore (8.3.24) implies that

$$(p'_{k+1}(x_{k,i}))^T t_k^{-1}Q_kp'_{k+1}(x_{k,i}) \le \{1 - O(q^4)\}^iO(mq^{-6}), \tag{8.3.26}$$

whence

$$l \le l^* = O(q^{-4})\{\ln(m/q) + \ln(1/\omega)\}. \tag{8.3.27}$$

Let

$$\nabla p_{k+1}(x) = t_k^{-1}Q_kp'_{k+1}(x)$$

be the gradient of $p_{k+1}$ with respect to the above Euclidean structure. Then

$$\| \nabla p_{k+1}(x_{k+1}) \|^2 \le \omega^2$$

(this is the termination rule in b)). Hence

$$\| x_{k+1} - x^*(t_{k+1}) \| \le \rho\eta q^{-2}\omega. \tag{8.3.28}$$

Moreover, we have (see (8.3.22), (8.3.23))

$$\Psi_{t_{k+1}}(x^*(t_{k+1})) = p''_{k+1}(x^*(t_{k+1})) \le \rho\eta q^{-2}(t_k Q_k^{-1}),$$

and (8.3.28) implies that

$$(x_{k+1} - x^*(t_{k+1}))^T \Psi_{t_{k+1}}(x^*(t_{k+1}))(x_{k+1} - x^*(t_{k+1})) \le \rho^3\eta^3 q^{-6}\omega^2, \tag{8.3.29}$$

which, by choice of $\omega$ and by our standard arguments, leads to $(\mathrm{K}_{k+1})$.

Note that $(\mathrm{K}_{k^*+1})$ implies (8.3.2). The relation $\tau'/\tau \ge 2$ (at the main stage), $\tau'/\tau \le \frac{1}{2}$ (at the preliminary stage) immediately follows from the termination rule (see d)). Thus, our procedure solves $\mathcal{P}$.

$3^0$. It remains to evaluate the arithmetic cost of the procedure. It is easily seen that the total cost $M'$ of all computations, *excluding computation of the matrices* $Q_k$, does not exceed $O(K^* l^* mn)$, $K^* = k^* + 1$. Now let us evaluate the total cost $M''$ of computing the matrices. First, $Q_0$ can be computed at the cost $\le O_\gamma(mn^{1+\gamma})$ (see Lemma 8.1.2). Now let

$$A_k = \{i \mid d_i^k \ne d_i^{k+1}\}, \quad r_k = |A_k|, \quad r = \sum_{k=0}^{k^*} r_k.$$

Then, by virtue of Lemma 8.1.2,

$$\begin{aligned} M'' &\le O_\gamma(mn^{1+\gamma}) + \sum_{k=0}^{k^*} O_\gamma(l[n, r_k]) \\ &\le O_\gamma(mn^{1+\gamma} + mK^*) + \sum_{k\in\mathcal{I}} O_\gamma(r_k n^{1+\gamma}) + \sum_{k\in\mathcal{J}} O_\gamma(n^2 r_k^\gamma), \end{aligned} \tag{8.3.30}$$

where

$$\mathcal{I} = \{k \mid 0 \le k \le k^*, r_k > n\}, \qquad \mathcal{J} = \{k \mid 0 \le k \le k^*, r_k \le n\}.$$

Let

$$h^0 = 0 \in \mathbf{R}^m, \quad (h^k)_i = \left| \ln\left( \frac{d_i(t_{k+1}, x_{k+1})}{d_i(t_k, x_k)} \right) \right|, \ 1 \le i \le m, \ 1 \le k \le k^*.$$

It is clear from (d) that

$$r = \sum_{k=0}^{k^*} r_k \le O\left(\frac{1}{\rho - 1}\right) \sum_{k=0}^{k^*} \| h^k \|_1 . \tag{8.3.31}$$

Furthermore,

$$\sum_{k=0}^{k^*} \| h^k \|_1 \le m^{1/2} \sum_{k=0}^{k^*} \| h^k \|_2 .$$

We have (see (8.3.5) and Corollary 8.2.1)

(8.3.32)
$$\begin{aligned}\| h^k \|_2^2 &= \sum_{i=1}^{m} 4\ln^2(\delta_i(t_{k+1}, x_{k+1})/\delta_i(t_k, x_k)) \\ &\leq 4\mu_1^2 \sum_{i=1}^{m} \{\delta_i(t_{k+1}, x_{k+1}) - \delta_i(t_k, x_k)\}^2 \{\delta_i(t_{k+1}, x_{k+1})\delta_i(t_k, x_k)\}^{-1} \\ &\leq O(m(\eta-1)^2 + 1).\end{aligned}$$

Since $k^* \leq O\left((\eta-1)^{-1}\right)$, (8.3.32) implies that

$$r \leq O\left(\frac{1}{\rho-1}\right) O\left(m + \frac{m^{1/2}}{\eta-1}\right) \leq O\left(\frac{m}{\rho-1}\right) \tag{8.3.33}$$

(the latter inequality holds by (8.3.16)). Hence, (8.3.30) implies that

$$\begin{aligned}&\sum_{k\in\mathcal{I}} O_\gamma(r_k n^{1+\gamma}) \leq O\left(\frac{1}{\rho-1}\right) O_\gamma(mn^{1+\gamma}), \\ &\sum_{k\in\mathcal{J}} O_\gamma(n^2 r_k^\gamma) \leq O_\gamma\left(|\mathcal{J}| \left(\frac{r}{|\mathcal{J}|}\right)^\gamma n^2\right) \\ &\qquad\qquad \leq O((\rho-1)^{-\gamma}) O_\gamma\left((K^*)^{1-\gamma} m^\gamma n^2\right) \\ &\qquad\qquad \leq O_\gamma\left((\eta-1)^{\gamma-1} m^\gamma n^2\right) O\left((\rho-1)^{-\gamma}\right).\end{aligned}$$

Thus, (8.3.30)–(8.3.32), (8.3.27), and (8.3.15) imply that

$$\begin{aligned}M^{(1)} \leq O_\gamma\,&(q^{-4}(\eta-1)^{-1} mn \ln(mq^{-1}\lambda^{-1}) \\ &+ mn^{1+\gamma} + (\eta-1)^{-1+\gamma}(\rho-1)^{-\gamma} m^\gamma n^2). \qquad \square\end{aligned}$$

Note that, in the case of the traditional linear algebra techniques ($\gamma = 1$), relations (8.3.20), (8.3.21) become

$$M^*(\varepsilon) \leq O\left((n^*)^2 m \ln\left(\frac{m}{\alpha(G:w)\varepsilon}\right)\right), \qquad n^* = \max\{n, m^{1/2}\ln m\}.$$

### 8.3.3 Acceleration based on the optimal method

It is known that the rate of convergence of the gradient descent method as applied to strongly convex problems can be improved. Therefore, it is reasonable to replace in the above scheme the gradient method by the so-called "optimal" method for smooth convex optimization [Ns 83], [Ns 88d]. Now we describe the corresponding procedure $\mathcal{M}^*(\rho, \eta, \lambda)$. The parameters of the procedure are subject to the same restrictions as in §8.3.2. The procedure is as follows.

*Initialization.* Let $\beta$ be the positive root of the equation

$$\beta^2/(1-\beta) = \lambda^2,$$

$$q = (1 + \mu_2[m \ln^2 \eta + \omega^2(\lambda) + 1])^{-1},$$
$$M = q^{-2}\rho\eta,$$
$$\omega = q^3 \beta(\rho\eta)^{-3/2},$$
$$x_0 = y, \qquad t_0 = \tau.$$

Compute $d^0 = d(t_0, x_0)$ and the matrices

$$M^{\phi}_{t_0}(x_0), \qquad Q_0 = (M^{\phi}_{t_0})^{-1}.$$

*The kth step,* $k \geq 0$. Assume that, after $k-1$ steps of the procedure, we have formed a point $x_k \in G'$, a vector $d^k \in \mathcal{S}$, a number $t_k > 0$, and the matrix

(8.3.34) $\qquad (\mathrm{L}_k): \qquad Q_k = (M(\phi, D_k))^{-1}, \qquad D_k \equiv \mathrm{Diag}\,\{d^k\},$

such that

(8.3.35) $\qquad (\mathrm{M}_k): \qquad \nu(d^k, d(t_k, x_k)) \leq \rho$

(note that the initialization rules ensure $(\mathrm{L}_0)$, $(\mathrm{M}_0)$).

At the $k$th step, we

(a) Set

$$t_{k+1} = \begin{cases} t_k\eta, & \text{at the main stage,} \\ t_k/\eta, & \text{at the preliminary stage;} \end{cases}$$
$$p_{k+1}(x) = t_{k+1}\phi(x) + F_{x_{k,q}}(x);$$

(b) Minimize the function $p_{k+1}$ over $x \in \mathbf{R}^n$ by the "optimal" method for smooth convex optimization (the method corresponds to the metric defined by the matrix $Q_k^{-1}$), i.e., set $x_{k,0} = x_k$, $A_{k,0} = M$, $v_{k,0} = x_k$; at the $i$th step of the minimization process ($i \geq 0$), we

1. Compute $a_{k,i} > 0$ as a root of the equation

$$a_{k,i}^2 = (1 - a_{k,i})A_{k,i}M^{-1};$$

2. Set

$$y_{k,i} = a_{k,i}v_{k,i} + (1 - a_{k,i})x_{k,i},$$
$$A_{k,i+1} = a_{k,i}q^2 + (1 - a_{k,i})A_{k,i},$$
$$x_{k,i+1} = y_{k,i} - M^{-1}t_k^{-1}Q_k p'_{k+1}(y_{k,i}),$$
$$v_{k,i+1} = (1 - a_{k,i})A_{k,i}A_{k,i+1}^{-1}v_{k,i} + a_{k,i}q^2 A_{k,i+1}^{-1} y_{k,i} - a_{k,i}A_{k,i+1}^{-1}t_k^{-1}Q_k p'_{k+1}(y_{k,i}).$$

The process is terminated at the earliest step $l$ when the condition

$$(p'_{k+1}(x_{k,l}))^T t_{k+1}^{-1} Q_k p'_{k+1}(x_{k,l}) \leq \omega^2$$

holds;

(c) Set $x_{k+1} = x_{k,l}$;

(d) Compute $d(t_{k+1}, x_{k+1})$ and $d^{k+1} \in \mathcal{S}$,

$$d_i^{k+1} = \begin{cases} d_i^k, & d_i^k/\rho \le d_i(t_{k+1}, x_{k+1}) \le d_i^k \rho, \\ d_i(t_{k+1}, x_{k+1}), & \text{otherwise} \end{cases}$$

(note that this updating ensures $(\mathrm{M}_{k+1})$) and use $Q_k$ to compute $Q_{k+1}$ in accordance with $(\mathrm{L}_{k+1})$.

The $k$th step of $\mathcal{M}^*(\rho, \eta, \lambda)$ is completed. If $t_{k+1}/t_0 \ge 2$ (at the main stage) or $t_{k+1}/t_0 \le \frac{1}{2}$ (at the preliminary stage), then set

$$k^* = k, \quad \tau' = t_{k+1}, \quad y' = x_{k+1}$$

and terminate; otherwise, perform the $(k+1)$th step.

**Theorem 8.3.2** *Each of the points $x_k$, $0 \le k \le k^*+1$ produced by $\mathcal{M}^*(\rho, \eta, \lambda)$, belongs to $G'$ and satisfies the relation*

$$(N_k): \qquad \lambda(F_{t_k}^{\phi}, x_k) \le \lambda, \tag{8.3.36}$$

*and the procedure solves $\mathcal{P}$.*

*Moreover,*

$$k^* \le O\left(\frac{1}{\eta - 1}\right). \tag{8.3.37}$$

*Let $\eta$ satisfy the condition*

$$m^{1/2}(\eta - 1) \ge 1. \tag{8.3.38}$$

*Then the arithmetic cost of $\mathcal{M}^*(\rho, \eta, \lambda)$ does not exceed the quantity*

$$\begin{aligned} M^{(2)} &= c(\lambda) O_\gamma(q^{-2}(\eta-1)^{-1} mn \ln(mq^{-1}\lambda^{-1}) \\ &\quad + mn^{1+\gamma} + (\eta-1)^{-1+\gamma}(\rho-1)^{-1} m^\gamma n^2), \end{aligned} \tag{8.3.39}$$

*where $c(\lambda)$ depends on $\lambda$ only.*

*In particular, under the choice*

$$\eta = 1 + m^{-(3-\gamma)/(4-\gamma)} \left(\frac{n^*}{\ln m}\right)^{1/(4-\gamma)}, \quad \rho = 1.5, \quad \lambda = 0.1, \tag{8.3.40}$$

*where*

$$n^* = \max\{n, m^{(2-\gamma)/2} \ln m\}, \tag{8.3.41}$$

*we have*

$$\begin{aligned} M^{(2)} &\le O_\gamma\left(m^{3/(4-\gamma)} (n^*)^{(7-\gamma)/(4-\gamma)} (\ln m)^{(1-\gamma)/(4-\gamma)}\right) \\ &\le O_\gamma\left(m^{(10-\gamma)/(4-\gamma)} (\ln m)^{(1-\gamma)/(4-\gamma)}\right). \end{aligned} \tag{8.3.42}$$

*Hence, under the choice* (8.3.40), *we have*

$$
\begin{aligned}
M^*(\varepsilon) &\le O_\gamma\left(M^{(2)}\ln\left(\frac{m}{\alpha(G:w)\varepsilon}\right)\right) \\
&\le O_\gamma\left(m^{(10-\gamma)/(4-\gamma)}\{\ln m\}^{(1-\gamma)/(4-\gamma)}\ln\left(\frac{m}{\alpha(G:w)\varepsilon}\right)\right).
\end{aligned} \tag{8.3.43}
$$

The proof of this theorem is quite similar to the proof of Theorem 8.3.1; the only difference is that now we use the results about the rate of convergence of the "optimal" method for smooth convex minimization (see [Ns 83], [Ns 88d]) instead of similar results for the gradient descent method.

Note that, in the case of $n = O(m)$ and of $\gamma = 1$, the optimal step in $t$ is

$$t \to (1 \pm O(m^{-1/3}))t$$

instead of $t \to (1 \pm O(m^{-3/7}))t$, as was the case in §8.3.2.

### 8.3.4 Fixed-point-based acceleration

*Henceforth, we assume that ($\mathcal{LQ}$) is an* LP *problem, i.e., that the function $\psi$ is linear.* This implies that the function $\phi$ involved into $\mathcal{P}$ also is linear.

The procedures described in §§8.3.2, 8.3.3 are based on the idea to transform an approximation of $x^*(t_k)$ into an approximation of $x^*(t_{k+1})$ for some *prescribed* $t_{k+1}$. Our new procedure $\mathcal{M}^{**}(\rho, \alpha, \lambda)$ is based on another idea: Roughly speaking, $t_{k+1}$ is as large as possible under the restriction $x^*(t_{k+1}) \in K_\alpha(x_k)$. Thus, we follow the trajectory until it leaves the domain in which $\Psi_t(x)$ is close to $\Psi_{t_k}(x_k)$. The idea of the procedure is very simple. The point $x^*(t)$ is the fixed point of the mapping

$$T_t : x \to x - Q_k \nabla F_t^\phi(x),$$

where $Q_k$ is (an arbitrary positive-definite) "scaling" matrix. If $\alpha$ is a small absolute constant and $Q_k$ is close enough to $(\nabla^2 F_t^\phi(x_k))^{-1}$, then $T_t$ proves to be a contraction (with respect to an appropriate norm $\|\cdot\|$) on the set

$$K_\alpha(x_k) : \| T(x) - T(y) \| \le \tfrac{1}{4} \| x - y \|, \quad x, y \in K_\alpha(x_k).$$

Note that here it is important that $\phi$ is linear, since otherwise $\nabla^2 F_t^\phi(x_k)$, and consequently the contracting properties of $T_t$ would depend on $t$. Now, if the fixed point $x^*(t)$ of the mapping $T_t$ belongs to a "smaller" neighbourhood $K_{\alpha'}(x_k)$ of $x_k$ ($\alpha' < \alpha$ is an appropriate absolute constant), then the sequence $u_{i+1}^t = T_t(u_i^t)$, $u_0^t = x_k$, due to contracting properties of $T_t$, does not leave the "domain of contractness" $K_\alpha(x_k)$ and converges linearly to $x^*(t)$. Since the rate of convergence is high, "a few" steps of the procedure

$$u_{i+1}^t = T_t(u_i^t)$$

allow us either to conclude that $x^*(t)$ is "close" to $K_\alpha(x_k)$ and to find a good approximation to this point, or to conclude that $x^*(t)$ is "far" from $K_\alpha(x_k)$. Combining this scheme with dichotomy in $t$, we can "quickly" (in $O(\ln^2 m)$ computations of values of $T$.) find a good approximation to $x^*(t_{k+1})$, where $t_{k+1}$ is something like the largest $t$ for which $x^*(t)$ belongs to $K_\alpha(x_k)$. Now, the computational cost of the updating $(t_k, x_k) \to (t_{k+1}, x_{k+1})$ is basically the cost of computing $Q_k$, since, given $Q_k$, we can compute a value of $T$. in $O(mn)$ operations, and it suffices to compute only $O(\ln^2 m)$ of these values; at the same time, the average cost of updating $Q_k \to Q_{k+1}$ even under the Karmarkar acceleration dominates in order the above $O(mn \ln^2 m)$. Thus, here, as in the Karmarkar-accelerated basic method, "almost all" computations are spent to update the approximate inverse Hessians. Note that our present strategy is somewhat similar to the basic one; the only difference is that in the latter strategy the polyhedral sets $K_\alpha$ are replaced by ellipsoids contained in $K_\alpha$ (these ellipsoids are defined by the Hessians of the barrier), which allows us to use a single computation of $T$. instead of $O(\ln^2 m)$ computations of $T$..

The procedure $\mathcal{M}^{**}(\rho, \alpha, \lambda)$ is as follows. The parameters of the procedure are subject to the restrictions

$$1 < \rho \le 1.1; \quad 0 < \alpha; \quad 0 < \lambda \le 0.1;$$

$$\lambda/8 \ge \alpha > 3\mu_2\omega(\lambda); \quad \rho(1+\alpha)^2 \le 1.25 \tag{8.3.44}$$

(it is clear that (8.3.44) can be satisfied by an appropriate choice of *absolute constants* $\rho, \alpha, \lambda$).

For $u \in G'$, $t > 0$, and $d \in \mathcal{S}$, let

$$Q^d = [M(\phi, \mathrm{Diag}\,\{d\})]^{-1};$$

$$\Theta_{u,t}(x) = \nabla(t\phi(x) + F_{u,q(\alpha)}(x)) : \mathbf{R}^n \to \mathbf{R}^n,$$

$$q(\alpha) = (1+\alpha)^{-1};$$

$$\Omega_{u,d,t}(x) = x - (t^{-1}Q^d)\Theta_{u,t}(x) : \mathbf{R}^n \to \mathbf{R}^n.$$

The following statement is important for us.

**Lemma 8.3.3** *Assume that* $\rho$, $\alpha$, $\lambda$ *satisfy* (8.3.44). *Let* $u \in G'$, $s, t > 0$, $d \in \mathcal{S}$ *be such that*

$$\nu(d, d(t,u)) \le \rho, \quad |\ln(s/t)| \le 0.1 \tag{8.3.45}$$

*and*

$$\lambda(F_t^\phi, u) \le \lambda. \tag{8.3.46}$$

*Then*

(i) *For given* $Q^d$, $s, u, x$ *the vector* $\Omega_{u,d,s}(x)$ *can be computed at the arithmetic cost* $O(mn)$;

(ii) *The relation*

$$\| S^{-1/2}\Omega'_{u,d,s}(x)S^{1/2} \| \leq 0.25 \quad \forall x \in \mathbf{R}^n, \tag{8.3.47}$$

*where*

$$S = s^{-1}Q^d,$$

*holds* ($\| \cdot \|$ *is the usual operator norm corresponding to the standard Euclidean structure on* $\mathbf{R}^n$);

(iii) *The implication*

$$\begin{aligned} &\left\{ |\ln(\sigma/t)| \leq m^{-1/2}(\alpha/(3\mu_2)),\ \ 1/2 \leq \sigma/t \leq 2 \right\} \\ &\Rightarrow \left\{ x^*(\sigma) \in K_{\alpha/2}(u), \quad \Omega_{u,d,\sigma}(x^*(\sigma)) = x^*(\sigma) \right\} \end{aligned} \tag{8.3.48}$$

*holds.*

**Proof.** (i) This part is evident.

(ii) By (8.3.11), we have

$$q^2(\alpha)F''(u) \leq F''_{u,q(\alpha)}(x) \leq q^{-2}(\alpha)F''(u);$$

thus, for

$$f(x) = s\phi(x) + F_{u,q(\alpha)}(x),$$

we have

$$q^2(\alpha)f''(u) \leq f''(x) \leq q^{-2}(\alpha)f''(u).$$

Furthermore, (8.3.45) implies that

$$\rho^{-1}s(Q^d)^{-1} \leq f''(u) \leq \rho s(Q^d)^{-1}.$$

Hence

$$\rho^{-1}q^2(\alpha)S^{-1} \leq f''(x) \leq \rho q^{-2}(\alpha)S^{-1}$$

or

$$\rho^{-1}q^2(\alpha)I_n \leq S^{1/2}f''(x)S^{1/2} \leq \rho q^{-2}(\alpha)I_n.$$

Since

$$I_n - S^{1/2}f''(x)S^{1/2} = S^{-1/2}\Omega'_{u,d,s}(x)S^{1/2},$$

we have

$$\| S^{-1/2}\Omega'_{u,d,s}(x)S^{1/2} \| \leq \rho q^{-2}(\alpha) - 1 = \rho(1+\alpha)^2 - 1 \leq 0.25$$

(the latter inequality holds by (8.3.44)). Part (ii) is proved.

(iii) Let $\sigma$ satisfy the premise in (8.3.48). Then, by (8.3.44), we have

$$\alpha > 2\mu_2\{m\ln^2(\sigma/t) + \omega^2(\lambda)\}^{1/2}, \qquad \tfrac{1}{2} \leq \sigma/t \leq 2,$$

whence, by Lemma 8.3.1(ii) applied with $x' = x^*(\sigma)$, $t' = \sigma$, $x = u$, we obtain

$$x^*(\sigma) \in K_{\alpha/2}(u).$$

The latter inclusion means that $\Theta_{u,\sigma}(x^*(\sigma)) = 0$ or $\Omega_{u,d,\sigma}(x^*(\sigma)) = x^*(\sigma)$. □

Now we describe the procedure $\mathcal{M}^{**}(\rho, \alpha, \lambda)$ as applied to $\mathcal{P}(\tau, y)$.
*Initialization.* Let

$$\eta = \min\{\exp\{m^{-1/2}(\alpha/(3\mu_2))\}, \exp\{0.05\}\};$$

$$\Xi = \max\{(1+\alpha)^2 - 1, \exp\{0.1\} - 1\};$$

$$N = \rfloor \ln^{-1}(16) \ln\{4\rho^2 m \alpha^2 \lambda^{-2} (1-\Xi)^{-1}\} \lfloor,$$

$$L = \rfloor \ln m \lfloor;$$

$$x_0 = y, \qquad t_0 = \tau.$$

Compute $d^0 = d(t_0, x_0)$ and the matrices

$$M_{t_0}^{\phi}(x_0), \qquad Q_0 = (M_{t_0}^{\phi})^{-1}. \tag{8.3.49}$$

*The kth step,* $k \geq 0$. Assume that, after $k-1$ steps of the procedure, we have formed a point $x_k \in G'$, a vector $d^k \in \mathcal{S}$, a number $t_k$ ($\tau \leq t_k \leq 2\tau$ at the main stage and $\tau \geq t_k \geq \tau/2$ at the preliminary stage) and the matrix

$$(\mathrm{M}_k): \qquad Q_k = (M(\phi, D_k))^{-1}, \qquad D_k \equiv \mathrm{Diag}\{d^k\}, \tag{8.3.50}$$

such that

$$(\mathrm{N}_k): \qquad \nu(d^k, d(t_k, x_k)) \leq \rho \tag{8.3.51}$$

(note that the initialization rules ensure $(\mathrm{M}_0)$, $(\mathrm{N}_0)$).

At the $k$th step, we

(a) Set

$$i = 1,$$

$$\tau_1 = \begin{cases} t_k \eta, & \text{at the main stage,} \\ t_k/\eta, & \text{at the preliminary stage,} \end{cases}$$

$$\tau^1 = \begin{cases} \tau_1 \exp\{0.05\}, & \text{at the main stage,} \\ \tau_1 \exp\{-0.05\}, & \text{at the preliminary stage,} \end{cases}$$

$$y_{0,0} = x_k; \quad y_{0,j} = \Omega_{x_k, d^k, \tau_1}(y_{0,j-1}), \quad 1 \leq j \leq N,$$

$$y^0 = y_{0,N},$$

$$\sigma^0 = \tau_1;$$

(b) Set

$$\sigma_i = \{\tau^i \tau_i\}^{1/2},$$

$$y_{i,0} = x_k; \quad y_{i,j} = \Omega_{x_k,d^k,\sigma_i}(y_{i,j-1}), \qquad 1 \le j \le N.$$

Set

$$\tau_{i+1} = \begin{cases} \sigma_i, & \text{if } y_{i,N} \in K_\alpha(x_k), \\ \tau_i, & \text{otherwise,} \end{cases}$$

$$\tau^{i+1} = \begin{cases} \tau^i, & \text{if } y_{i,N} \in K_\alpha(x_k), \\ \sigma_i, & \text{otherwise,} \end{cases}$$

$$y^i = \begin{cases} y_{i,N}, & \text{if } y_{i,N} \in K_\alpha(x_k), \\ y^{i-1}, & \text{otherwise,} \end{cases}$$

$$\sigma^i = \begin{cases} \sigma_i, & \text{if } y_{i,N} \in K_\alpha(x_k), \\ \sigma^{i-1}, & \text{otherwise.} \end{cases}$$

If $i < L$, then set $i = i + 1$ and go to (b); otherwise, set

$$x_{k+1} = y^L, \quad t_{k+1} = \sigma^L$$

and go to (c).

(c) Compute $d(t_{k+1}, x_{k+1})$ and $d^{k+1} \in \mathcal{S}$,

$$d_i^{k+1} = \begin{cases} d_i^k, & \rho^{-1} d_i^k \le d_i(t_{k+1}, x_{k+1}) \le \rho d_i^k, \\ d_i(t_{k+1}, x_{k+1}), & \text{otherwise} \end{cases}$$

(note that this updating ensures ($\mathrm{N}_{k+1}$)) and use $Q_k$ to compute $Q_{k+1}$ in accordance with ($\mathrm{M}_{k+1}$).

The $k$th step of $\mathcal{M}^{**}(\rho, \alpha, \lambda)$ is completed. If $t_{k+1}/t_0 \ge 2$ (at the main stage) or $t_{k+1}/t_0 \le \frac{1}{2}$ (at the preliminary stage), then set

$$k^* = k, \quad \tau' = t_{k+1}, \quad y' = x_{k+1}$$

and terminate; otherwise, perform the $(k+1)$th step.

**Comment.** Let $y(x, d, s)$ be $N$th point of the sequence

$$y_0 = x; \qquad y_j = \Omega_{x,d,s}(y_{j-1}), \quad 1 \le j \le N.$$

Then (a) and (b) describe the usual $L$-step dichotomy as applied to the problem

($\mathcal{L}_k$) : Given $x_k$, $t_k$, find the point $\zeta$ of the segment $[\ln(\tau_1/t_k), \ln(\tau^1/t_k)]$, nearest to $\ln(\tau^1/t_k)$, such that

$$y(x_k, d^k, s(\zeta)) \in K_\alpha(x_k), \qquad s(\zeta) = t_k \exp\{\zeta\}. \tag{8.3.52}$$

Indeed, $i$ is the number of a dichotomy step. It can be proved that (8.3.52) holds for $\zeta = \ln(\tau_1/t_k)$, i.e., for $s(\zeta) = \tau_1$. It is clear that the latter statement leads to

$$y(x_k, d^k, t_{k+1}) \in K_\alpha(x_k). \tag{8.3.53}$$

Note that the choice of $\eta$ and $\tau^1$ leads to

$$\ln \eta \le |\ln(\tau^1/t_k)|. \tag{8.3.54}$$

It is not difficult to derive from (a), (b) and (8.3.54) that either

$$|\ln(\tau^1/t_{k+1})| \le 2^{-L}$$

or

$$|\ln(t/t_{k+1})| \le 2^{-L}$$

for some $t$ such that $y(x_k, d^k, t) \notin K_\alpha(x_k)$.

Note also that (8.3.54) implies the relations

$$t_{k+1}/t_k \begin{cases} \ge \eta & \text{at the main stage,} \\ \le 1/\eta & \text{at the preliminary stage.} \end{cases} \tag{8.3.55}$$

**Theorem 8.3.3** *Each of the points $x_k$, $0 \le k \le k^*+1$, produced by $\mathcal{M}^{**}(\rho, \alpha, \lambda)$, belongs to $G'$ and satisfies the relation*

$$(\mathrm{O}_k): \qquad \lambda(F^\phi_{t_k}, x_k) \le \lambda \tag{8.3.56}$$

*and the procedure solves $\mathcal{P}$. Moreover,*

$$k^* \le O(m^{1/2}). \tag{8.3.57}$$

*The arithmetic cost of $\mathcal{M}^{**}(\rho, \alpha, \lambda)$ does not exceed the quantity*

$$\begin{aligned} &c(\rho, \alpha, \lambda, \gamma)(mn^{1+\gamma} + nm^{3/2}\ln^2 m + m^{(1+\gamma)/2}n^2) \\ &\le C(\rho, \alpha, \lambda, \gamma)m^{(5+\gamma)/2}, \end{aligned} \tag{8.3.58}$$

*where $c(\rho, \alpha, \lambda, \gamma)$, $C(\rho, \alpha, \lambda, \gamma)$ depends on $\rho$, $\alpha$, $\lambda$, $\gamma$ only.*

**Proof.** $1^0$. Let us prove $(\mathrm{O}_k)$. $(\mathrm{O}_0)$ is true in view of the initialization rules and (8.3.1). Assume that $(\mathrm{O}_k)$ holds for some $k \le k^*$ and let us prove that $(\mathrm{O}_{k+1})$ also holds. Let us fix $i$, $1 \le i \le L$ and let

$$S = \sigma_i^{-1} Q_k, \quad y_j = y_{i,j}, \quad v_j = S^{-1/2} y_j.$$

Then, for $j \ge 1$, we have

$$v_j = R(v_{j-1}), \quad R(v) = S^{-1/2}\Omega_{x_k, d^k, \sigma_i}(S^{1/2}v).$$

We have (henceforth $\| \cdot \|$ denotes the usual Euclidean norm on $\mathbf{R}^n$, as well as the standard operator norm of a linear mapping from $\mathbf{R}^n$ into itself)

$$\| R'(v) \| = \| S^{-1/2} \Omega'_{x_k, d^k, \sigma_i}(S^{1/2} v) S^{1/2}) \| \le 0.25$$

(Lemma 8.3.3 as applied with $u = x_k$, $d = d^k$, $t = t_k$, $s = \sigma_i$; the assumptions of the lemma hold in view of $(\mathrm{O}_k)$, $(\mathrm{N}_k)$ and since $|\ln(\sigma_i/t_k)| \le |\ln(\sigma_i/\tau_1)| + |\ln(\tau_1/t_k)| \le |\ln(\tau^1/\tau_1)| + \ln \eta \le 0.05 + 0.05 = 0.1$).

Thus,

$$\| R(v) - R(v') \| \le 0.25 \| v - v' \| . \tag{8.3.59}$$

In particular, there exists an unique point $v^*$, such that

$$R(v^*) = v^*,$$

and, for each $v$, we have

$$\| v - v^* \| \le \tfrac{4}{3} \| v - R(v) \| . \tag{8.3.60}$$

We also have, for $j > 0$, $\| v_j - v^* \| \le 4^{-j} \| v_0 - v^* \|$, whence, in view of $N \ge 1$,

$$\tfrac{4}{3} \| v_0 - v_N \| \ge \| v_0 - v^* \|$$

or

$$\| v_N - v^* \| \le \tfrac{4}{3} 4^{-N} \| v_0 - v_N \| . \tag{8.3.61}$$

Of course,

$$\| v_N - v^* \| \le 4^{-N} \| v_0 - v^* \| . \tag{8.3.62}$$

$2^0$. Let us prove that

$$y_N \in K_\alpha(x_k) \;\Rightarrow\; \lambda(F^\phi_{\sigma_i}, y_N) \le \lambda \tag{8.3.63}$$

and

$$x^*(\sigma_i) \in K_{\alpha/2}(x_k) \;\Rightarrow\; y_N \in K_\alpha(x_k). \tag{8.3.64}$$

Assume that the premise in (8.3.63) holds. Then we have

$$\begin{aligned} \| v_0 - v_N \|^2 &= \| S^{-1/2}(y_0 - y_N) \|^2 = (y_0 - y_N)^T S^{-1} (y_0 - y_N) \\ &= (y_0 - y_N)^T \sigma_i Q_k^{-1} (y_0 - y_N) \le \rho (y_0 - y_N)^T (F^\phi_{\sigma_i})''(x_k)(y_0 - y_N) \end{aligned}$$

(we considered $(\mathrm{M}_k)$, $(\mathrm{N}_k)$). Hence

$$\| v_0 - v_N \|^2 \le \rho \sum_{i=1}^m (f_i(x_k) - f_i(y_N))^2 f_i^{-2}(x_k) \le \rho m \alpha^2$$

(the latter relation holds since $y_N \in K_\alpha(x_k)$). Hence (8.3.61) implies that

$$\| v_N - v^* \| \le \tfrac{4}{3} 4^{-N} \rho^{1/2} m^{1/2} \alpha. \tag{8.3.65}$$

We have $R(v^*) = v^*$, or

$$\| R(v_N) - v_N \| \leq 1.34 \| v_N - v^* \| \leq 2(4^{-N}\rho^{1/2}m^{1/2}\alpha).$$

Since

$$\Omega_{x_k,d^k,\sigma_i}(y) = y - S(F^{\phi}_{\sigma_i})'(y) \quad \text{for } \; y \in K_\alpha(x_k), \tag{8.3.66}$$

our inequality implies that

$$\| S^{-1/2}\{y_N - S(F^{\phi}_{\sigma_i})'(y_N)\} - S^{-1/2}y_N \| \leq 2(4^{-N}\rho^{1/2}m^{1/2}\alpha)$$

or

$$[(F^{\phi}_{\sigma_i})'(y_N)]^T S[(F^{\phi}_{\sigma_i})'(y_N)] \leq 4^{-2N+1}\rho m\alpha^2. \tag{8.3.67}$$

We also have, for $y \in K_\alpha(x_k)$ (see (8.3.6)),

$$(F^{\phi}_{t_k})''(x_k)(1-\theta) \leq (F^{\phi}_{\sigma_i})''(y) \leq (F^{\phi}_{t_k})''(x_k)(1+\theta),$$

$$\theta = \max\{(1+\alpha)^2 - 1, |1 - \sigma_i/t_k|\} < 1$$

or, since $\rho^{-1}S^{-1} \leq (F^{\phi}_{t_k})''(x_k) \leq \rho S^{-1}$,

$$\rho^{-1}(1-\theta)S^{-1} \leq (F^{\phi}_{\sigma_i})''(y) \leq \rho(1+\theta)S^{-1}.$$

Thus,

$$[(F^{\phi}_{\sigma_i})''(y)]^{-1} \leq \rho(1-\theta)^{-1}S,$$

and (8.3.67) implies that

$$\lambda^2(F^{\phi}_{\sigma_i}, y_N) = [(F^{\phi}_{\sigma_i})'(y_N)]^T[(F^{\phi}_{\sigma_i})''(y_N)]^{-1}[(F^{\phi}_{\sigma_i})'(y_N)] \leq 4^{-2N+1}\rho^2(1-\theta)^{-1}m\alpha^2.$$

By definition of $N$ and by virtue of $\Xi \geq \theta$, the resulting inequality proves the conclusion in (8.3.63).

Now let us prove (8.3.64). Assume that

$$x^* \equiv x^*(\sigma_i) \in K_{\alpha/2}(x_k).$$

By (8.3.66), we have $\Omega_{x_k,d^k,\sigma_i}(x^*) = x^*$, or $S^{-1/2}x^* = v^*$. Thus,

$$\| v_0 - v^* \|^2 = \| S^{-1/2}(x_k - x^*) \|^2 = (x_k - x^*)^T S^{-1}(x_k - x^*)$$

$$\leq \rho(x_k - x^*)^T(F^{\phi}_{t_k})''(x_k)(x_k - x^*) = \rho\sum_{i=1}^{m}\{f_i(x_k) - f_i(x^*)\}^2 f_i^{-2}(x_k) \leq \rho m\alpha^2,$$

which, combined with (8.3.62), leads to

$$\| v_N - v^* \|^2 \leq 4^{-2N}\rho m\alpha^2 \tag{8.3.68}$$

or $(y_N - x^*)^T S^{-1}(y_N - x^*) \le 4^{-2N}\rho m\alpha^2$. The latter inequality, as above, leads to

$$(y_N - x^*)^T (F^{\phi}_{t_k})''(x_k)(y_N - x^*) \le 4^{-2N}\rho^2 m\alpha^2$$

or

$$\sum_{i=1}^{m}\{f_i(y_N) - f_i(x^*)\}^2 f_i^{-2}(x_k) \le 4^{-2N}\rho^2 m\alpha^2. \tag{8.3.69}$$

Hence

$$|f_i(x^*)/f_i(x_k) - f_i(y_N)/f_i(x_k)| \le 4^{-N}\rho m^{1/2}\alpha$$

or, by virtue of $x^* \in K_{\alpha/2}(x_k)$,

$$f_i(y_N)/f_i(x_k) \le f_i(x^*)/f_i(x_k) + 4^{-N}\rho m^{1/2}\alpha \le 1 + \alpha/2 + 4^{-N}\rho m^{1/2}\alpha \tag{8.3.70}$$

and

$$f_i(x_k)/f_i(y_N) \le \{1 - \alpha/2 - 4^{-N}\rho m^{1/2}\alpha\}^{-1}. \tag{8.3.71}$$

Relations (8.3.70) and (8.3.71), combined with (8.3.44) and the definition of $N$, imply the conclusion of (8.3.64).

$3^0$. Note that, by choice of $\eta$, we have $x^*(\tau_1) \in K_{\alpha/2}(x_k)$ (see Lemma 8.3.3(iii)). Thus, (8.3.64) means that $y^0 \in K_\alpha(x_k)$ (this was announced in comment). By virtue of the comment, we have

$$x_{k+1} = y(x_k, d^k, t_{k+1}) \in K_\alpha(x_k),$$

which, by (8.3.63), implies $(\mathrm{O}_{k+1})$.

Thus, $(\mathrm{O}_k)$ holds for all $k$, $0 \le k \le k^* + 1$.

$4^0$. $(\mathrm{O}_{k^*+1})$, combined with (8.3.55) and the termination rule, proves that the procedure solves $\mathcal{P}$.

$5^0$. It remains to evaluate the arithmetic cost of the procedure. By Lemma 8.1.2, the total number of operations *excluding computation of the matrices* $Q_k$ does not exceed

$$M' = O(mnNLK^*),\ K^* = k^* + 1 \le c_1(\rho, \alpha, \lambda)m^{1/2} \tag{8.3.72}$$

(the latter inequality holds in view of (8.3.55); henceforth, $c_i(\rho, \alpha, \lambda)$ denote quantities depending on $\rho, \alpha, \lambda$ only).

Let us evaluate the total number $M''$ of operations needed by computation of the matrices. As in the situation of Theorem 8.3.1, we have

$$\begin{aligned} M'' &\le c_2(\rho, \alpha, \lambda)O_\gamma(mn^{1+\gamma} + mK^* + (\rho - 1)^{-\gamma}(K^*)^{1-\gamma}m^\gamma n^2 + (\rho - 1)^{-1}mn^{1+\gamma}) \\ &\le c_3(\rho, \alpha, \lambda)O_\gamma(mn^{1+\gamma} + m^{3/2} + m^{(1+\gamma)/2}n^2) \end{aligned}$$

(the latter inequality holds by (8.3.72)). The resulting inequality combined with (8.3.72) completes the proof. □

## 8.4 Conjugate-gradient-based acceleration

### 8.4.1 Preliminary remarks

In this section, we present one more acceleration strategy for the basic method. Here, as in the initial method, steps in $t$ are of the form $t \to t^+ = (1 \pm O(m^{-1/2}))t$, and the updating of $x$ is $x \to x^+ = x - e(x)$, where $e(x)$ is an approximate solution to the Newton system

$$(*) \qquad \nabla^2 F^{\phi}_{t^+}(x)e = \nabla F^{\phi}_{t^+}(x).$$

The difference with the Karmarkar-type accelerated basic method is that, instead of computing $e(x)$ as

$$e(x) = Q\nabla F^{\phi}_{t^+}(x),$$

where $Q$ is a "close" approximation to the inverse Hessian

$$(**) \qquad \rho^{-1}\nabla^2 F^{\phi}_{t^+}(x) \le Q^{-1} \le \rho\nabla^2 F^{\phi}_{t^+}(x)$$

for an appropriate absolute constant $\rho$ close to 1, we solve (*) by a kind of preconditioned conjugate gradient method (PCG). The reader should be immediately aware that our PCG is not the same as what is usually called "preconditioned conjugate gradient." The standard PCG as applied to a $n \times n$ system $Ae = b$ with a symmetric $A$ makes use of a representation

$$A = A_0 + A_1,$$

where $A_0$ is positive definite and can be easily inverted, while $k \equiv \operatorname{Rank} A_1$ is $\ll n$ (a typical example: a block-diagonal matrix spoiled by a small amount of dense columns and rows). To solve such a system, we can reduce it to the system $(I + A_0^{-1/2}A_1A_0^{-1/2})u = A_0^{-1/2}b$ and then solve the latter system by the standard conjugate gradient method; if $u_i$ are the approximate solutions formed by the latter method, then $e_i = A_0^{-1/2}u_i$ are approximate solutions to the initial system, and the updatings $e_i \to e_{i+1}$ require only matrix-vector multiplications involving known vectors and the matrices $A_0$, $A_0^{-1}$, $A_1$. On the other hand, the matrix $(I + A_0^{-1/2}A_1A_0^{-1/2})$ has no more than $k$ eigenvalues different from 1, so that the standard conjugate gradient method solves the system with this matrix in no more than $k+1$ (instead of the usual $n$) steps. Thus, the essence of the usual PCG is that it solves *exactly* systems of *an appropriate structure* at the cost that is the less the better is the structure of the system. As far as $(\mathcal{LQ})$ is concerned, to apply this method, it is necessary to assume that the matrix of linear constraints possesses an appropriate structure, which generally is not the case.

We exploit another type of preconditioning. Assume that when solving $(*)$ we know a symmetric matrix $Q$, which satisfies $(**)$ with certain known $\rho$

(which, in contrast to the Karmarkar acceleration scheme, can now be large). Then we can reduce $(*)$ to the system $Au = b$, where

$$A = Q^{-1/2}\nabla^2 F^{\phi}_{t^+}(x)Q^{-1/2}, \qquad b = Q^{-1/2}\nabla F_{t^+}(x).$$

Let us solve the latter system by the standard conjugate gradient method; as above, the vectors $e_i = Q^{-1/2}u_i$, $u_i$ being the approximate solutions formed by this method, are approximate solutions to $(*)$, and the updating $e_i \to e_{i+1}$ requires only matrix-vector multiplications involving known vectors and the matrices $Q$, $Q^{-1}$, $\nabla^2 F^{\phi}_{t^+}(x)$. Now the rate of convergence of the method can be controlled, since, in view of $(**)$, the condition number of $A$ does not exceed $\rho^2$. Therefore, we can terminate the process when the desired accuracy is attained.

Of course, to use the above scheme, we should maintain relation $(**)$ along the sequence of our iterates $(t, x) = (t_i, x_i)$. This, as always, can be done with the aid of the Karmarkar scheme; an important point is that now we can use large $\rho$ instead of $\rho$ close to 1, so that we can save on the updatings of $Q$ (increasing the effort required by the conjugate gradient method). The initial Karmarkar acceleration of the basic barrier method is, in a sense, disbalanced: The average cost of updating $Q$ is $O(m^{2.5})$, and the cost of computing $e$ after $Q$ is found is only $O(m^2)$ (we mean the case of $n = O(m)$). In what follows, we balance the costs of updating $Q$ and computing $e$, which results in the lowering of the total computational effort.

### 8.4.2 Description of the method

Similarly to the basic barrier method (§3.2), the accelerated method is defined by the parameters $\lambda_1', \lambda_1, \lambda_2, \lambda_3', \lambda_3$, subject to the restrictions

$$0 < \lambda^+(\lambda_1) < \lambda_1' < \lambda_2 < \lambda_3 < \lambda_*,$$

$$\lambda_1' < \lambda_1 < \lambda_*, \qquad \lambda^+(\lambda_3) < \lambda_3' < \lambda_3; \tag{8.4.1}$$

$$\zeta(\lambda_1') \le \tfrac{1}{9}, \quad (1-\omega(\lambda_2))^{-2}\omega^2(\lambda_2) < \tfrac{1}{9},$$

$$(1-\omega(\lambda_3'))^{-2}\omega^2(\lambda_3') < 1 \tag{8.4.2}$$

(recall that $\lambda^+(\lambda) = \lambda^2/(1-\lambda)^2$, $\omega(\lambda) = 1-(1-3\lambda)^{1/3}$, $\zeta(\lambda) = \omega^2(\lambda)(1+\omega(\lambda))/(1-\omega(\lambda)))$ and by a starting point $w \in \operatorname{int} G$.

In what follows, we regard $\lambda_1', \lambda_1, \lambda_2, \lambda_3', \lambda_3$ as *absolute constants* satisfying (8.4.1) and (8.4.2).

Let

$$n^* = \max\{n, m^{(1+\gamma+\gamma^2)/(1+\gamma)}\},$$

$$\rho = 10\{(n^*)^{(1+\gamma)} m^{-(1+\gamma-\gamma^2)}\}^{2/(2+3\gamma-\gamma^2)},$$

$$M = \rfloor \{m^{\gamma/2} n^*\}^{\gamma/(2+3\gamma-\gamma^2)} \lfloor,$$

(8.4.3) $$K = \rfloor m^{1/2} M^{-1} \lfloor, \qquad L = KM.$$

Note that $n^* \le m$.

Denote

$$F(x) = \sum_{i=1}^{m} \ln(1/f_i(x)) : G' \to \mathbf{R}.$$

Then, as we know, $F$ is an $m$-self-concordant barrier for $G$.

*Structure of the method.* Similarly to the basic barrier method, the accelerated method consists of the preliminary and the main stages. Each of the stages corresponds to a set comprised of

(a) a convex quadratic form $\phi$ defined on $\mathbf{R}^n \equiv E$,
(b) a number $\kappa > 0$,
(c) an initial value, $t_0$, of the parameter $t$,
(d) an initial point $x_{-1} \in G'$,
(e) a pair of numbers $\lambda, \lambda' \in (0, \lambda_*)$.

For the preliminary stage, these data are

$$\phi(x) = DF(w)[w - x];$$

$$t_0 = 1; \qquad x_{-1} = w;$$

$$\kappa = \exp\left\{-\frac{\lambda_1 - \lambda_1'}{\lambda_1 + m^{1/2}}\right\}; \quad \lambda = \lambda_1, \quad \lambda' = \lambda_1'.$$

For the main stage, the data are

$$\phi(x) = f_0(x);$$

$$x_{-1} = u$$

($u$ is the result of the preliminary stage);

$$t_0 = \frac{\lambda_3 - \lambda(F, u)}{\| \nabla f_0(u) \|_{u,F}}$$

($\| \cdot \|_{u,F}$ is the norm in $\mathbf{R}^n$ induced by the scalar product $D^2F(u)[\cdot,\cdot]$, and $\nabla$ is the gradient with respect to the corresponding Euclidean structure);

$$\kappa = \exp\left\{\frac{\lambda_3 - \lambda_3'}{\lambda_3 + m^{1/2}}\right\}; \quad \lambda = \lambda_3, \quad \lambda' = \lambda_3'.$$

Each of the two stages corresponds to a family

$$\mathcal{F}_\phi = \{G', F_t^\phi(x) = t\phi(x) + F(x), \mathbf{R}^n\}_{t>0};$$

this family is strongly self-concordant. The metrics associated with the family are

(8.4.4) $$\rho_\nu(\mathcal{F}_\phi; t, t') = (m^{1/2} + \nu)\nu^{-1} |\ln(t/t')|$$

(Proposition 3.2.1). Moreover, it is clear that, for $x \in G'$, $t > 0$, we have

$$(8.4.5) \qquad (F_t^\phi)''(x) = tM_t^\phi(x).$$

At each stage of the accelerated method, we compute approximations, $x_i$, of the points

$$x_i^* = \operatorname{argmin}\{F_{t_i}^\phi(x) \mid x \in G'\},$$

where $t_i = t_0\kappa^i, i \geq 0$.

*Auxiliary constructions.* Let us start with some definitions. Let $h$ be a $m$-dimensional vector with positive entries $h_l$, $1 \leq l \leq m$. For $1 \leq l \leq m$, let

$$\Gamma_j(h,l) = \{s > 0 \mid \rho^{2j-1}h_l \leq s < \rho^{2j+1}h_l\}, \qquad j \in \mathbf{Z}$$

($\mathbf{Z}$ is the set of all integers). The number $h_l\rho^{2j}$ is called the *center* of the *zone* $\Gamma_j(h,l)$. For a positive vector $d \in \mathbf{R}^m$, the vector $\mathfrak{q}_{d,h}$ is defined as a vector from $\mathbf{R}^m$, which the $l$th component equals the center of that one of the zones $\{\Gamma_j(h,l) \mid j \in \mathbf{Z}\}$, which contains the number $d_l$.

To describe a stage, we should classify its iterations. Let us split the set of iteration numbers into sequential $L$-element *segments*; each of the segments is in turn split into $K$ sequential $M$-element *groups* (see (8.4.3)).

*Stage of the accelerated method.* Consider one of the two stages. At the $i$th iteration of the stage, we have positive $m$-dimensional vectors $h^{i-1}$, $d^{i-1}$, an $n \times n$ matrix

$$(8.4.6) \qquad (I_{i-1}): \qquad Q_{i-1} = [M(\phi, \operatorname{Diag}\{d^{i-1}\})]^{-1},$$

and a point $x_{i-1} \in G'$.

These quantities are updated according to the following rules (in the below description, the numbers in angle brackets are the arithmetic costs of implementation of the rules):

**I.** *Computation of $h^i$, $d^i$, $Q_i$.*

**I.1.** (a) Compute

$$(8.4.7) \qquad d^{*i} = d(t_i, x_{i-1})$$

$\langle O(mn),$ Lemma 8.1.2$\rangle$.

If $i$ is not an initial point of a segment, go to I.2.

(b) If $i$ is an initial point of a segment, set

$$(8.4.8) \qquad h^i = d^i = d^{*i}$$

and compute the matrix $M_{t_i}^\phi(x_{i-1})$ $\langle O_\gamma(mn^{1+\gamma}),$ Lemma 8.1.2$\rangle$.

Compute $Q_i$ in accordance with ($\mathrm{I}_i$) (see 8.4.6) $\langle O_\gamma(n^{\gamma+2})$, the definition of $\gamma\rangle$.

At the main stage, go to **II.**

At the preliminary stage, also compute

$$\lambda(F, x_{i-1}) = t_i^{-1}(F'(x_{i-1}))^T Q_i F'(x_{i-1})$$

(the equality holds by (8.4.7), ($\mathrm{I}_i$), (8.4.5), and (8.4.8)).

If $\lambda(F, x_{i-1}) \leq \lambda_2$, then terminate the preliminary stage and define its result as $u \equiv x_{i-1}$; otherwise, go to **II** $\langle O(mn)\rangle$

**I**.2. (a) Compute

$$d^{+i} = \mathrm{q}_{d^{*i}, h^{i-1}} \tag{8.4.9}$$

$\langle O(m)\rangle$ and use $Q_{i-1}$, $d^{+i}$, $d^{i-1}$ to compute

$$Q_i^+ = [M(\phi, \mathrm{Diag}\,\{d^{+i}\}]^{-1} \tag{8.4.10}$$

$\langle O_\gamma\,(m + l[n, k(i)])\rangle$, where

$$k(i) = |\{l \in \{1, 2, \dots, m\} \mid d_l^{+i} \neq d_l^{i-1}\}|,$$

by virtue of ($\mathrm{I}_{i-1}$) and Lemma 8.1.2 (the quantity $l[n, k]$ was defined in Lemma 8.1.2).

If $i$ is not an initial element of a group, set

$$h^i = h^{i-1}; \quad d^i = d^{+i}; \quad Q_i = Q_i^+ \tag{8.4.11}$$

(hence ($\mathrm{I}_i$) holds) and go to **II.**

(b) If $i$ is an initial element of a group, set, for $1 \leq l \leq m$, in the case of $|\ln(d_l^{*i}/h_l^{i-1})| > 1$,

$$d_l^i = h_l^i = d_l^{*i} \tag{8.4.12}$$

(henceforth, the subscripts at $d$ and $h$ denote indices of components of the vectors); in the case of $|\ln(d_l^{*i}/h_l^{i-1})| \leq 1$,

$$h_l^i = h_l^{i-1}; \qquad d_l^i = d_l^{+i} \tag{8.4.13}$$

$\langle O(m)\rangle$.

Then use $Q_i^+$, $d^{+i}$, $d^i$ to compute $Q_i$ according to ($\mathrm{I}_i$) and go to **II** $\langle O_\gamma\,(l[n, p(i)] + m)\rangle$, where

$$p(i) = |\{l \mid d_l^i \neq d_l^{+i}\}|,$$

by virtue of (8.4.10) and Lemma 8.1.2.

**II**. *Computation of* $x_i$.

**II**.1. Perform

$$N(\rho) = \rfloor \left(\frac{1}{2}\rho \ln\left\{\frac{2\lambda}{(1-\lambda)(\lambda' - \lambda^+)}\right\}\right)_+ \lfloor +1 \tag{8.4.14}$$

steps of the process

$$u_0 = 0; \qquad s_1 = r_0 = -Q_i b(t_{i+1}, x_{i-1});$$

$$u_j = u_{j-1} + \alpha_j s_j;$$
$$r_j = r_{j-1} + \alpha_j Q_i H(t_i, x_{i-1}) s_j;$$

(8.4.15) $$s_{j+1} = r_j + \beta_j s_j;$$

Herein,

$u.$, $r.$, $s.$ are vectors from $\mathbf{R}^n$;

$$\alpha_j = -\frac{r_{j-1}^T S_i r_{j-1}}{s_j^T H(t_i, x_{i-1}) s_j}; \qquad \beta_j = \frac{r_j^T S_i r_j}{r_{j-1}^T S_i r_{j-1}^T};$$

$b(t, x)$ is the gradient of $F_t^\phi$ at $x$;

$H(t, x)$ is the Hessian of $F_t^\phi$ at $x$

(the gradient and the Hessian are taken with respect to the standard Euclidean structure on $\mathbf{R}^n$);

$S_i = M(\phi, \mathrm{Diag}\,\{d^i\}) = Q_i^{-1}$.

**II**.2. Compute

(8.4.16) $$x_i = x_{i-1} - u_{N(\rho)}.$$

This completes the $i$th iteration.

*Comment to* **II**. Process (8.4.15) is the conjugate gradient method (corresponding to the metric induced by the matrix $S_i$) as applied to the equation

(8.4.17) $$H(t_i, x_{i-1})y = b(t_i, x_{i-1}).$$

It is easily seen that, under the notation

$$b = b(t_i, x_{i-1}); \quad H = H(t_i, x_{i-1}); \quad S = S_i; \quad Q = Q_i;$$

(8.4.18) $$b_* = S^{-1/2} b; \quad H_* = S^{-1/2} H S^{-1/2}; \quad z_j = S^{1/2} u_j$$

($i$ is fixed), the sequence $\{z_j\}$ is the trajectory of the standard conjugate gradient method for the minimization of the quadratic form

(8.4.19) $$\Psi(z) = z^T H_* z - 2b_*^T z,$$

(the starting point is $z_0 = 0$). Note that the computations are interrupted after $N(\rho)$ steps.

### 8.4.3 The main result

The main result is as follows.

**Theorem 8.4.1** *Let the linearly constrained quadratic programming problem* ($\mathcal{LQ}$) *satisfy the assumptions from the beginning of* §8.1.1 *and let the accelerated barrier method be applied to the problem,* $w \in G'$ *being the starting point. Then*

(i) *The amount of segments at the preliminary stage does not exceed the quantity*

$$N_1 = O(1) \ln \left( \frac{m}{1 - \pi_{x(F)}(w)} \right) \leq O(1) \ln \left( \frac{m}{\alpha(G : w)} \right)$$

*(recall that $x(F)$ is the minimizer of $F$ over $G'$, $\pi_{x(F)}$ is the Minkowsky function of $G$ with the pole at $x(F)$, and $\alpha(G:w)$ is the asymmetry coefficient of $G$ with respect to $w$).*

(ii) *For each $\varepsilon \in (0,1)$, the number $N(\varepsilon)$ of segments of the main stage that is required to find an approximate solution $x^\varepsilon \in \operatorname{int} G$ such that*

$$\psi(x^\varepsilon) - \min_G \psi \leq \varepsilon V_F(\psi),$$

*where*

$$V_F(\psi) = \max\{\psi(x) \mid x \in W_{1/2}(x(F))\} - \min\{\psi(x) \mid x \in W_{1/2}(x(F))\},$$

$$W_r(x) = \{y \mid \langle F''(x)(y-x), y-x\rangle \leq r^2\},$$

*satisfies the inequality*

$$N(\varepsilon) \leq O(\ln(m/\varepsilon)).$$

(iii) *The arithmetic cost $\mathsf{M}$ of each segment of iterations (at the preliminary and at the main stages) satisfies the inequality*

$$\mathsf{M} \leq O_\gamma(m^{r_1(\gamma)}(n^*)^{r_2(\gamma)} + mn^{1+\gamma}) \leq O_\gamma(m^{r(\gamma)}), \tag{8.4.20}$$

*where*

$$r_1(\gamma) = \frac{2+5\gamma+\gamma^2}{4+6\gamma-2\gamma^2}, \qquad r_2(\gamma) = \frac{4+5\gamma-\gamma^2}{2+3\gamma-\gamma^2},$$

*and*

$$r(\gamma) = \frac{10+15\gamma-\gamma^2}{4+6\gamma-2\gamma^2}.$$

*In particular, the total arithmetic cost $M^*(\varepsilon)$ of finding $x^\varepsilon$ satisfies the inequality*

$$\begin{aligned} M^*(\varepsilon) &\leq O_\gamma(1)\left\{m^{r_1(\gamma)}(n^*)^{r_2(\gamma)} + mn^{1+\gamma}\right\}\ln\left(\frac{2m}{\alpha(G:w)\varepsilon}\right) \\ &\leq O_\gamma(1)m^{r(\gamma)}\ln\left(\frac{2m}{\alpha(G:w)\varepsilon}\right). \end{aligned} \tag{8.4.21}$$

**Proof.** A. Let us start with the following result.

**Lemma 8.4.1** *For each iteration number $i$, we have*

$$(\mathrm{J}_i): \qquad x_{i-1} \in \operatorname{int} G; \tag{8.4.22}$$

*the matrices $S_i$ and $Q_i$ are symmetric positive-definite and*

$$(\mathrm{K}_i): \qquad t_i\rho^{-1}S_i \leq H(t_i, x_{i-1}) \leq t_i\rho S_i; \tag{8.4.23}$$

*the relations*

$$(\mathrm{L}_i): \qquad \lambda(F^\phi_{t_i}, x_{i-1}) \leq \lambda, \tag{8.4.24}$$

$$(\mathrm{M}_i): \qquad \lambda(F^\phi_{t_i}, x_i) \leq \lambda' \tag{8.4.25}$$

*hold.*

**Proof.** *Induction on i (first for the preliminary stage and then for the main stage).*

*Base*:

($J_0$) is evident for the preliminary stage. This relation holds for the main stage since such a relation holds for the iterations of the preliminary stage.

($L_0$) is evident for the preliminary stage (by definition of $\phi$ for the stage). This relation holds for the main stage by virtue of the termination rule used at the preliminary stage (see the proof of Proposition 3.2.4).

*Step*:

To complete the induction, it suffices to prove the implications

$$(J_j), j \le i \;\Rightarrow\; (K_i), \tag{8.4.26}$$

$$(K_i) \text{ and } (L_i) \;\Rightarrow\; (M_i) \text{ and } (J_{i+1}), \tag{8.4.27}$$

$$(M_i) \;\Rightarrow\; (L_{i+1}). \tag{8.4.28}$$

Relation (8.4.28) holds by Theorem 3.1.1, (8.4.4), and the definition of $\kappa$ (see the proofs of Propositions 3.2.3 and 3.2.4).

Let us verify (8.4.26). We have

$$S_i = M(\phi, \mathrm{Diag}\,\{d^i\}), \qquad H_i \equiv H(t_i, x_{i-1}) = t_i M(\phi, \mathrm{Diag}\,\{d^{*i}\}).$$

In the case of (8.4.9), we have $d^{*i} = d^i$; in each of the cases (8.4.11)–(8.4.13), we have

$$\rho^{-1} d^{*i} \le d^i \le \rho d^{*i}.$$

These inequalities, combined with the definition of $S_i$, imply ($K_i$).

To prove (8.4.27), let us fix $i$. In what follows, we use notation (8.4.18) and the abbreviations

$$t = t_i, \quad x = x_{i-1}, \quad F_t = F^{\phi}_{t_i}.$$

Note that $F_t \in S_1^+(G', \mathbf{R}^n)$ (by Proposition 3.2.2; recall that $\phi$ is a quadratic form and therefore is 0-compatible with $F$).

$1^0$. Let $z_* = H_*^{-1} b_*$. By the standard properties of the conjugate gradient method, we have $z_j H_*(z_* - z_j) = 0$ for each $j$ (recall that $z_0 = 0$), so that

$$z_j^T H_* z_j = z_*^T H_* z_* - (z_* - z_j)^T H_* (z_* - z_j). \tag{8.4.29}$$

We also have $z_*^T H_* z_* = u_* H u_*$ ($u_*$ is the solution to (8.4.17)), or

$$z_*^T H_* z_* = \lambda^2(F_t, x) \le \lambda^2 < \tfrac{1}{9}$$

(we considered ($L_i$) and the relation

$$\lambda^2(F_t, x) = (F_t'(x))^T [F_t''(x)]^{-1} F_t'(x) = b^T H^{-1} b = u_*^T H u_*).$$

Thus, (8.4.29), combined with the equality $z_j^T H_* z_j = u_j^T H u_j$, implies that

$$u_j^T H u_j \le z_*^T H_* z_* \le \lambda^2 < \tfrac{1}{9}, \tag{8.4.30}$$

so that (see Theorem 2.1.1(ii))

(8.4.31) $$x - u_j \in G', \qquad j \geq 0,$$

and, in particular, $x_i \in G'$, as required in ($J_{i+1}$).

$2^0$. Let

$$\varepsilon_j = \left\{(z_* - z_j)^T H_*(z_* - z_j)\right\}^{1/2} \quad \left(= \left\{(u_* - u_j)^T H(u_* - u_j)\right\}^{1/2}\right)$$

and let

$$g_j(\tau) = H^{-1/2} F_t'(x - \tau u_j)$$

for $0 \leq \tau \leq 1$. Then

$$g_j'(\tau) + H^{1/2} u_j = \left\{H^{-1/2}\left(F_t''(x) - F_t''(x - \tau u_j)\right) H^{-1/2}\right\} \{H^{1/2} u_j\},$$

whence (Theorem 2.1.1 combined with (8.4.30) and the inclusion $F_t \in S_1^+(G', \mathbf{R}^n)$)

$$\| g_j'(\tau) + H^{1/2} u_j \|_2 \leq \lambda((1 - \tau\lambda)^{-2} - 1), \qquad 0 \leq \tau \leq 1,$$

which leads to

$$\| H^{-1/2} F_t'(x - u_j) - H^{-1/2} F_t'(x) + H^{1/2} u_j \|_2 \leq \lambda^2/(1 - \lambda).$$

The latter relation, in view of $F_t'(x) = H u_*$, implies that

(8.4.32) $$\| H^{-1/2} F_t'(x - u_j) \|_2 \leq \frac{\lambda^2}{1 - \lambda} + \| H^{1/2}(u_* - u_j) \|_2 = \lambda^2/(1 - \lambda) + \varepsilon_j.$$

By Theorem 2.1.1 and (8.4.30), we have

$$(1 - \lambda)^2 H \leq F_t''(x - u_j) \leq (1 - \lambda)^{-2} H,$$

which, combined with (8.4.32), implies that

$$\| (F_t''(x - u_j))^{-1/2} F_t'(x - u_j) \|_2 \leq \frac{\lambda^2}{(1 - \lambda)^2} + \frac{\varepsilon_j}{1 - \lambda}$$

or

(8.4.33) $$\lambda(F_t, x - u_j) \leq \lambda^2/(1 - \lambda)^2 + \varepsilon_j/(1 - \lambda).$$

$3^0$. Let $\mathsf{P}_j$ be the space of real polynomials on the axis of degree less than $j$. Then

$$z_j = p_j(H_*) H_* z_*,$$

where

$$p_j \in \operatorname{Argmin}\{(p(H_*)H_*z_*)^T H_*(p(H_*)H_*z_*) - 2z_*^T H_* p(H_*)H_*z_* \mid p \in \mathsf{P}_j\},$$

or, which is the same,

$$p_j \in \operatorname{Argmin}\{\| H_*^{1/2}(I_n - H_* p(H_*))z_* \|_2^2 \mid p \in \mathsf{P}_j\}.$$

The latter relation implies that

$$\forall p \in \mathsf{P}_j: \quad \varepsilon_j^2 = (z_* - z_j)^T H_*(z_* - z_j) \le \| H_*^{1/2}(I_n - H_* p(H_*))z_* \|_2^2 .$$

Thus, for each $p \in \mathsf{P}_j$ we have

$$\varepsilon_j^2 \le \max_{\tau\in\Sigma}{}^2 |1 - \tau p(\tau)| \; \| H_*^{1/2} z_* \|_2^2,$$

where $\Sigma$ is the spectrum of the matrix $H_*$. By (8.4.30), we have

$$\varepsilon_j \le \lambda \max_{\tau\in\Sigma} |1 - \tau p(\tau)| \quad \forall p \in \mathsf{P}_j. \tag{8.4.34}$$

Let $q_j \in \mathsf{P}_j$ be such that

$$1 - \tau q_j(\tau) = T_j\left(\frac{\rho - 2\tau t_i^{-1} + \rho^{-1}}{\rho - \rho^{-1}}\right)\left\{T_j\left(\frac{\rho + \rho^{-1}}{\rho - \rho^{-1}}\right)\right\}^{-1},$$

where $T_j(s) = \mathrm{ch}(j\, \mathrm{ch}^{-1}(s))$ is the Tchebyshev polynomial of the degree $j$. For $p \equiv q_j$, (8.4.34) leads to

$$\varepsilon_j \le \lambda T_j^{-1}\left(\frac{\rho + \rho^{-1}}{\rho - \rho^{-1}}\right)$$

(we considered $(\mathrm{K}_i)$), which immediately implies that

$$\varepsilon_j \le 2\lambda \exp\{-2j/\rho\}, \qquad j \ge 1.$$

The latter inequality, combined with (8.4.33) and the definition of $N(\rho)$, leads to the relation

$$\lambda(F_t, x - u_{N(\rho)}) \le \lambda',$$

which is required in $(\mathrm{M}_i)$. $\square$

B. It is not difficult to conclude from (8.4.3) that, if

$$M^* = \left\{m^{\gamma/2} n^*\right\}^{\gamma/(2+3\gamma-\gamma^2)},$$

$$K^* = m^{1/2}(M^*)^{-1},$$

$$L^* = K^* M^* = m^{1/2},$$

then

$$M = O(M^*), \quad K = O(K^*), \quad L = O(L^*). \tag{8.4.35}$$

Relation (8.4.35) immediately implies that, if the numbers $i$, $i'$ belong to a common segment of iterations, then

$$|\ln(t_i/t_{i'})| \le O(1), \tag{8.4.36}$$

and, if these numbers belong to a common group of iterations, then

$$|\ln(t_i/t_{i'})| \le O(\Delta), \qquad \Delta = 1/K^* \in [O(m^{-1/2}), O(1)]. \tag{8.4.37}$$

The results of Lemma 8.4.1 and relation (8.4.36) prove statements of Theorems 8.4.1(i) and 8.4.1(ii) (compare with the proofs of Propositions 3.2.3 and 3.2.4).

C. It remains to prove Theorem 8.4.1(iii). Note that (8.4.21) is an immediate consequence of the preceding statements of the theorem, so that it suffices to prove (8.4.20).

$1^0$. Let $T = \{t_i \mid i \ge 0\}$. For $t = t_i$, we write $x(t)$ instead of $x_{i-1}$ and $F_t$ instead of $F_t^\phi$.

Also, let

$$\psi_i(t) = t^{1/2} f_i(x(t)), \qquad t \in T,$$

$$\psi_{*,i}(t) = t^{1/2} f_i(x_*(t)),$$

$$x_*(t) = \operatorname{argmin}\{F_t(x) \mid x \in G'\}, \qquad t > 0.$$

By Corollary 8.2.1, we have

$$\begin{aligned}(\forall t, t' \in T) : &\sum_{i=1}^{m} (\psi_i(t) - \psi_i(t'))^2 (\psi_i(t)\psi_i(t'))^{-1} \\ &\le O\left(m(t^{1/2} - (t')^{1/2})^2 (tt')^{-1/2}\right).\end{aligned} \tag{8.4.38}$$

$2^0$. Let us fix a segment of iterations and let $I$ be the corresponding set of values of the iteration number $i$. Denote the sets of values of $i$ for the groups of the segment $I$ by $J_1, \dots, J_K$, respectively. The remarks on the arithmetic cost of the rules forming the method (see the description of the method) imply that

$$\begin{aligned}\mathsf{M} \le\ & O_\gamma\left(mn^{1+\gamma} + mnN(\rho)L\right) \\ &+ \sum_{j=1}^{K} \sum_{i \in J_j} O_\gamma(l[n, k(i)]) + \sum_{j=1}^{K} \sum_{i \in J_j} O_\gamma(l[n, p(i)]).\end{aligned} \tag{8.4.39}$$

In view of rules **I**.1 and **I**.2, we have

$$k(i) = \begin{cases} 0, & i \text{ starts a segment}, \\ |U(i)|, \ U(i) = \{l \mid 1 \le l \le m, \ d_l^{+i} \ne d_l^{i-1}\}, & \text{otherwise}; \end{cases} \tag{8.4.40}$$

$$(8.4.41)\quad p(i) = \begin{cases} 0, & i \text{ does not start a group or starts the group } J_1, \\ |V(i)|, \quad V(i) = \{l \mid 1 \le l \le m,\ d^i_l \ne d^{+i}_l\}, & \text{otherwise.} \end{cases}$$

$3^0$. First, let us evaluate the numbers $k(i)$, $i \in I$. Let $i(q)$ be the initial element of the group $J_q$, $1 \le q \le K$. From the description of the method, it is clear that

$$(8.4.42)\qquad h^{i(q)} = h^{i(q)+1} = \cdots = h^{i(q)+M-1};$$

$$(8.4.43)\qquad d^{+i} = \mathsf{q}_{d^{*i},h^{i(q)}}, \quad i(q)+1 \le i \le i(q)+M, \quad i \in I;$$

$$(8.4.44)\qquad d^{i} = \mathsf{q}_{d^{*i},h^{i(q)}}, \qquad i(q) \le i < i(q)+M.$$

Let $I_q = \{i \in I \mid i(q) \le i \le i(q)+M\}$. By(8.4.7) and by the definition of $\psi_l(t)$, we have

$$d^{*i} = d(t_i, x_{i-1}) = (\psi_1^{-2}(t_i), \ldots, \psi_m^{-2}(t_i))^T.$$

Thus, in view of (8.4.38), (8.4.35), we have

$$(8.4.45)\quad (\mathrm{P}_i): \qquad \sum_{i=1}^{m} \left\{(d^{*i}_l)^{1/2} - (d^{*i(q)}_l)^{1/2}\right\}^2 \left\{d^{*i}_l d^{*i(q)}_l\right\}^{-1/2} \le O_\gamma(m\Delta^2)$$

for $1 \le q \le K$ and $i \in I_q$.

Let us fix $q$. We call a pair $(i,l) \in I_q \times \{1,2,\ldots,m\}$ *an event*, if the number $d^{*i}_l$ does not belong to the zone $\Gamma_0(h^{i(q)}, l)$. Let us prove that, if $i \in I_q \backslash \{i(q)\} \equiv I^0_q$ and $l \in U(i)$, then either $(i,l)$ or $(i-1,l)$ is an event. Indeed, if, for some $i \in I^0_q$ and $l$, neither $(i,l)$ nor $(i-1,l)$ is an event, then $d^{*i-1}_l \in \Gamma_0(h^{i(q)}, l)$, which, by (8.4.44), implies that $d^{i-1}_l = h^{i(q)}_l$. Since $d^{*i}_l \in \Gamma_0(h^{i(q)}, l)$, (8.4.43) implies that $d^{+i}_l = h^{i(q)}_l$. Thus, $d^{+i}_l = d^{i-1}_l$, and $l \notin U(i)$ (see (8.4.40)). □

The above arguments mean that, for $i \in I^0_q$, the quantity $k(i)$ does not exceed the total number of events of the form $(i,l)$ or $(i-1,l)$. If $(i,l)$ is an event, then

$$\left|\ln\left(\frac{d^{*i}_l}{h^{i(q)}_l}\right)\right| \ge \ln(\rho),$$

while (8.4.12), (8.4.13) imply that

$$\left|\ln\left(\frac{d^{*i(q)}_l}{h^{i(q)}_l}\right)\right| \le 1.$$

Thus,

$$\left|\ln\left(\frac{d^{*i}_l}{d^{*i(q)}_l}\right)\right| \ge \ln\frac{\rho}{\mathrm{e}} \ge 1$$

(we considered that $\rho \ge 10$). The latter inequality means that the term in $(\mathrm{P}_i)$ corresponding to the value of $l$ under consideration is $> O(\rho^{1/2})$. Thus, the number of events of the form $(i, l)$ does not exceed $O_\gamma(m\Delta^2\rho^{-1/2})$. The number of events of the form $(i-1, l)$ admits a similar upper bound. Thus,

$$k(i) \leq O_\gamma(m\Delta^2\rho^{-1/2}), \quad i \in I_q^0, \quad 1 \leq q \leq K.$$

Since $\bigcup_{q=1}^K = I\backslash\{i^*\}$ ($i^*$ is the initial element of the segment $I$) and $k(i^*) = 0$ (see (8.4.40)), we obtain

$$\sum_{j=1}^K \sum_{i \in J_j} O_\gamma\left(l[n, k(i)]\right) \leq O_\gamma\left((n^*)^2 m^\gamma \Delta^{2\gamma}\rho^{-\gamma/2} L^*\right) \tag{8.4.46}$$

(note that $m\Delta^2\rho^{-1/2} \leq n^*$ by virtue of (8.4.3), so that $l[n, k(i)] \leq (n^*)^2 k^\gamma(i)$).

$4^0$. Now let us evaluate the latter sum in the right-hand side of (8.4.39). The sum is of the form

$$S = \sum_{j=1}^K \sum_{i \in J_j} O_\gamma\left(l[n, p(i)]\right).$$

Note that, by (8.4.41), we have

$$S = \sum_{q=2}^K O_\gamma\left(l[n, p(i(q))]\right) \leq O_\gamma\left(n^2 P^\gamma K^{1-\gamma}\right) + O_\gamma\left(n^{1+\gamma} P\right),$$

$$P = \sum_{q=2}^K p(i(q)). \tag{8.4.47}$$

For $1 \leq q \leq K$, let $s^q$ be $m$-dimensional vectors with the components $\ln(h_l^{i(q)})$; let $r^q$ be vectors with the components $\ln(d_l^{*i(q)})$, $1 \leq l \leq m$. From the description of the method, it is clear that the evolution of these vectors looks as follows:

$$s^1 = r^1; \tag{8.4.48}$$

$$q > 1 \Rightarrow s_l^q = \begin{cases} s_l^{q-1}, & |s_l^{q-1} - r_l^q| \leq 1; \\ r_l^q, & \text{otherwise.} \end{cases} \tag{8.4.49}$$

Let us prove that, for $q \geq 2$, we have

$$p(i(q)) \leq |V^*(q)|,$$

$$V^*(q) = \{l \in \{1, \ldots, m\} \mid |s_l^{q-1} - r_l^q| > 1\}. \tag{8.4.50}$$

Indeed, if $l \notin V^*(q)$, then, in view of $h^{i(q)-1} = h^{i(q-1)}$ ($h^\cdot$ is constant at the iterations of a group) and, by **I**.2.(b), we have

$$|\ln(h_l^{i(q)-1}/d_l^{*i(q)})| \leq 1.$$

Thus, by (8.4.14), $d_l^{i(q)} = d_l^{+i(q)}$. Therefore $l \notin V(q)$. Hence $V(q) \subset V^*(q)$, which, by virtue of (8.4.41), proves (8.4.50).

It is clear that

$$|\ln(s/s')|^2 \le O\left(\{s^{1/2} - (s')^{1/2}\}^2 \{ss'\}^{-1/2}\right), \qquad s, s' > 0.$$

As in $3^0$, we have

$$\sum_{l=1}^{m} \left\{(d_l^{*i(q-1)})^{1/2} - (d_l^{*i(q)})^{1/2}\right\}^2 \left\{d_l^{*i(q-1)} d_l^{*i(q)}\right\}^{-1/2} \le O_\gamma(m\Delta^2),$$

which implies that

$$\| r^q - r^{q-1} \|_2^2 \le O_\gamma(m\Delta^2)$$

or $\| r^q - r^{q-1} \|_1 \le O_\gamma(m\Delta)$. Since (8.4.48)–(8.4.50) imply that

$$\sum_{q=2}^{K} p(i(q)) \le \sum_{q=2}^{K} \| r^q - r^{q-1} \|_1,$$

we obtain

$$P \le O_\gamma(m\Delta K) \le O_\gamma(m). \tag{8.4.51}$$

Relations (8.4.51), (8.4.46), (8.4.47), (8.4.39), combined with (8.4.35) and (8.4.3), prove (8.4.20). □

*This page intentionally left blank*

# Appendix 1

**Proposition 9.1.1** *Let $H[x_1, \dots, x_k]$ be a symmetric $k$-linear form on $\mathbf{R}^n$ and $A[x, y]$ be a positive-semidefinite quadratic form on $\mathbf{R}^n$. If*

$$\forall x \in \mathbf{R}^n : \quad |H[x, \dots, x]| \leq \{A[x, x]\}^{k/2}, \tag{9.1.1}$$

*then*

$$\forall (x_1, \dots, x_k \in \mathbf{R}^n) : \quad |H[x_1, \dots, x_k]| \leq \prod_{i=1}^{k} \{A[x_i, x_i]\}^{1/2}. \tag{9.1.2}$$

**Proof.** Of course, it suffices to prove the proposition in the case of positive-definite $A$ (if it is not the case from the very beginning, we can replace $A[x, x]$ by $A_\varepsilon[x, x] = A[x, x] + \varepsilon x^T x$ and then pass to limit as $\varepsilon \to +0$). Thus, we henceforth assume that $A$ is positive-definite.

The proof is given by induction on $k$. *Base* $k = 2$ is evident. Assume that the statement holds for a given $k = l - 1 \geq 2$ and let us prove that it holds for $k = l$.

$1^0$. Consider $A[x, y]$ as a scalar product (we henceforth write $(x, y)$ instead of $A[x, y]$ and $\| x \|$ instead of $(A[x, x])^{1/2}$). Since both sides in (9.1.2) are homogeneous in each $x_i$, it suffices to prove that

$$\omega \equiv \sup\{|H[x_1, \dots, x_l]| \mid \| x_i \| \leq 1, 1 \leq i \leq l\} \leq 1. \tag{9.1.3}$$

$2^0$. Let us call a collection $\mathcal{T} = \{T_1, \dots, T_l\}$ of one-dimensional subspaces of $\mathbf{R}^n$ *an extremal,* if for some (and then for each) choice of *unit* vectors $e_i \in T_i$, we have $|H[e_1, \dots, e_l]| = \omega$. Clearly, extremals exist; let $\mathsf{T}$ be the set of all extremals. Proving (9.1.3) is the same as proving that $\mathsf{T}$ contains an extremal of the type $\{T, \dots, T\}$.

**Lemma 9.1.2** *Let $\{T_1, \dots, T_l\} \in \mathsf{T}$ and $T_1 \neq T_2$. Let $e_i \in T_i$ be unit vectors, $h = e_1 + e_2$, $q = e_1 - e_2$. Then $\{\mathbf{R}h, \mathbf{R}h, T_3, \dots, T_l\} \in \mathsf{T}$ and $\{\mathbf{R}q, \mathbf{R}q, T_3, \dots, T_l\} \in \mathsf{T}$.*

**Proof.** First, $e_1$ and $e_2$ are linearly independent since $T_1 \neq T_2$; therefore $h \neq 0$, $q \neq 0$. Let $(Qx, y) = H[x, y, e_3, \dots, e_l]$; then $Q$ is a symmetric matrix.

Since $\{T_1, \dots, T_l\}$ is an extremal, we have

$$\omega = |(Qe_1, e_2)| = \max\{|(Qu, v)| \mid \| u \|, \| v \| \leq 1\}.$$

Therefore, if $E^+ = \{x \in \mathbf{R}^n \mid Qx = \omega x\}$, $E^- = \{x \in \mathbf{R}^n \mid Qx = -\omega x\}$ and $E = (E^+ + E^-)^\perp$, then at least one of the subspaces $E^+, E^-$ is nonzero, $\| Qx \| \leq \omega' \| x \|$, $x \in E$, where $\omega' < \omega$. $\mathbf{R}^n$ is the direct sum of $E^+, E^-$, and $E$. Let $x = x^+ + x^- + x'$ be the decomposition of $x \in \mathbf{R}^n$ corresponding to the decomposition $\mathbf{R}^n = E^+ + E^- + E$. Since each of the subspaces $E^+$, $E^-$, and $E$ is invariant for $Q$,

$$\begin{aligned}
\omega = |(Qe_1, e_2)| &\leq |\omega(e_1^+, e_2^+) - \omega(e_1^-, e_2^-)| + \omega' \| e_1' \| \| e_2' \| \\
&\leq \omega(\| e_1^+ \| \| e_2^+ \| + \| e_1^- \| \| e_2^- \|) + \omega' \| e_1' \| \| e_2' \| \\
&\leq \omega\{\| e_1^+ \|^2 + \| e_2^+ \|^2\}^{1/2}\{\| e_1^- \|^2 + \| e_2^- \|^2\}^{1/2} + \omega' \| e_1' \| \| e_2' \| \leq \omega
\end{aligned}$$

(we considered that $\| e_i^+ \|^2 + \| e_i^- \|^2 + \| e_i' \|^2 = 1$, $i = 1, 2$). We see that all the inequalities in the above chain are equalities. Therefore we have

$$\| e_1' \| = \| e_2' \| = 0; \quad \| e_1^+ \| = \| e_2^+ \|; \quad \| e_1^- \| = \| e_2^- \|;$$

moreover, $|(e_1^+, e_2^+)| = \| e_1^+ \| \| e_2^+ \|$ and $|(e_1^-, e_2^-)| = \| e_1^- \| \| e_2^- \|$, which means that $e_1^+ = \pm e_2^+$ and $e_1^- = \pm e_2^-$. Since $e_1$ and $e_2$ are linearly independent, only two cases are possible:

(a) $e_1^+ = e_2^+ \neq 0$, $e_1^- = -e_2^- \neq 0$, $e_1' = e_2' = 0$;

(b) $e_1^+ = -e_2^+ \neq 0$, $e_1^- = e_2^- \neq 0$, $e_1' = e_2' = 0$.

In case (a), $h$ is proportional to $e_1^+$, and $q$ is proportional to $e_1^-$; therefore

$$\{\mathbf{R}h, \mathbf{R}h, T_3, \ldots, T_l\} \in \mathsf{T}$$

and

$$\{\mathbf{R}q, \mathbf{R}q, T_3, \ldots T_l\} \in \mathsf{T}.$$

The same arguments can be used in case (b). □

$3^0$. Let $\mathsf{T}^*$ be the subset of $\mathsf{T}$ formed by the extremals of the type

$$\{\overbrace{T, \ldots, T}^{t \text{ times}}, \overbrace{S, \ldots, S}^{s \text{ times}}\}$$

for some $t$ and $s$ (depending on the extremal). By virtue of the inductive assumption, $\mathsf{T}^*$ is nonempty (in fact, $\mathsf{T}^*$ contains an extremal of the type $\{T, \ldots, T, S\}$). For

$$\mathcal{T} = \{\overbrace{T, \ldots, T}^{t \text{ times}}, \overbrace{S, \ldots, S}^{s \text{ times}}\} \in \mathsf{T}^*,$$

let $\alpha(\mathcal{T})$ denote the angle between $T$ and $S$ and let $e \in T$ and $f \in S$ be unit vectors with the angle between them being equal to $\alpha(\mathcal{T})$. Without loss of generality, we can assume that $t \leq s$ (note that reordering of an extremal leads to an extremal, since $H$ is symmetric). By virtue of Lemma 9.1.2, in the case of $\alpha(\mathcal{T}) \neq 0$, the collection

$$\mathcal{T}' = \{\overbrace{\mathbf{R}(e+f), \ldots, \mathbf{R}(e+f)}^{2t \text{ times}}, \overbrace{S, \ldots, S}^{s-t \text{ times}}\}$$

belongs to $\mathsf{T}^*$, and clearly $\alpha(\mathcal{T}') = \alpha(\mathcal{T})/2$. Thus, either $\mathsf{T}^*$ contains an extremal $\mathcal{T}$ with $\alpha(\mathcal{T}) = 0$, or we can find a sequence $\{\mathcal{T}_i \in \mathsf{T}^*\}$ with $\alpha(\mathcal{T}_i) \to 0$. In the latter case, the sequence $\{\mathcal{T}_i\}$ contains a subsequence converging (in the natural sense) to certain collection $\mathcal{T}$, which clearly belongs to $\mathsf{T}^*$, and $\alpha(\mathcal{T}) = 0$. Thus, $\mathsf{T}$ contains an extremal $\mathcal{T}$ with $\alpha(\mathcal{T}) = 0$, or, which is the same, an extremal of the type $\{T, \ldots, T\}$. $\square$

*This page intentionally left blank*

# Appendix 2

## Notation

We henceforth regard $\mathbf{R}^n$ as an Euclidean space with the standard scalar product $\langle \cdot, \cdot \rangle$, so that the space is identified with its conjugate space. We use the following notation: If a lowercase letter (like $a$) denotes a vector from $\mathbf{R}^n$, then the same uppercase letter (like $A$) denotes the diagonal matrix with the diagonal elements being equal to the corresponding entries of $a$. $e$ denotes the $n$-dimensional vector of ones (and, consequently, $E$ - the unit $n \times n$ matrix). If $x$ is a (nonnegative) vector and $a$ is a vector, then $X^a$ denotes the diagonal matrix with the diagonal entries $x_i^{a_i}$. The nonnegative $n$-dimensional orthant is denoted by $\mathbf{R}^n_+$, and its interior is denoted by $\mathbf{R}^n_{++}$.

Let $a = (a_1, \dots, a_n)^T$ be a fixed vector satisfying $0 < a_i < 1, \;\; i = 1, \dots, n$ and let $c = (c_1, \dots, c_n)^T$ be a fixed positive vector. These data define the concave function

$$f(x) \equiv f_{a,c}(x) = \sum_{i=1}^{n} c_i x_i^{a_i} : \mathbf{R}^n_+ \to \mathbf{R}; \tag{10.1.1}$$

in the above notation,

$$f(x) = \langle C X^a e, e \rangle . \tag{10.1.2}$$

We study the mapping

$$p(x) = \lambda(x) f'(x) : \mathbf{R}^n_{++} \to \mathbf{R}^n_{++}, \; \lambda(x) = \langle f'(x), x \rangle^{-1} . \tag{10.1.3}$$

Note that $x > 0$ is (clearly, the unique) solution to the concave program

$$\text{maximize } \; f(u) \quad \text{subject to } \; u \in \mathbf{R}^n_+, \; \langle u, p(x) \rangle \le 1;$$

it follows that $x \to p(x)$ is a one-to-one mapping of $\mathbf{R}^n_{++}$ onto itself, so that the inverse mapping $x(p) : \mathbf{R}^n_{++} \to \mathbf{R}^n_{++}$ is well defined,

$$p(x(u)) \equiv u, \qquad u \in \mathbf{R}^n_{++}. \tag{10.1.4}$$

Note also that $p(\cdot)$ is $C^\infty$ smooth.

## Antimonotonicity of $p(\cdot)$

The fact that $-p(\cdot)$ is monotone was proved in [PM 78]. The following statement establishes some useful estimates of $p'$.

**Lemma 10.1.3** *$-p(\cdot)$ is monotone, $p'(\cdot)$ is nonsingular, and*

$$\begin{aligned}(10.1.5)\qquad &\tfrac{1}{2}\lambda(x)\,\langle -f''(x)h,h\rangle \le \langle -p'(x)h,h\rangle \\ &\le \tfrac{3}{2}(1-a_{\max})^{-1}\lambda(x)\,\langle -f''(x)h,h\rangle\,,\end{aligned}$$

$$\begin{aligned}(10.1.6)\qquad &\tfrac{1}{2}(1-a_{\max})\lambda^{-1}(x)\,\langle[-f''(x)]^{-1}h,h\rangle \le \langle[-p'(x)]^{-1}h,h\rangle \\ &\le \tfrac{3}{2}\lambda^{-1}(x)\,\langle[-f''(x)]^{-1}h,h\rangle\,,\end{aligned}$$

*where*

$$a_{\max} = \max\{a_1,\ldots,a_n\}.$$

**Proof.** We have (everything is taken at $x$)

$$(10.1.7)\qquad p' = \lambda f'' - \lambda^2\{(f')(f')^T + ((f')x^T)f''\}.$$

$f''$ is symmetric negative defined diagonal matrix, so that there exists a nonsingular diagonal matrix $Q = Q(x)$ with

$$(10.1.8)\qquad f'' = -Q^{-2},$$

whence

$$(10.1.9)\qquad Q^{-1} = C^{1/2}A^{1/2}(E-A)^{1/2}X^{a/2-e}.$$

Let $W = Qp'Q$. Then (see (10.1.7))

$$(10.1.10)\qquad W = -\lambda\left(E + \lambda\{(Qf')(Qf')^T - (Qf')(Q^{-1}x)^T\}\right).$$

Denote

$$(10.1.11)\qquad \phi = \lambda^{1/2}Qf', \qquad \xi = \lambda^{1/2}Q^{-1}x,$$

so that

$$W = -\lambda U, \qquad U = E + \{\phi\phi^T - \phi\xi^T\},$$

$$(10.1.12)\qquad \langle\phi,\xi\rangle = \lambda\,\langle f',x\rangle = 1.$$

Let $z$, $w$ be orthogonal to each other unit vectors such that $\phi = lz$ and $\xi$ is a linear combination of $z$ and $w$. By virtue of the last relation in (10.1.12), we have $\xi = l^{-1}z + mw$. Let $L$ be the subspace generated by $z$ and $w$; then $L$ and $L^\perp$ are invariant with respect to $W$, $W$ is the identity on $L^\perp$, and the restriction of $W$ onto $L$ in the orthogonal basis $z$, $w$ is the matrix

$$Y = \begin{pmatrix} l^2 & -lm \\ 0 & 1 \end{pmatrix},$$

so that

$$Y^{-1} = l^{-2}\begin{pmatrix} 1 & lm \\ 0 & l^2 \end{pmatrix}.$$

For a two-dimensional vector $r = (s,t)^T$ we have $r^T Y r = l^2 s^2 - lmst + t^2$. Let us prove that $(lm)^2 \le l^2$, or, which is the same, $m^2 \le 1$. Indeed,

$$\begin{aligned} m^2 \le \xi^T \xi = \lambda \langle Q^{-1}x, Q^{-1}x\rangle &= \langle f'(x), x\rangle^{-1} \langle -f''x, x\rangle \\ &= \left\langle CA(E-A)X^{a-2e}x, x\right\rangle \langle CAX^{a-e}e, x\rangle^{-1} \\ &= \langle CA(E-A)X^a e, e\rangle \langle CAX^a e, e\rangle^{-1} \le 1. \end{aligned}$$

We conclude that

$$\tfrac{3}{2}(l^2s^2+t^2) \ge r^T Y r \ge \tfrac{1}{2}(l^2s^2+t^2).$$

Now,

$$\begin{aligned} l^2 &= \lambda\langle Qf', Qf'\rangle = -\langle (f'')^{-1}f', f'\rangle \langle f', x\rangle^{-1} \\ &= \langle C^{-1}A^{-1}(E-A)^{-1}X^{2e-a}CAX^{a-e}e, CAX^{a-e}e\rangle \langle CAX^{a-e}e, x\rangle^{-1} \\ &= \langle CA(E-A)^{-1}X^a e, e\rangle \langle CAX^a e, e\rangle^{-1}, \end{aligned}$$

so that

$$(1-a_{\max})^{-1} \ge l^2 \ge 1.$$

Thus,

$$\tfrac{1}{2}\langle h,h\rangle \le \langle Uh,h\rangle \le \tfrac{3}{2}(1-a_{\max})^{-1}\langle h,h\rangle. \tag{10.1.13}$$

Similarly,

$$\tfrac{3}{2}l^{-2}(l^2t^2+s^2) \ge r^T Y^{-1} r \ge \tfrac{1}{2}l^{-2}(l^2t^2+s^2),$$

whence

$$\tfrac{3}{2}\langle h,h\rangle \ge \langle U^{-1}h,h\rangle \ge \tfrac{1}{2}(1-a_{\max})\langle h,h\rangle. \tag{10.1.14}$$

Relations (10.1.13) and (10.1.14) combined with $-p' = \lambda Q^{-1}UQ^{-1}$ immediately lead to (10.1.5)–(10.1.7). □

## Main result

Since $p(x)$ is a one-to-one $C^\infty$-smooth mapping of $\mathbf{R}^n_{++}$ onto itself and $-p'$ is positive-definite (although nonsymmetric), the inverse mapping $x(\cdot)$ is also $C^\infty$-smooth and antimonotone.

**Theorem 10.1.2** *The mapping $-x'(p) : \mathbf{R}^n_{++} \to \mathbf{R}^n_{++}$ is $\beta$-compatible with the standard logarithmic barrier*

$$F(u) = -\sum_{i=1}^n \ln u_i$$

*for the nonnegative orthant $\mathbf{R}^n_+$, where*

$$\beta = 3\frac{(a_{\max}/a_{\min})^{1/3}}{(1-a_{\max})^{5/2}}\sqrt{2n},$$

$$a_{\min} = \min\{a_1, \dots, a_n\}.$$

**Remark 10.1.1** *From the above statement, it follows immediately that the mapping*

$$p \to \operatorname{argmin}\{f(x) \mid x \geq 0,\ \langle p, x\rangle \leq w\},$$

*$w$ being a positive constant, is $\beta$-compatible with $F$; this is precisely the fact used in the Application Example; see* §6.3.7.

**Proof of the theorem.** $1^0$. Let us fix a positive vector $p \in \mathbf{R}^n$ and a triple of vectors $\eta,\ \rho,\ \tau \in \mathbf{R}^n$. Let us compute $x'(p)$ and $x''(p)\eta,\ \eta \in \mathbf{R}^n$. Let $x = x(p)$, so that $p = p(x)$, and let $\lambda = \lambda(x)$. We have (derivatives of $x(\cdot)$ are taken at $p$, derivatives of $p(\cdot)$ are taken at $x$)

$$x' = (p')^{-1}; \tag{10.1.15}$$

$$x''[\eta] = -(p')^{-1}p''[(p')^{-1}\eta](p')^{-1}; \tag{10.1.16}$$

henceforth, for a smooth mapping, $S : \mathbf{R}^n \to \mathbf{R}^n$ $S''[\eta]$ denotes the derivative in the direction $\eta$ of the mapping $u \to S'(u)$; this mapping takes values in the space of $n \times n$ matrices, so that the value $S''[\eta]$ is an $n \times n$ matrix linearly depending on $\eta$.

$2^0$. Now let $\Phi$ be the Hessian of $F$ at $p$,

$$\langle \Phi h, h\rangle = \left\langle P^{-2}h, h\right\rangle. \tag{10.1.17}$$

We prove that $-x(\cdot)$ is $\beta$-compatible with the barrier $F$, or, which is the same, that

$$|\langle -x''[\eta]\rho, \tau\rangle| \leq 3^{3/2}\beta g(\eta)g(\rho)g(\tau) \equiv 3^{3/2}\beta\Theta, \tag{10.1.18}$$

where

$$g(\sigma) = \langle -x'\sigma, \sigma\rangle^{1/3}\langle \Phi\sigma, \sigma\rangle^{1/6}. \tag{10.1.19}$$

Let

$$(p')X = \mathcal{D}, \qquad (p')^T X = \mathcal{E}.$$

$3^0$. Let us prove some lemmas.

**Lemma 10.1.4** *We have*

(10.1.20)
$$g(\mathcal{D}s) \geq \left(\tfrac{\lambda}{2}\right)^{1/3}\langle CA(E-A)X^a s, s\rangle^{1/3}(2n)^{-1/6}\parallel s \parallel_2^{1/3}(1-a_{\max})^{1/3},$$

$$\begin{aligned} g(\mathcal{E}s) \geq & \left(\frac{\lambda}{2}\right)^{1/3}\langle CA(E-A)X^a s, s\rangle^{1/3} \\ & \cdot(2n)^{-1/6}\parallel s \parallel_2^{1/3}(1-a_{\max})^{1/3}\left(\frac{a_{\min}}{a_{\max}}\right)^{1/3}. \end{aligned} \tag{10.1.21}$$

**Proof.** We have

$$\begin{aligned} g(\mathcal{D}s) &= \langle -x'(p')Xs, (p')Xs\rangle^{1/3} \langle \Phi(p')Xs, (p')Xs\rangle^{1/6} \\ &= \langle Xs, (-p')Xs\rangle^{1/3} \left\langle P^{-2}(p')Xs, (p')Xs\right\rangle^{1/6} \\ &\geq \left(\frac{\lambda}{2}\right)^{1/3} \langle -f''Xs, Xs\rangle^{1/3} \left\langle P^{-2}(p')Xs, (p')Xs\right\rangle^{1/6} \end{aligned}$$

(the second equality holds true by virtue of $x' = (p')^{-1}$; the inequality follows from (10.1.5)). Thus,

$$g(\mathcal{D}s) \geq \left(\frac{\lambda}{2}\right)^{1/3} \langle -f''Xs, Xs\rangle^{1/3} \left\langle P^{-2}(p')Xs, (p')Xs\right\rangle^{1/6}. \tag{10.1.22}$$

We have

$$\langle -f''Xs, Xs\rangle = \langle CA(E-A)X^a s, s\rangle. \tag{10.1.23}$$

Now, from (10.1.7) and relation $p = \lambda f'(x)$, it follows that

$$\begin{aligned} P^{-1}(-p') &= \lambda^{-1}\mathrm{Diag}^{-1}\{f'\}(-p') = (E-A)X^{-e} + \lambda e(f')^T - \lambda e((E-A)f')^T \\ &= (E-A)X^{-e} + \lambda(ee^T)CA^2X^{a-e}; \end{aligned}$$

thus,

$$\begin{aligned} (-p')^T P^{-2}(-p') &= \{(E-A)X^{-e} + \lambda CA^2X^{a-e}(ee^T)\} \\ &\quad \cdot \{(E-A)X^{-e} + \lambda(ee^T)CA^2X^{a-e}\} \\ &= (E-A)^2X^{-2e} + \lambda CA^2X^{a-e}(ee^T)(E-A)X^{-e} \\ &\quad + \lambda(E-A)X^{-e}(ee^T)CA^2X^{a-e} \\ &\quad + \lambda^2 nCA^2X^{a-e}(ee^T)CA^2X^{a-e} \end{aligned}$$

and

$$\begin{aligned} X(-p')^T P^{-2}(-p')X &= (E-A)^2 + \lambda CA^2X^a(ee^T)(E-A) + \lambda(E-A)(ee^T)CA^2X^a \\ &\quad + \lambda^2 nCA^2X^a(ee^T)CA^2X^a. \end{aligned}$$

Let $\phi = \lambda CA^2X^a e$, $\psi = (E-A)e$. Then

$$\begin{aligned} X(-p')^T P^{-2}(-p')X &= (E-A)\{E - n^{-1}ee^T\}(E-A) \\ &\quad + (n^{-1/2}\psi + n^{1/2}\phi)(n^{-1/2}\psi + n^{1/2}\phi)^T. \end{aligned}$$

Let

$$\gamma = (E-A)^{-1}(n^{-1/2}\psi + n^{1/2}\phi) \equiv n^{-1/2}e + n^{1/2}\lambda CA^2(E-A)^{-1}X^a e,$$

so that

$$X(-p')^T P^{-2}(-p')X = (E-A)\{E - n^{-1}ee^T + \gamma\gamma^T\}(E-A). \tag{10.1.24}$$

**Lemma 10.1.5** *For all $s \in \mathbf{R}^n$, we have*

$$\left\langle \{E - n^{-1}ee^T + \gamma\gamma^T\}s, s\right\rangle \geq \frac{1}{2n} \parallel s \parallel_2^2 . \tag{10.1.25}$$

**Proof.** Let $L$ be the subspace generated by $e$ and $\gamma$; then $L$ and $L^\perp$ are invariant with respect to the symmetric operator $S = E - n^{-1}ee^T + \gamma\gamma^T$. Let $s = s' + s''$, $s' \in L$, $s'' \in L^\perp$. We have

$$\langle Ss, s\rangle = \langle Ss', s'\rangle + \parallel s'' \parallel_2^2 .$$

The vector $\gamma$ clearly is positive and $\parallel \gamma \parallel_2 \geq 1$. If $L$ is one-dimensional, i.e., if $\gamma = \omega e$, then $Ss' = \parallel \gamma \parallel_2^2 s'$, so that $\langle Ss', s'\rangle \geq \parallel s' \parallel_2^2$ . Now let $L$ be two-dimensional and let $v = n^{-1/2}\, e$ and $w$ be the unit vector in $L$ orthogonal to $v$. Let $\alpha$ be the angle between $\gamma$ and e; since $\gamma$ is positive, we have $\cos\alpha \geq n^{-1/2}$. Let $r' = \langle s', v\rangle$ , $r'' = \langle s', w\rangle$ , so that

$$\begin{aligned}\langle Ss', s'\rangle &= (r'')^2 + \parallel \gamma \parallel_2^2 (r'\cos\alpha + r''\sin\alpha)^2 \\ &\geq (r'')^2 + (r'\cos\alpha + r''\sin\alpha)^2.\end{aligned}$$

Let $r' = r\cos\beta$, $r'' = r\sin\beta$, where $r = \parallel s' \parallel_2$;, then

$$\langle Ss', s'\rangle = r^2\{\sin^2\beta + \cos^2(\alpha - \beta)\}.$$

The minimum of the latter quantity in $\beta$ is $2r^2\sin^2(\pi/4 - \alpha/2) \geq (2n)^{-1}r^2$. It follows that $\langle Ss, s\rangle \geq (2n)^{-1} \parallel s \parallel_2^2$ . □

Relations (10.1.22)–(10.1.25) lead to (10.1.20).

Now let us prove (10.1.21). We have

$$\begin{aligned} g(\mathcal{E}s) &= \left\langle -x'(p')^T Xs, (p')^T Xs\right\rangle^{1/3} \left\langle \Phi(p')^T Xs, (p')^T Xs\right\rangle^{1/6} \\ &= \langle Xs, (-p')Xs\rangle^{1/3} \left\langle P^{-2}(p')^T Xs, (p')^T Xs\right\rangle^{1/6} \\ &\geq \left(\frac{\lambda}{2}\right)^{1/3} \langle -f''Xs, Xs\rangle^{1/3} \left\langle P^{-2}(p')^T Xs, (p')^T Xs\right\rangle^{1/6} \end{aligned}$$

(the second equality holds true by virtue of $x' = (p')^{-1}$, the inequality follows from (10.1.5)). The resulting inequality combined with (10.1.23) leads to

$$g(\mathcal{E}s) \geq \left(\frac{\lambda}{2}\right)^{1/3} \langle CA(E - A)X^a s, s\rangle^{1/3} \left\langle P^{-2}(p')^T Xs, (p')^T Xs\right\rangle^{1/6}. \tag{10.1.26}$$

Now, from (10.1.7) and relation $p = \lambda f'(x)$, it follows that

$$\begin{aligned} P^{-1}(-p')^T &= \lambda^{-1}\mathrm{Diag}^{-1}\{f'\}(-p')^T \\ &= (E - A)X^{-e} + \lambda(ee^T)CAX^{a-e} \\ &\quad - \lambda(C^{-1}A^{-1}X^{-a+e}CA(E - A)X^{a-2e}(x(f')^T)) \\ &= (E - A)X^{-e} + \lambda(ee^T)CAX^{a-e} - \lambda(E - A)X^{-e}(Xee^T)CAX^{a-e} \\ &= (E - A)X^{-e} + \lambda A(ee^T)CAX^{a-e}. \end{aligned}$$

Thus,

$$\begin{aligned}(-p')P^{-2}(-p')^T &= \{(E-A)X^{-e}+\lambda CAX^{a-e}(ee^T)A\}\\ &\quad\times\{(E-A)X^{-e}+\lambda A(ee^T)CAX^{a-e}\}\\ &= (E-A)^2X^{-2e}+\lambda CAX^{a-e}(ee^T)A(E-A)X^{-e}\\ &\quad+\lambda(E-A)X^{-e}A(ee^T)CAX^{a-e}\\ &\quad+\lambda^2\langle Ae,Ae\rangle CAX^{a-e}(ee^T)CAX^{a-e}\end{aligned}$$

and

$$\begin{aligned}X(-p')P^{-2}(-p')^TX &= (E-A)^2+\lambda CAX^a(ee^T)A(E-A)+\lambda(E-A)A(ee^T)CAX^a\\ &\quad+\lambda^2\langle Ae,Ae\rangle CAX^a(ee^T)CAX^a.\end{aligned}$$

Let $\theta=\lambda CX^aAe$, $\zeta=(E-A)Ae$. Then

$$\begin{aligned}X(-p')P^{-2}(-p')^TX &= (E-A)^2+\theta\zeta^T+\zeta\theta^T+\langle Ae,Ae\rangle\,\theta\theta^T\\ &= (E-A)^2+\{\langle Ae,Ae\rangle^{1/2}\theta+\langle Ae,Ae\rangle^{-1/2}\zeta\}\\ &\quad\cdot\{\langle Ae,Ae\rangle^{1/2}\theta+\langle Ae,Ae\rangle^{-1/2}\zeta\}^T\\ &\quad-\langle Ae,Ae\rangle^{-1}(E-A)((Ae)(Ae)^T)(E-A)\\ &= (E-A)\{E-\langle Ae,Ae\rangle^{-1}(Ae)(Ae)^T+\mu\mu^T\}(E-A),\end{aligned}$$

where

$$\mu=(E-A)^{-1}\{\langle Ae,Ae\rangle^{1/2}\theta+\langle Ae,Ae\rangle^{-1/2}\zeta\}.$$

Thus,

$$\begin{aligned}&X(-p')P^{-2}(-p')^TX\\ &\quad=(E-A)\{E-\langle Ae,Ae\rangle^{-1}(Ae)(Ae)^T+\mu\mu^T\}(E-A),\end{aligned}\tag{10.1.27}$$

$$\mu=\langle Ae,Ae\rangle^{1/2}(E-A)^{-1}\lambda CX^aAe+\langle Ae,Ae\rangle^{-1/2}Ae.\tag{10.1.28}$$

Let $\nu=\langle Ae,Ae\rangle$, $\varepsilon=Ae$.

**Lemma 10.1.6** *For all $s\in\mathbf{R}^n$, we have*

$$\left\langle\{E-\nu^{-1}\varepsilon\varepsilon^T+\mu\mu^T\}s,s\right\rangle\geq\frac{1}{2n}\left(\frac{a_{\min}}{a_{\max}}\right)^2\|\,s\,\|_2^2\,.\tag{10.1.29}$$

**Proof.** Let $L$ be the subspace generated by $\varepsilon$ and $\mu$; then $L$ and $L^\perp$ are invariant with respect to the symmetric operator $\Sigma=E-\nu^{-1}\varepsilon\varepsilon^T+\mu\mu^T$. Let $s=s'+s''$, $s'\in L$, $s''\in L^\perp$. We have $\langle\Sigma s,s\rangle=\langle\Sigma s',s'\rangle+\|\,s''\,\|_2^2$. The vector $\mu$ clearly is positive and $\|\,\mu\,\|_2\geq 1$. If $L$ is one-dimensional, i.e., if $\mu=\omega\varepsilon$, then $\Sigma s'=\|\,\mu\,\|_2^2\,s'$, so that

$$\langle Ss',s'\rangle\geq\|\,s'\,\|_2^2\,.$$

Now let $L$ be two-dimensional and let $v = \nu^{-1/2}\varepsilon$ and $w$ be the unit vector in $L$ orthogonal to $v$. Let $\delta$ be the angle between $\mu$ and $\varepsilon$; since $\mu$ and $\varepsilon$ are positive, $\cos\delta$ is not less than the minimal entry of $\| \varepsilon \|_2^{-1} \varepsilon = \nu^{-1}\varepsilon$, so that $\cos\delta \ge n^{-1/2}(a_{\min}/a_{\max})$. Let $r' = \langle s', v\rangle$, $r'' = \langle s', w\rangle$, so that

$$\langle \Sigma s', s'\rangle = (r'')^2 + \| \mu \|_2^2 (r'\cos\delta + r''\sin\delta)^2 \ge (r'')^2 + (r'\cos\delta + r''\sin\delta)^2.$$

Let $r' = r\cos\beta$, $r'' = r\sin\beta$, where $r = \| s' \|_2$;, then $\langle \Sigma s', s'\rangle = r^2\{\sin^2\beta + \cos^2(\delta - \beta)\}$. The latter quantity is at least

$$2r^2\sin^2\left(\frac{\pi}{4} - \frac{\delta}{2}\right) \ge \frac{1}{2n}\left(\frac{a_{\min}}{a_{\max}}\right)^2.$$

It follows that

$$\langle \Sigma s, s\rangle \ge \frac{1}{2n}\left(\frac{a_{\min}}{a_{\max}}\right)^2 \| s \|_2^2 .$$

Relations (10.1.26)–(10.1.29) lead to (10.1.21). □

$4^0$. We have

$$\Omega \equiv -\langle x''[\eta]\rho, \tau\rangle = \left\langle p''[(p')^{-1}\eta](p')^{-1}\rho, ((p')^T)^{-1}\tau\right\rangle. \tag{10.1.30}$$

Let

$$\mathsf{h} = (p')^{-1}\eta,\quad \mathsf{r} = (p')^{-1}\rho,\quad \mathsf{t} = (p')^{-1}\tau,\quad B = ((p')^T)^{-1}p'; \tag{10.1.31}$$

then

$$\Omega = \langle p''[\mathsf{h}]\mathsf{r}, B\mathsf{t}\rangle. \tag{10.1.32}$$

We have (see (10.1.7))

(10.1.33)

$$\begin{aligned} p''[\mathsf{h}] &= \lambda(f')''[\mathsf{h}] - \lambda^2\{\langle f''\mathsf{h}, x\rangle + \langle f', \mathsf{h}\rangle\}f'' \\ &\quad +2\lambda^3\{\langle f''\mathsf{h}, x\rangle + \langle f', \mathsf{h}\rangle\}\{(f')(f')^T + ((f')x^T)f''\} \\ &\quad -\lambda^2\{(f''\mathsf{h})(f')^T + f'(f''\mathsf{h})^T + ((f''\mathsf{h})x^T)f'' + (f'\mathsf{h}^T)f'' + (f'x^T)(f')''[\mathsf{h}]\} \\ &= \sum_{k=1}^{9} A_k[\mathsf{h}], \end{aligned}$$

where

$$\begin{aligned} \mathbf{A}_1[\mathsf{h}] &= \lambda(f')''[\mathsf{h}], \\ \mathbf{A}_2[\mathsf{h}] &= -\lambda^2\langle f''\mathsf{h}, x\rangle f'', \\ \mathbf{A}_3[\mathsf{h}] &= -\lambda^2\langle f', \mathsf{h}\rangle f'', \\ \mathbf{A}_4[\mathsf{h}] &= 2\lambda^3\{\langle f''\mathsf{h}, x\rangle + \langle f', \mathsf{h}\rangle\}(f')(f')^T, \\ \mathbf{A}_5[\mathsf{h}] &= 2\lambda^3\{\langle f''\mathsf{h}, x\rangle + \langle f', \mathsf{h}\rangle\}((f')x^T)f'', \\ \mathbf{A}_6[\mathsf{h}] &= -\lambda^2\{(f''\mathsf{h})(f')^T + f'(f''\mathsf{h})^T\}, \\ \mathbf{A}_7[\mathsf{h}] &= -\lambda^2((f''\mathsf{h})x^T)f'', \end{aligned}$$

$$\mathbf{A}_8[\mathsf{h}] = -\lambda^2(f'\mathsf{h}^T)f'',$$
$$\mathbf{A}_9[\mathsf{h}] = -\lambda^2(f'x^T)(f')''[\mathsf{h}].$$

Let

$$h = X^{-1}\mathsf{h}, \quad r = X^{-1}\mathsf{r}, \quad t = X^{-1}B\mathsf{t},$$

so that

$$\eta = \mathcal{D}h, \rho = \mathcal{D}r, \tau = \mathcal{E}t. \tag{10.1.34}$$

**A.** We have

$$\begin{aligned} |\langle \mathbf{A}_1[\mathsf{h}]\mathsf{r}, B\mathsf{t}\rangle| &= \lambda\, |\langle CA(E-A)(2E-A)X^a HRe, Te\rangle| \\ &\le 2\lambda \,\langle CA(E-A)X^a|H||R||T|e, e\rangle\,, \end{aligned} \tag{10.1.35}$$

where $|Y|$ denotes the matrix obtaining from a matrix $Y$ by replacing each entry by its absolute value. By virtue of Holder inequality, we have

(10.1.36)
$$\begin{aligned} &2\lambda\,\langle CA(E-A)X^a|H||R||T|e, e\rangle \\ &\le 2\lambda \left\langle CA(E-A)X^a|H|^3e, e\right\rangle^{1/3} \left\langle CA(E-A)X^a|R|^3e, e\right\rangle^{1/3} \\ &\quad\times \left\langle CA(E-A)X^a|T|^3e, e\right\rangle^{1/3} \\ &\le 2\,\{\lambda\,\langle CA(E-A)X^a h, h\rangle\}^{1/3}\,\{\lambda\,\langle CA(E-A)X^a r, r\rangle\}^{1/3} \\ &\quad\times \{\lambda\,\langle CA(E-A)X^a t, t\rangle\}^{1/3} \parallel h \parallel_\infty^{1/3} \parallel r \parallel_\infty^{1/3} \parallel t \parallel_\infty^{1/3} \equiv 2\Lambda. \end{aligned}$$

At the same time, (10.1.20) and (10.1.34) demonstrate that

$$\{\lambda\,\langle CA(E-A)X^a h, h\rangle\}^{1/3} \parallel h \parallel_\infty^{1/3} \le 2^{1/3}(1-a_{\max})^{-1/3}(2n)^{1/6}g(\eta), \tag{10.1.37}$$

$$\{\lambda\,\langle CA(E-A)X^a r, r\rangle\}^{1/3} \parallel r \parallel_\infty^{1/3} \le 2^{1/3}(1-a_{\max})^{-1/3}(2n)^{1/6}g(\rho), \tag{10.1.38}$$

while (10.1.21) leads to

$$\begin{aligned} &\{\lambda\,\langle CA(E-A)X^a t, t\rangle\}^{1/3} \parallel t \parallel_2^{1/3} \\ &\le 2^{1/3}(1-a_{\max})^{-1/3}(2n)^{1/6}(a_{\max}/a_{\min})^{1/3}g(\tau). \end{aligned} \tag{10.1.39}$$

It follows (see (10.1.35)–(10.1.39), (10.1.18)) that

$$\Lambda \le (a_{\max}/a_{\min})^{1/3}(1-a_{\max})^{-1}(2n)^{1/2}\Theta \tag{10.1.40}$$

and

$$|\,\langle \mathbf{A}_1[\mathsf{h}]\mathsf{r}, B\mathsf{t}\rangle\,| \le 2\Lambda. \tag{10.1.41}$$

**B.** We have

$$\{\lambda\,\langle CA(E-A)X^a h, h\rangle\}^{1/2}\{\lambda\,\langle CA(E-A)X^a r, r\rangle\}^{1/2}$$
$$\cdot\{\lambda\,\langle CA(E-A)X^a t, t\rangle\}^{1/2}$$

$$= \{\lambda \langle CA(E-A)X^a h, h\rangle\}^{1/3}\{\lambda \langle CA(E-A)X^a r, r\rangle\}^{1/3}$$
$$\cdot\{\lambda \langle CA(E-A)X^a t, t\rangle\}^{1/3}$$
$$\cdot\{\lambda \langle CA(E-A)X^a h, h\rangle\}^{1/6}\{\lambda \langle CA(E-A)X^a r, r\rangle\}^{1/6}$$
$$\cdot\{\lambda \langle CA(E-A)X^a t, t\rangle\}^{1/6}.$$

Since $\lambda = \langle CAX^a e, e\rangle^{-1}$, we clearly have $\{\lambda \langle CA(E-A)s, s\rangle\}^{1/6} \leq \| s \|_\infty^{1/3}$. Thus,

$$\begin{aligned}&\{\lambda \langle CA(E-A)X^a h, h\rangle\}^{1/2}\{\lambda \langle CA(E-A)X^a r, r\rangle\}^{1/2}\\ &\quad\cdot\{\lambda \langle CA(E-A)X^a t, t\rangle\}^{1/2} \leq \Lambda.\end{aligned} \tag{10.1.42}$$

Note also that

$$\lambda \langle CAX^a s, s\rangle \leq (1-a_{\max})^{-1}\{\lambda \langle CA(E-A)X^a s, s\rangle\}. \tag{10.1.43}$$

**C.** We have

$$\begin{aligned}
|\langle \mathbf{A}_2[\mathsf{h}]\mathsf{r}, B\mathsf{t}\rangle| &= \lambda^2 |\langle f''\mathsf{h}, x\rangle|\,|\langle f''\mathsf{r}, B\mathsf{t}\rangle| = \lambda^2 |\langle f''Xh, Xe\rangle|\;|\langle f''Xr, Xt\rangle| \\
&\quad\times \lambda^2 \left|\left\langle CA(E-A)X^{a-2e}XHe, Xe\right\rangle\right| \\
&\quad \lambda^2 \left|\left\langle CA(E-A)X^{a-2e}XRe, XTe\right\rangle\right| \\
&= \lambda^2 \langle CA(E-A)X^a|H|e, e\rangle\,\langle CA(E-A)X^a|R|e, |T|e\rangle \\
&= \lambda^2 \left\langle \{CA(E-A)X^a\}^{1/2}|H|e, \{CA(E-A)X^a\}^{1/2}e\right\rangle \\
&\quad\times \left\langle \{CA(E-A)X^a\}^{1/2}|R|e, \{CA(E-A)X^a\}^{1/2}|T|e\right\rangle \\
&\leq \lambda^2 \left\langle \{CA(E-A)X^a\}^{1/2}e, \{CA(E-A)X^a\}^{1/2}e\right\rangle^{1/2} \\
&\quad\times \left\langle \{CA(E-A)X^a\}^{1/2}h, \{CA(E-A)X^a\}^{1/2}h\right\rangle^{1/2} \\
&\quad\times \left\langle \{CA(E-A)X^a\}^{1/2}r, \{CA(E-A)X^a\}^{1/2}r\right\rangle^{1/2} \\
&\quad\times \left\langle \{CA(E-A)X^a\}^{1/2}t, \{CA(E-A)X^a\}^{1/2}t\right\rangle^{1/2} \\
&= \{\lambda \langle CA(E-A)X^a e, e\rangle\}^{1/2}\{\lambda \langle CA(E-A)X^a h, h\rangle\}^{1/2} \\
&\quad\times \{\lambda \langle CA(E-A)X^a r, r\rangle\}^{1/2}\{\lambda \langle CA(E-A)X^a t, t\rangle^{1/2}\} \leq \Lambda;
\end{aligned}$$

the latter inequality follows from (10.1.42) in view of

$$\lambda = \langle CAX^a e, e\rangle^{-1} \leq \langle CA(E-A)X^a e, e\rangle^{-1}.$$

Thus,

$$|\langle \mathbf{A}_2[\mathsf{h}]\mathsf{r}, B\mathsf{t}\rangle| \leq \Lambda. \tag{10.1.44}$$

**D.** We have (compare with **B**)

$$
\begin{aligned}
|\langle \mathbf{A}_3[\mathsf{h}]\mathsf{r}, B\mathsf{t}\rangle| &= \lambda^2\, |\langle f', \mathsf{h}\rangle|\; |\langle f''\mathsf{r}, \mathsf{Bt}\rangle| \\
&= \lambda^2\, |\langle CAX^{a-e}e, Xh\rangle| \left|\left\langle CA(E-A)X^{a-2e}XRe, XTe\right\rangle\right| \\
&= \lambda^2\, |\langle CAX^{a}He, e\rangle| \\
&\quad \times |\langle CA(E-A)X^{a}Re, Te\rangle| \\
&\le \lambda^2\, \langle CAX^{a}e, e\rangle^{1/2}\, \langle CAX^{a}h, h\rangle^{1/2} \\
&\quad \times \langle CA(E-A)X^{a}r, r\rangle^{1/2}\, \langle CA(E-A)X^{a}t, t\rangle^{1/2} \\
&= \{\lambda\, \langle CAX^{a}h, h\rangle\}^{1/2}\{\lambda\, \langle CA(E-A)X^{a}r, r\rangle\}^{1/2} \\
&\quad \times \{\lambda\, \langle CA(E-A)X^{a}t, t\rangle\}^{1/2} \\
&\le (1-a_{\max})^{-1/2}\Lambda
\end{aligned}
$$

(the latter inequality follows from (10.1.42), (10.1.43)). Thus,

$$
|\, \langle \mathbf{A}_3[\mathsf{h}]\mathsf{r}, B\mathsf{t}\rangle \,| \le \Lambda(1-a_{\max})^{-1/2}. \tag{10.1.45}
$$

**E.** We have $f''X^2e + Xf' = CA^2X^ae$. Now,

$$
\begin{aligned}
|\langle \mathbf{A}_4[\mathsf{h}]\mathsf{r}, B\mathsf{t}\rangle| &= 2\lambda^3\, |\langle f''\mathsf{h}, x\rangle + \langle f', \mathsf{h}\rangle|\, |\langle f', \mathsf{r}\rangle\,||\, \langle f', B\mathsf{t}\rangle| \\
&= 2\lambda^3\, |\langle f''XHe, Xe\rangle + \langle f', XHe\rangle|\; |\langle f', XRe\rangle|\; |\langle f', XTe\rangle| \\
&= 2\lambda^3 \left|\left\langle CA^2X^{a}h, e\right\rangle\right|\; |\langle CAX^{a}e, r\rangle|\; |\langle CAX^{a}e, t\rangle| \\
&\le 2\lambda^3\, \langle CAX^{a}h, h\rangle^{1/2} \\
&\quad \times \langle CAX^{a}e, e\rangle^{1/2}\, \langle CAX^{a}r, r\rangle^{1/2}\, \langle CAX^{a}e, e\rangle^{1/2} \\
&\quad \times \langle CAX^{a}t, t\rangle^{1/2}\, \langle CAX^{a}e, e\rangle^{1/2} \\
&= 2\{\lambda\, \langle CAX^{a}h, h\rangle\}^{1/2}\{\lambda\, \langle CAX^{a}r, r\rangle\}^{1/2}\{\lambda\, \langle CAX^{a}t, t\rangle\}^{1/2} \\
&\le 2(1-a_{\max})^{-3/2}\Lambda
\end{aligned}
$$

(we have used (10.1.42), (10.1.43)). Thus,

$$
|\, \langle \mathbf{A}_4[\mathsf{h}]\mathsf{r}, B\mathsf{t}\rangle \,| \le 2(1-a_{\max})^{-3/2}\Lambda. \tag{10.1.46}
$$

**F.** We have (compare with **E**)

$$
\begin{aligned}
|\langle \mathbf{A}_5[\mathsf{h}]\mathsf{r}, B\mathsf{t}\rangle| &= 2\lambda^3\, |\langle f''\mathsf{h}, x\rangle + \langle f', \mathsf{h}\rangle|\; |\langle f', B\mathsf{t}\rangle|\; |\langle f''\mathsf{r}, x\rangle| \\
&= 2\lambda^3 \left|\left\langle CA^2X^{a}h, e\right\rangle\right|\; |\langle CAX^{a}t, e\rangle|\; |\langle CA(E-A)X^{a}r, e\rangle|\,,
\end{aligned}
$$

and the latter quantity, as in **E**, proves to be $\le 2(1-a_{\max})^{-1}\Lambda$. Thus,

$$
|\langle \mathbf{A}_5[\mathsf{h}]\mathsf{r}, B\mathsf{t}\rangle| \le 2(1-a_{\max})^{-1}\Lambda. \tag{10.1.47}
$$

**G.** We have

$$
\begin{aligned}
|\langle \mathbf{A}_6[\mathsf{h}]\mathsf{r}, B\mathsf{t}\rangle| &= \lambda^2 \,|\langle f''\mathsf{h}, B\mathsf{t}\rangle \langle f', \mathsf{r}\rangle + \langle f', B\mathsf{t}\rangle \langle f''\mathsf{h}, \mathsf{r}\rangle| \\
&= \lambda^2 \,|\langle CA(E-A)X^a h, t\rangle \langle CAX^a e, r\rangle \\
&\quad + \langle CA(E-A)X^a h, r\rangle \langle CAX^a e, t\rangle| \\
&\le \lambda^2 \{\langle CA(E-A)X^a h, h\rangle^{1/2} \langle CA(E-A)t, t\rangle^{1/2} \\
&\qquad \cdot \langle CAX^a r, r\rangle^{1/2} \langle CAX^a e, e\rangle^{1/2} \\
&\qquad + \langle CA(E-A)X^a h, h\rangle^{1/2} \langle CA(E-A)X^a r, r\rangle^{1/2} \\
&\qquad \times \langle CAX^a t, t\rangle^{1/2} \langle CAX^a e, e\rangle^{1/2}\} \\
&= \{\lambda \langle CA(E-A)X^a h, h\rangle\}^{1/2} \{\lambda \langle CA(E-A)X^a t, t\rangle\}^{1/2} \\
&\qquad \cdot \{\lambda \langle CAX^a r, r\rangle\}^{1/2} \\
&\qquad + \{\lambda \langle CA(E-A)X^a h, h\rangle\}^{1/2} \{\lambda \langle CA(E-A)X^a r, r\rangle\}^{1/2} \\
&\qquad \times \{\lambda \langle CAX^a t, t\rangle\}^{1/2} \le 2(1-a_{\max})^{-1/2}\Lambda
\end{aligned}
$$

(see (10.1.42), (10.1.43)). Thus,

$$
|\langle \mathbf{A}_6[\mathsf{h}]\mathsf{r}, B\mathsf{t}\rangle| \le 2(1-a_{\max})^{-1/2}\Lambda. \tag{10.1.48}
$$

**H.** Now

$$
\begin{aligned}
|\langle \mathbf{A}_7[\mathsf{h}]\mathsf{r}, B\mathsf{t}\rangle| &= \lambda^2 \,|\langle f''\mathsf{h}, B\mathsf{t}\rangle \langle f''x, \mathsf{r}\rangle| = \lambda^2 \,|\langle f''Xh, Xt\rangle \langle f''Xe, Xr\rangle| \\
&= \lambda^2 \,|\langle CA(E-A)X^a h, t\rangle \langle CA(E-A)X^a r, e\rangle| \\
&\le \lambda^2 \langle CA(E-A)X^a h, h\rangle^{1/2} \langle CA(E-A)X^a t, t\rangle^{1/2} \\
&\qquad \times \langle CA(E-A)X^a r, r\rangle^{1/2} \langle CA(E-A)X^a e, e\rangle^{1/2} \\
&\le \{\lambda \langle CA(E-A)X^a h, h\rangle\}^{1/2} \{\lambda \langle CA(E-A)X^a t, t\rangle\}^{1/2} \\
&\qquad \cdot \{\lambda \langle CA(E-A)X^a r, r\rangle\}^{1/2} \le \Lambda
\end{aligned}
$$

(see (10.1.42)). Thus,

$$
|\langle \mathbf{A}_7[\mathsf{h}]\mathsf{r}, B\mathsf{t}\rangle| \le \Lambda. \tag{10.1.49}
$$

**I.** We have

$$
\begin{aligned}
|\langle \mathbf{A}_8[\mathsf{h}]\mathsf{r}, B\mathsf{t}\rangle| &= \lambda^2 \,|\langle f', B\mathsf{t}\rangle \langle f''\mathsf{h}, \mathsf{r}\rangle| = \lambda^2 \,|\langle f', Xt\rangle \langle f''Xh, Xr\rangle| \\
&= \lambda^2 \,|\langle CAX^a e, t\rangle \langle CA(E-A)X^a h, r\rangle| \\
&\le \lambda^2 \langle CAX^a e, e\rangle^{1/2} \langle CAX^a t, t\rangle^{1/2} \\
&\qquad \times \langle CA(E-A)X^a h, h\rangle^{1/2} \langle CA(E-A)X^a r, r\rangle^{1/2} \\
&= \{\lambda \langle CAX^a t, t\rangle\}^{1/2} \{\lambda \langle CA(E-A)X^a h, h\rangle\}^{1/2} \\
&\qquad \times \{\lambda \langle CA(E-A)X^a r, r\rangle\}^{1/2} \le (1-a_{\max})^{-1/2}\Lambda
\end{aligned}
$$

(see (10.1.42), (10.1.43)). Thus,

$$|\langle \mathbf{A}_8[\mathsf{h}]\mathsf{r}, B\mathsf{t}\rangle| \leq (1 - a_{\max})^{-1/2}\Lambda. \tag{10.1.50}$$

**J.** We have ($\mathsf{H} = \operatorname{Diag}\{\mathsf{h}\}$)

$$\begin{aligned}
&|\langle \mathbf{A}_9[\mathsf{h}]\mathsf{r}, B\mathsf{t}\rangle| = \lambda^2 \left|\langle f', B\mathsf{t}\rangle \langle (f')''[\mathsf{h}]\mathsf{r}, x\rangle\right| \\
&= \lambda^2 \left|\langle f', Xt\rangle \left\langle CA(E-A)(2E-A)X^{a-3e}\mathsf{H}\mathsf{r}, x\right\rangle\right| \\
&= \lambda^2 \left|\langle CAX^a e, t\rangle \langle CA(E-A)(2E-A)X^a HRe, e\rangle\right| \\
&= \lambda^2 \left|\langle CAX^a e, t\rangle \langle CA(E-A)(2E-A)X^a h, r\rangle\right| \leq \lambda^2 \langle CAX^a t, t\rangle^{1/2} \\
&\times \langle CAX^a e, e\rangle^{1/2} \langle CA(E-A)(2E-A)X^a h, h\rangle^{1/2} \\
&\cdot \langle CA(E-A)(2E-A)X^a r, r\rangle^{1/2} \\
&\leq 2\lambda^{3/2} \langle CAX^a t, t\rangle^{1/2} \langle CA(E-A)X^a h, h\rangle^{1/2} \langle CA(E-A)X^a r, r\rangle^{1/2} \\
&\leq 2(1 - a_{\max})^{-1/2}\Lambda
\end{aligned}$$

(see (10.1.42), (10.1.43)). Thus,

$$|\langle \mathbf{A}_9[\mathsf{h}]\mathsf{r}, B\mathsf{t}\rangle| \leq 2(1 - a_{\max})^{-1/2}\Lambda. \tag{10.1.51}$$

$5^0$. From (10.1.30), (10.1.32), (10.1.33), and (10.1.41)–(10.1.51), it follows that

$$|- \langle x''[\eta]\rho, \tau\rangle| \leq 14(1 - a_{\max})^{-3/2}\Lambda$$

or, in view of (10.1.40),

$$|- \langle x''[\eta]\rho, \tau\rangle| \leq 14(a_{\max}/a_{\min})^{1/3}(1 - a_{\max})^{-5/2}(2n)^{1/2}\Theta;$$

by definition of $\beta$, the latter quantity is $\leq 3^{3/2}\beta\Theta$, so that (10.1.18) does hold true. $\square$

*This page intentionally left blank*

# Bibliography comments

What follows is a very brief historical and bibliography comment to the monograph. We attempted to mention all contributions to the subject known to us, but, since the area of interior-point polynomial-time methods is extremely popular now and has attracted attention of many researchers, we understand that the comment is far from being complete. The most complete bibliography (over 1,200 entries) on interior-point methods was recently created by Dr. E. Kranich and is available via NETLIB.

## 1 Historical remarks

### 1.1 Classical interior-point methods

Polynomial-time interior-point methods are closely related to quite traditional areas of mathematical programming, meaning the interior penalty function approach combined with the Newton method for minimizing the resulting (penalized) objective (see the classical monograph of Fiacco and McCormick [FMcC 68]). The traditional analysis of this scheme, however, gives no theoretical understanding of the preferable types of penalties and reasonable handling the penalty parameter (recall that in 1968 almost nobody was interested in the key for this understanding complexity issues). The logarithmic barrier

$$F(x) = -\sum_i \ln f_i(x)$$

for a polytope

$$G = \{x \in \mathbf{R}^n \mid f_i(x) = a_i^T x - b_i \geq 0,\ i = 1, \dots, m\}$$

was used in the context of linear programming by Frisch in the early 1950s (see [Frs 56]).

In more specific sense, the "prehistory" of the polynomial-time interior-point methods originates from the paper of Dikin [Di 67] where the so-called Dikin method for LP problems was suggested. In this method, a given strictly feasible solution $y$ to the problem

$$\text{(LP)}: \qquad \text{minimize } c^T x \quad \text{s.t. } x \in G$$

is updated as follows: We compute the minimizer $x'$ of the objective over the ellipsoid $\{z \mid \langle F''(y)(z-y), z-y\rangle \leq 1\}$ defined by the logarithmic barrier for $G$, and the iterate of $x$ is obtained by a step from $x$ in the direction $x' - x$, the stepsize being a fixed fraction of the largest stepsize preserving the feasibility of the shifted point (compare with the first "half-step" in the parallel trajectories method §2.5). Dikin proved the convergence of this procedure and established the asymptotic rate of convergence (under the assumption that the problem is nondegenerate). To our knowledge, there are no polynomial-time results about this method; simultaneously, it looks similar to the polynomial-time interior-point algorithms, and it is reported that the practical performance of Dikin's algorithm is close to, say, that one of Karmarkar's method. Dikin's activity in this area was completely unknown in the western world and almost unknown in the USSR until the birth of modern polynomial-time interior-point methods.

### 1.2 Polynomial-time interior-point methods

The history of the methods begins with the landmark paper of Karmarkar [Ka 84], where the author's method for LP problems was suggested. The excellent complexity result of this paper, as well as the claim that the performance of the new method considerably overcomes that one of the simplex method, made this work a sensation and inspired very intensive further studies. Two years later, Renegar [Re 86] suggested a new (theoretically, even more efficient) polynomial-time interior-point method for LP, which appeared a quite traditional method of centers. This paper was especially important for clarifying the nature of polynomial-time interior-point methods for LP, and it predetermined further development in the field.

#### 1.2.1 Linear programming

Until now, the activity in the field of interior-point methods focuses mainly on linear programming. It seems to be impossible to mention most of the contributions to the area done by various researchers; our reference list (although far from being complete) does give some impression on the activity in the area. In what follows, we restrict ourselves to the minimal amount of comments compatible with the contents of the monograph.

Below, $n$ denotes the larger size of an LP (or an LCQP—linearly constrained convex quadratic) problem in the canonical form, and $L$ denotes the total length of the input data (in the case of problems with integer data) or the quantity $\ln(1/\varepsilon)$, if the problem under consideration is to be solved to the relative (with respect to an appropriate scale) accuracy $\varepsilon$ (in the latter case, the coefficients can be reals). To simplify the estimates, we express them in terms of the larger size only.

The method of Karmarkar solves an LP problem in $O(nL)$ steps of the total arithmetic cost $O(n^4L)$ (the basic version) or $O(n^{3.5}L)$ (the accelerated version). Renegar [Re 86] developed the first path-following polynomial-time method for LP with the efficiency estimate of $O(n^{1/2}L)$ steps ($O(n^{3.5}L)$ operations totally); independent, earlier similar method of centers were proposed by Sonnevend (along with results on ellipsoidal approximations and nonlinear applications) [So 85], however without any complexity estimates. Vaidya [Va 87] accelerated Renegar's method, with the aid of the Karmarkar speed-up; the resulting complexity estimate ($O(n^3L)$ operations) remains the best-known complexity bound for LP, provided that the traditional linear algebra technique is used. The same complexity estimate was independently and simultaneously obtained by Gonzaga in the paper [Go 87], where the first barrier path-following method for LP was developed. This complexity bound is valid for all the following interior-point methods for LP (all exceptions will be specified).

Most of the papers connected with the path-following approach are based on the method of centers. As already mentioned, the barrier path-following method originates from Gonzaga [Go 87]. The parallel trajectories method for LP and LCQP was developed by Nesterov [Ns 88b], [Ns 88c].

Primal-dual interior-point methods for LP with the complexity of $O(n^{1/2}L)$ steps were developed by Kojima, Mizuno, and Yoshise [KMY 87] and Monteiro and Adler [MA 87a]. Important contributions to this field belong to Carpenter, Choi, Lustig, Monma, Mulvey, and Shanno [CMS 88], [CLMS 90]. The potential reduction primal-dual $O(n^{1/2}L)$-step LP method was suggested by Todd and Ye [TY 87] and then improved by Ye [Ye 88a], [Ye 89].

### 1.2.2 Linearly constrained quadratic programming

This is the closest to LP class of nonlinear problems: As far as the path-following methods are concerned, the technique created for LP usually can be extended without essential difficulties to this class. Polynomial-time methods for LCQP were developed by Kapoor and Vaidya [KV 86], Ye and Tse [YT 89], and Mehrotra and Sun [MS 87]. As far as we know, the first $O(n^{1/2}L)$-step methods of $O(n^3L)$ arithmetic cost for LCQP were suggested by Goldfarb and Liu [GL 88], Monteiro and Adler [A 87b], and Nesterov [Ns 88b], [Ns88c].

### 1.2.3 Complementarity problems

A special area in the theory of interior-point methods is formed by studies devoted to the interior-point and path-following technique for complementarity problems, mainly linear ones (Kojima, Megiddo, Mizuno, Noma, Pardalos, Ye, and Yoshise [KMgY 88], [KMY 88], [KMgNY 90], [Ye 88b], [YP 89]). Polynomial-time interior-point methods for nonlinear complementarity problems were developed by Kojima, Megiddo, Mizuno, and Noma [KMgN 88], [KMzN 88a], [KMzN 88b], [KMgM 90], [KMgNY 90].

### 1.2.4 "Anticipated" behaviour of interior-point methods for LP and LCQP

Another special area is the "anticipated" behavior of the interior-point methods for LP problems. Computational experience demonstrates that the number $N$ of steps of, say, Karmarkar's method on real-world problems is essentially less than that one prescribed by the worst-case efficiency estimate. While the latter is proportional to the size $n$ of an LP problem, in practice, $N$ seems to be proportional to $\ln n$. Although no rigorous justifications of this phenomenon are known, there are some arguments that make it not too surprising. Indeed, assume that the search direction $\xi$ defined at a given step by the Karmarkar algorithm with line search, is random and uniformly distributed in the corresponding subspace (the latter subspace is provided by the Euclidean structure defined by the second-order differential of the barrier). It can be proved that under this assumption the amount at which the Karmarkar potential function is reduced at the step typically is $\Omega(n/\ln n)$ instead of the worst-case $\Omega(1)$, which results in the "anticipated" efficiency estimate of $O(L \ln n)$ steps instead of the worst-case estimate of $O(nL)$ steps (Mizuno, Nemirovsky, Todd, and Ye, [Nm 87], [To 89], [MTY 90a], [MTY 90b], [Ye 90b]). The main drawback of this approach is that we cannot point out (and cannot even prove the existence) of an a priori probabilistic distribution on the set of LP problems of the size $n$, which results in a "good" distribution of the search direction at each step. It is interesting to note that $O(L \ln n)$-anticipated behavior of potential reduction interior-point methods is established for the methods with the worst-case efficiency estimate $O(nL)$ only, while the anticipated behavior of the methods with $O(n^{1/2}L)$-worst case efficiency is $O(n^{1/4}L)$ (see [MTY 90a], [MTY 90b] for a detailed discussion of these issues).

We should also mention the work of Sonnevend, Stoer, and Zhao [SSZ 89], [SSZ 90], where it is proved that, for some *special* classes of LP problems, the worst-case efficiency estimate for path-following methods is better than $O(n^{1/2}L)$ iterations.

**1.2.5 Nonlinear convex problems**

There is a number of papers devoted to investigation of trajectories generated by logarithmic penalty functions (Sonnevend [So 85], who also gives results on centers of polyhedra; Jarre [Ja 87]; and Mehrotra and Sun [MS 88b]) within general convex context. To our knowledge, there were rather few results [Ja 87], [Ja 89a], [Ja 89b], [MS 88a], [MS 88b] on polynomial-time interior-point algorithms for essentially nonlinear problems.

In our opinion, the reason for relatively restricted activity in the field of polynomial-time interior-point methods for nonlinear convex programs is that the initial technique developed for LP in the seminal papers of Karmarkar [Ka 84] and Renegar [Re 86] heavily depends on the specific properties of LP.[1]

To extend interior-point polynomial-time methods on "more nonlinear" problems, new approaches are needed. The first results here dealt with convex quadratically constrained quadratic programs; path-following methods for these problems were developed by Jarre [Ja 87] and Mehrotra and Sun [MS 88a]. These authors also extended their approaches to general convex problems (see [Ja 89a], [Ja 89b], [MS 88b]); below, we give more detailed presentation of these general results. We should also mention a recent paper of Monteiro and Adler [MA 90] on separable problems and the papers of Alizadeh [Al 91a], [Al 91b], [Al 92] and Jarre [Ja 91] on semidefinite programming (compare with §5.4). The approach to explanation and design of interior-point polynomial-time methods underlying this monograph is based on the concepts of self-concordant functions and self-concordant barriers introduced by Nesterov [Ns 88b], [Ns 88c]. These ideas were developed in a number of papers of the authors (see [NN 88], [NN 89], [NN 90a], [NN 90b], [NN 90c], [NN 90d],

[1]Note, anyway, that the technique developed for LP can be quite straightforwardly used for semidefinite programming (see [Al 91a], [Al 91b]).

[NN 91a], [NN91b]; as far as polynomial path-following methods are concerned, basically, all applications presented in this monograph were already given in [NN 88]).

As already mentioned, general interior-point approaches to nonlinear convex problems were also developed by Mehrotra and Sun [MS 88b], Jarre [Ja 89a], [Ja 89b], Monteiro and Adler [MA 90], and Kortanek and Zhu [Z 90], [KZ 91]. Mehrotra and Sun deal with convex problems of the type

$$\text{minimize } f_0(x) \quad \text{s.t. } f_i(x) \le 0,\ i = 1, \dots, m,$$

with bounded feasible set, and make rather restrictive assumption that the functions $f_i$, $i = 0, \dots, m$, satisfy the *curvature condition*

$$\kappa^{-1} f_i''(y) \le f_i''(x) \le \kappa f_i''(y)$$

for all feasible $x, y$; here $\kappa$ is certain "curvature constant." Under this assumption, they establish polynomiality of a path-following Renegar-type method associated with the barrier $\sum_{i=1}^m \ln(-f_i)$. In fact, the curvature condition is very close to the notion of strong convexity and shares its main shortcoming: The numerical value of the curvature constant $\kappa$ usually depends not only on the analytical structure and sizes of the problem, but also on the diameter of the feasible domain, on numerical values of coefficients, and so forth.

The approach developed by Jarre [Ja 89a], [Ja 89b] is based on the "relative Lipschitz condition" (RLC): a convex $C^2$ function $f$ that is negative on an open convex domain $Q$ and tends to 0 along each sequence converging to a boundary point of the domain satisfies this condition (with constant $M$) if

$$\forall z \in \mathbf{R}^n\ \forall y \in Q\ \forall h \text{ with } h^T H(y) h \le 0.25(1 + M^{1/3})^{-2} :$$

$$|z^T \left( D^2 f(x+h) - D^2 f(x) \right) z| \le M \{h^T H(y) h\}^{1/2} z^T D^2 f(y) z;$$

here $H(y)$ is the Hessian of the function $F_f(\cdot) = -\ln(-f(\cdot))$ at $y$. Jarre demonstrates that the problem

$$\text{minimize } c^T x \quad \text{s.t. } f_i(x) \le 0,\ i = 1, \dots, m$$

associated with the constraints $f_i$ satisfying the RLC on the domains $\{f_i(x) < 0\}$ can be solved in polynomial time by the path-following method associated with the barrier $F(x) = \sum_i F_{f_i}(x)$. As found by Jarre [Ja 90], the RLC implies self-concordance: For a $C^3$ function $f$ satisfying the RLC with constant $M$, the function $F_f$ proves to be $O((M+1)^2)$-self-concordant barrier for the Lebesque set $\{x \mid f(x) \le 0\}$. This observation allows us to derive all constructions and complexity results based on the RLC from those based on the theory of self-concordance. In our opinion, the main disadvantage of the RLC, as compared to self-concordance, is that it seems much more difficult to verify the former property in the case of nonquadratic $f$. Besides this, the width of the class of convex domains that can be represented in terms of functions satisfying the RLC is not known.

The schemes of Monteiro and Adler [MA 90] and Kortanek and Zhu [Z90], [KZ 91] deal with convex problems involving linear equality constraints; Monteiro and Adler, in addition, assume the objective to be separable. Under appropriate assumptions on the objective, the authors prove polynomiality of the barrier-type path-following methods associated with the natural logarithmic barriers for the epigraph of the objective. As shown by den Hertog [D-H 92], the assumptions of these papers imply self-concordance of the corresponding barriers (at least in the case of $C^3$-smooth objectives), so that the schemes in question also are covered by the general self-concordance-based approach.

In all of the above approaches, to find a barrier for the domain $\{x \mid f(x) \leq 0\}$ defined by a smooth convex function, we should take as barrier the function $-\ln(-f(x))$; to provide "nice" properties of the barrier, it requires us to impose certain restrictions on $f$. It means that the structure of the barrier is a priory fixed. In this book, we proceed differently. We point out the desired property of the barrier, prove that, in principle, such a barrier does exist for every domain, and develop a technique for obtaining "computable" barriers. The latter approach seems to be more flexible, since we do not restrict ourselves to barriers of any specific type. For example, we feel free to use the function

$$F(t,x) = -\ln(t^2 - \| x \|_2^2)$$

as a self-concordant barrier for the second-order cone

$$\{(t,x) \mid f(t,x) = \| x \|_2 - t \leq 0\}.$$

It seems rather difficult to find a direct correspondence between the functions $f$ and $F$. Of course, an arbitrary $\vartheta$-self-concordant barrier $F$ can be represented via the logarithm of a concave function $f(x)$ as

$$F(x) = -\vartheta \ln(f(x)),$$

with

$$f(x) = \exp\{-\vartheta^{-1} F(x)\},$$

($f$ is concave by Proposition 1.3.2(iv)), but this possibility seems useless for developing the theory of the interior point methods, same as for constructing "computable" barriers.

**2.** What follows are specific bibliography comments to different parts of the text.

*Chapter* 1. The basic concepts of self-concordant function and self-concordant barrier were introduced in [Ns 88b], [Ns 88c]. The results presented in the chapter originate from [NN 89], [NN 90b].

*Chapter* 2. The notion of a self-concordant family was introduced in [NN 89], where on the basis of this notion the authors explained and extended onto the nonlinear case path-following methods previously developed for LP and LCQP (the barrier method [Go 87], [Ns 88b], [Ns 88c], the method of centers [Re4 86] and the primal and dual methods of parallel trajectories [Ns 88a], [Ns 89]).

As compared to [NN 89], the only new result of Chapter 2 is the complexity analysis of the barrier-generated path-following method with large-step strategy (see Proposition 2.2.5). The authors were pleased to find that similar result was independently obtained by den Hertog [D-H 92].

*Chapter* 3. Duality for convex programs involving "nonnegativity constraints" defined by a general-type convex cone in a Banach space is a relatively old (and, possibly, slightly forgotten by the mathematical programming community) part of convex analysis (see, e.g., [ET 76]). The corresponding general results, as applied to the case of conic problems (i.e., finite-dimensional problems with general-type nonnegativity constraints and *affine* functional constraints), form the contents of §3.2. To our knowledge, in convex analysis, there was no special interest to conic problems, and consequently to remarkable symmetric form of the aforementioned duality in this particular case. The only previous result in spirit of this duality known to us is the dual characterization of the Lovasz capacity number $\theta(\Gamma)$ of a graph (see [Lo 79]).

The general approach presented in Chapter 3 regarding explanation and extension onto the general convex case of the potential reduction interior-point methods known for LP (the method of Karmarkar [Ka 84], the projective method [Nm 87], [NN 90a], the primal-dual method of Todd and Ye [TY 87], [Ye 88a], [Ye 89]), was developed in [NN 90b], [NN 90d]. Interpretation of the generalized method of Karmarkar in terms of the Newton minimization of a self-concordant barrier for an unbounded domain (§3.3.4) is motivated by the devoted to the LP case paper of Bayer and Lagarias [BL 91].

The authors are greatly indebted to Professor Stephen Boyd, who attracted their attention to the generalized linear-fractional problem and stimulated by this the research summarized in §3.4.

*Chapter* 4. The results of this chapter originate from [NN 88], [NN 89], [NN 90b]. Section 4.1.2 is new. Several interesting self-concordant barriers for two-dimensional convex sets can be also found in [D-HJRT 92].

An important open question concerning barriers concerns decreasing the gap between the theoretically best possible value $O(n)$ of the parameter of a self-concordant barrier for an $n$-dimensional convex domain $G$ (see Theorem 1.5.1) and the parameters of "computable" barriers for $G$ (for example, in the case of a polytope $G$ defined by $m$ linear inequalities the parameter of the standard logarithmic barrier for $G$ is $m$). A very important result in this direction was obtained by Vaidya [Va 89]; the generalization presented in §4.5 of this result seems to be new.

*Chapter* 5. Most of the results presented here were announced in [NN 88]; complete proofs were given in [NN 89], [NN 90b]. Among these applications, the most attractive, in our opinion, are those related to quadratically constrained quadratic programming and to semidefinite programming. The latter area is now very popular among those involved into the design of interior-point methods for nonlinear convex programming (Alizadeh and Jarre [Al 91a], [Al 91b], [Al 92], [Ja 91]), same as among those interested in applications of these methods, mainly to various control theory problems (Balakrishnan, Baratt, Barmish, Boyd, Doyle, El-Ghaoui, Fan, Nekooie, Packard, Tits, Yang, Zhou, among others; see [BB 90], [BBr 91], [BBK 89], [BG 92], [BY 89], [DPZ 91], [Do 82], [FN 91], [FT 86], [FT 88], [FT 91], [FTD 91], [GB 86], [KR 91]); detailed presentation of control theory applications of semidefinite programming can be found in the monograph of Boyd, El-Ghaoui, Feron, and Balakrishnan [BGFB 93]. The most of "semidefinite reformulations" of convex programs presented in §5.4 are quite straightforward; the only exception is the *pd*-representation of the function "the sum of $k$ largest eigenvalues of a symmetric matrix"; the representation given in §5.4 (it was independently found by Alizadeh) is based on the "convex-concave" description of the function given by Overton and Womersley in [OW 91].

As far as generalized linear-fractional problem on the cone of positive semidefinite matrices is concerned (§5.4.4.7), many applications of this problem, same as a special interior-point *method of analytic centers* for it, are given by Boyd and El-Ghaoui [BG 92].

The problems connected with extremal ellipsoids (§5.5) were, from different viewpoints, studied by many authors. We should especially mention the improvement (by a factor of $O(m)$) of the complexity bound from Theorem 5.5.1 for the problem of finding the inscribed ellipsoid of the maximum volume obtained by Khachiyan and Todd in [KhT 90].

*Chapter* 6. To our knowledge, polynomial-time interior-point methods for variational inequalities previously were developed mainly for inequalities with affine operators and only for polyhedral feasible domains (see references in item 1.2.3). We should also mention the paper of Guler [Gu 90], where some general properties of the penalty-type paths associated with a nonlinear monotone complementarity problem are investigated. The contents of the chapter originates from [NN 91a].

*Chapter* 7. Acceleration of polynomial-type interior-point methods for LP and LCQP begins in the landmark paper of Karmarkar [Ka 84], where the first method of this family was suggested. As already mentioned, Vaidya and Gonzaga were the first to accelerate path-following methods for LP, which resulted in the best-known—cubic with respect to the size of an LP problem—efficiency estimate for LP. Karmarkar's type acceleration was used by most of the researchers dealt with LP and LCQP (see references in item 1.2.1). The schemes presented in Chapter 7 originate from [NN 89]; see also [NN 91b].

# Bibliography


[AB 89] K. M. ANSTREICHER AND R. A. BOSCH, *Long steps in a $O(n^3L)$ algorithm for linear programming*, Yale School of Management, New Haven, CT, 1989.

[AHU 76] A. V. AHO, J. K. HOPCROFT, AND J. D. ULLMAN, *The Design and Analysis of Computer Algorithms*, Addison–Wesley, Reading, MA, 1976.

[AKRV 86] I. ADLER, N. KARMARKAR, M. G. C. RESENDE, AND G. VEIGA, *An implementation of Karmarkar's algorithm for linear programming*, Report No. ORC 86-8, Operations Research Center, University of California, Berkeley, CA, 1989; Math. Programming, 44 (1989), pp. 297–336.

[AKRV 90] ——, *Data structures and programming techniques for the implementation of Karmarkar's algorithm*, ORSA J. Comput., 1 (1990), pp. 84–106.

[Al 91a] F. ALIZADEH, *Optimization over the positive-semidefinite cone: interior-point methods and combinatorial applications*, Report, Dedicated to J. B. Rosen on the Occasion of His 70th Birthday, Math. Programming, submitted.

[Al 91b] ——, *Combinatorial Optimization with Interior Point Methods and Semi-Definite Matrices*, Ph.D. thesis, Graduate School of the University of Minnesota, October 1991.

[Al 92] ——, *Optimization over the positive-semidefinite cone: Interior point methods and combinatorial optimization*, in Advances in Optimization and Parallel Computing, P. Pardalos, ed., North–Holland, Amsterdam, 1992.

[An 86] K. M. ANSTREICHER, *A monotonic projective algorithm for fractional linear programming*, Algorithmica, 1 (1986), pp. 483–498.

[An 89a] ——, *A combined Phase I–Phase 2 projective algorithm for linear programming*, Math. Programming, 43 (1989), pp. 209–223.

[An 89b] K. M. ANSTREICHER, *A combined Phase* I–*Phase* 2 *Scaled Potential Algorithm for Linear Programming*, CORE Discussion Paper 8939, CORE Catholic University of Louvan, Belgium, 1989.

[An 89c] ——, *A worst-case step in Karmarkar's algorithm*, Math. Oper. Res., 14 (1989), pp. 294–302.

[An 90] ——, *On Monotonicity in Scaled Potential Algorithm for Linear Programming*, Report, CORE, Catholic University of Louvan, Belgium, 1990.

[AW 90] K. M. ANSTREICHER AND P. WATTEYNE, *A Family of Search Directions for Karmarkar's Algorithm*, CORE Discussion Paper 9030, CORE Catholic University of Louvan, Belgium, 1990.

[Ba 86] E. R. BARNES, *A variation on Karmarkar's algorithm for solving linear programming problems*, Math. Programming, 36 (1986), pp. 174–182.

[BB 90] S. BOYD AND V. BALAKRISHNAN, *A regularity result for the singular values of a transfer matrix and a quadratically converging algorithm for computing its $L_\infty$-norm*, Systems Control Lett., 15 (1990), pp. 1–7.

[BBK 89] S. BOYD, V. BALAKRISHNAN, AND P. KABAMBA, *A bisection method for computing the $H_\infty$ norm of a transfer matrix and related problems*, Math. Control Signals Systems, 2 (1989), pp. 278–295.

[BBr 91] S. BOYD AND C. BARATT, *Linear Controller Design: Limits of Performance*, Prentice–Hall, Englewood Cliffs, NJ, 1991.

[BDDW 90] P. BOGGS, P. D. DOMICH, J. R. DONALDSON, AND C. WITZGALL, *Algorithmic enhancements to the method of centers for linear programming problems*, ORSA J. Comput., 1 (1990), pp. 159–171.

[BG 92] S. BOYD AND L. EL-GHAOUI, *Method of centers for minimizing generalized eigenvalues*, Linear Algebra Appl., in preparation.

[BGFB 93] S. BOYD, L. EL GHAOUI, E. FERON, AND V. BALAKRISHNAN, *Linear Matrix Inequalities in System and Control Theory*, in preparation.

[BH 66] B. TRONG LIEU AND P. HUARD, *La methode des centres dans un espace topologique*, Numer. Mat., 8 (1966), pp. 56–67.

[BL 86] D. A. BAYER AND J. C. LAGARIAS, *The nonlinear geometry of linear programming.* I: *Affine and projective scaling trajectories*; II: *Legendre transform coordinates and central trajectories*, AT&T Bell Laboratories, Murray Hill, NJ, 1986; Trans. Amer. Math. Soc., 314 (1989), pp. 499–581.

[BL 91] ——, *Karmarkar's algorithm and Newton's method*, Math. Programming, 50 (1991), pp. 291–330.

[B-TN 92] A. BEN-TAL AND A. NEMIROVSKII, *Interior Point Polynomial Time Methods for Truss Topology Design*, Research Report # 3/92, Faculty of Industrial Engineering and Management, The Israel Institute of Technology, Haifa 32000, Israel, June 1992.

[BY 89] S. BOYD AND Q. YANG, *Structured and simultaneous Lyapunov functions for system stability problems*, Internat. J. Control, 49 (1989), pp. 2215–2240.

[CLMS 90] T. J. CARPENTER, I. J. LUSTIG, J. M. MULVEY, AND D. F. SHANNO, *A primal-dual interior point method for convex nonlinear programs*, RUTCOR Research Report 25-90, Rutgers Center for Operations Research, Rutgers University, New Brunswick, NJ, 1990.

[CMS 88] C. CHOI, C. L. MONMA, AND D. F. SHANNO, *Further development of primal-dual interior point methods*, Report 60-88, Rutgers Center for Operations Research, Rutgers University, New Brunswick, NJ, 1990.

[Co 65] A. COBHAM, *The intrinsic computational difficulty of functions*, in Proc. 1964 Internat. Congress for Logic, Methodology and Philosophy of Science, 24-30, Y. Bar-Hillel, ed., North-Holland, Amsterdam, 1965.

[CW 86] D. COPPERSMITH AND S. WINOGRAD, *Matrix Multiplication via Arithmetic Progression*, Dept. of Math. Sci., IBM Thomas J. Watson Res. Center, Nov. 1986, preprint.

[D-H 92] D. DEN HERTOG, *Interior Point Approach to Linear, Quadratic and Convex Programming—Algorithms and Complexity*, Ph.D. thesis, Technishe Universiteit Delft, September 1992.

[D-HJRT 92] D. DEN HERTOG, F. JARRE, C. ROOS, AND T. TERLAKY, *A sufficient condition for self-concordance, with application to some classes of structured convex programming problems*, Math. Programming Ser. B, 1992, to appear.

[Di 67] I. I. DIKIN, *Iterative solution of problems of linear and quadratic programming*, Soviet Mat. Dokl., 8 (1967), pp. 674–675.

[Do 82] J. C. DOYLE, *Analysis of feedback systems with structured uncertainties*, IEEE Proc., 129 (1982), pp. 242–250.

[DPZ 91] J. DOYLE, A. PACKARD, AND K. ZHOU, *Review of* LFTs, LMIs, *and $\mu$*, in Proc. of the 30th Conference on Decision and Control, Brighton, England, December 1991.

[Ed 65] J. EDMONDS, *Paths, trees and flowers*, Canad. J. Math., 17 (1965), pp. 449–467.

[ET 76] I. EKELAND AND R. TEMAM, *Convex Analysis and Variational Problems*, North–Holland, Amsterdam, 1976.

[FMcC 68] A. V. FIACCO AND G. P. MCCORMICK, *Nonlinear Programming*, John Wiley, New York, 1968.

[FN 91] M. K. H. FAN AND B. NEKOOIE, *On minimizing the largest eigenvalue over an affine family of Hermitian matrices*, Linear Algebra Appl., submitted.

[Fr 88a] R. M. FREUND, *Projective transformations for interior point methods, Part* I: *Basic theory and linear programming*, Working Paper OR 179-88, Operations Research Center, Massachusetts Institute of Technology, Cambridge, MA, 1988.

[Fr 88b] ——, *Projective transformations for interior point methods, Part* II: *An algorithm for finding the weighted center of a polyhedral system*, Working Paper OR 180-88, Operations Research Center, Massachusetts Institute of Technology, Cambridge, MA, 1988.

[Fr 88c] ——, *Polynomial-time algorithms for linear programming based only on primal scaling and projected gradients of a potential function*, Research Report OR 182-88, Sloan School of Management, Massachusetts Institute of Technology, Cambridge, MA, 1988.

[Frs 56] R. FRISCH, *La resolution des problemes de programme lineaire par la methode du potential logarithmique*, Cahiers du Seminaire D'Econometrie, 4 (1956), pp. 7–20.

[FT 86] M. K. H. FAN AND A. TITS, *Characterization and efficient computation of the structured singular value*, IEEE Trans. Automat. Control, AC-31 (1986), pp. 734–743.

[FT 88] M. K. H. FAN AND A. TITS, *m-form numerical range and the computation of the structured singular value*, IEEE Trans. Automat. Control, AC-33 (1988), pp. 284–289.

[FT 91] ——, *A measure of worst-case $H_\infty$ performance and of largest acceptable uncertainty*, Systems Control Lett., to appear.

[FTD 91] M. K. H. FAN, A. TITS, AND J. DOYLE, *Robustness in the presence of mixed parametric uncertainty and unmodeled dynamics*, IEEE Trans. Automat. Control, 36 (1991), pp. 25–38.

[Ga 87] D. M. GAY, *A variant of Karmarkar's linear programming algorithm for problems in standard form*, Math. Programming, 37 (1987), pp. 81–90.

[GB 86] A. GALMIDI AND D. BARMISH, *The constrained Lyapunov problem and its application to robust output feedback stabilization*, IEEE Trans. Automat. Control, AC-31 (1986), pp. 410–419.

[GJ 79] M. R. GAREY AND D. S. JOHNSON, *Computers and Intractability: A Guide to the Theory of* NP-*Completeness*, W. H. Freeman, San Francisco, 1979.

[GL 88] D. GOLDFARB AND S. LIU, *An $O(n^3L)$ primal interior point algorithm for convex quadratic programming*, Tech. Report, Dept. of IEOR, Columbia University, New York, 1988.

[GLS 81] M. GROTSCHEL, L. LOVASZ, AND A. SCHRIJVER, *The ellipsoid method and its consequences in combinatorial optimization*, Combinatorica, 1 (1981), pp. 125–156.

[GLS 86] ——, *The Ellipsoid Method and Combinatorial Optimization*, Springer-Verlag, Bonn, Germany, 1986.

[GLS 88] ——, *Geometric Algorithms and Combinatorial Optimization*, Springer-Verlag, Bonn, Germany, 1988.

[GMSTW 86] P. E. GILL, W. MURRAY, M. A. SAUNDERS, J. A. TOMLIN, AND M. H. WRIGHT, *On projected Newton barrier methods for linear programming and an equivalence to Karmarkar's projective method*, Math. Programming, 36 (1986), pp. 183–209.

[Go 87] C. C. GONZAGA, *An algorithm for solving linear programming problems in $O(nL)$ operations*, Tech. Report, Dept. of Electrical Engineering and Computer Sciences, Univ. of California, Berkeley, CA, 1987; in Progress in Mathematical Programming: Interior-Point and Related Methods, N. Megiddo, ed., Springer-Verlag, Berlin, New York, 1989, pp. 1–28.

[Go 88] C. C. GONZAGA, *Polynomial Affine Algorithms for Linear Programming*, Report ES-139/88, COPPE, Universidade Federal do Rio de Janeiro, 1988.

[Go 89] ——, *Conical projection algorithm for linear programming*, Math. Programming, 43 (1989), pp. 151–173.

[Gr 60] B. GRUNBAUM, *Partitions of mass-distributions and of convex bodies by hyperplanes*, Pacific J. Math., 10 (1960), pp. 1257-1261.

[GT 89] C. C. GONZAGA AND M. J. TODD, *An $O(n^{1/2}L)$-iteration large-step primal-dual affine algorithm for linear programming*, Tech. Report No. 862, School of Operations Research and Industrial Engineering, Cornell University, Ithaca, NY, 1989.

[GTr 89] E. G. GOLSHTEIN AND N. V. TRETYAKOV, *Modified Lagrangeans*, Moscow, Nauka, 1989. (In Russian.)

[Gu 90] O. GÜTLER, *Existence of interior points and interior paths in nonlinear monotone complementarity problems*, Working Paper No. 90-21, College of Business Administration, The University of Iowa, Iowa City, IA, October 1990.

[GV 86] G. DE GHELLINK AND J.-P. VIAL, *A polynomial Newton method for linear programming*, Algorithmica, 1 (1986), pp. 425–454.

[GV 87] ——, *An extension of Karmarkar's algorithm for problems in standard form*, Math. Programming, 37 (1987), pp. 81–90.

[II 86] M. IRI AND H. IMAI, *Multiplicative barrier function method for linear programming*, Algorithmica, 1 (1986), pp. 455–482.

[Ja 87] F. JARRE, *On the convergence of the method of analytic centers when applied to convex quadratic programs*, Report No. 35, Dec. 1987, Schwerpunktprogramm der Deutchen Forschungsgemeinschaft–Anwendungsbezogene Optimierung und Steuerung, revised Feb. 1988.

[Ja 89a] ——, *The method of analytic centers for smooth convex programs*, dissertation, Institut ftr Angewandte Mathematik und Statistics, Universitat Wtrzburg, 1989.

[Ja 89b] ——, *On the method of analytic centers for a class of convex programs*, in Lecture Notes in Mathematics, vol. 1405, S. Dolecki, ed., 1989; in Optimization—Fifth French German Conference Castel Novel 1988.

[Ja 90] F. JARRE, *Interior point methods for convex programming*, Report SOL 90-16, Dept. of Operations Research, Stanford University, Stanford, CA, 1990; Appl. Math. Optim., to appear.

[Ja 91] ——, *An interior point method for minimizing the maximum eigenvalue of a linear combination of matrices*, Report SOL 91-8, Dept. of Operations Research, Stanford University, Stanford, CA, 1991.

[Ka 84] N. KARMARKAR, *A new polynomial-time algorithm for linear programming*, Combinatorica, 4 (1984), pp. 373–395.

[KhT 90] L. G. KHACHIYAN AND M. J. TODD, *On the complexity of approximating the maximal inscribed ellipsoid for a polytope*, Tech. Report No. 893, School of Operations Research and Industrial Engineering, College of Engineering, Cornell University, Ithaca, NY, 1990.

[KhTE 88] L. G. KHACHIYAN, S. P. TARASOV, AND A. I. ERLICH, *The inscribed ellipsoid method*, Dokl. Akad. Nauk SSSR, Vol. 298, 1988. (Engl. trans. Soviet Math. Dokl.)

[KMgM 90] M. S. KOJIMA, N. MEGIDDO, AND S. MIZUNO, *A general framework of continuation methods for complementarity problems*, Research Report, IBM Almaden Research Center, San Jose, CA, 1990.

[KMgN 88] M. S. KOJIMA, N. MEGIDDO, AND T. NOMA, *Homotopy continuation methods for nonlinear complementarity problems*, Math. Oper. Res., 16 (1991), pp. 754–774.

[KMgNY 90] M. S. KOJIMA, N. MEGIDDO, T. NOMA, AND A. YOSHISE, *A unified approach to interior point algorithms for complementarity problems*, Research Report, RJ 7493 (70008), IBM Almaden Research Center, San Jose, CA, 1990.

[KMgY 88] M. S. KOJIMA, N. MEGIDDO, AND Y. YE, *An interior point potential reduction algorithm for the linear complementarity problem*, Research Report, RJ 6486, IBM Almaden Research Center, San Jose, CA, 1988.

[KMY 87] M. S. KOJIMA, S. MIZUNO, AND A. YOSHISE, *A primal-dual interior point algorithm for linear programming*, Report No. B-193, Department of Information Sciences, Tokyo Institute of Technology, Tokyo, 1987; in Progress in Mathematical Programming: Interior-Point and Related Methods, N. Megiddo, ed., Springer-Verlag, Berlin, New York, 1989, pp. 29–48.

[KMY 88] M. S. KOJIMA, S. MIZUNO, AND A. YOSHISE, *An $O(n^{1/2}L)$ iteration potential reduction algorithm for linear complementarity problems*, Research Reports on Information Sciences B-217, Dept. of Information Sciences, Tokyo Institute of Technology, Tokyo, 1988.

[KMY 89] ——, *A polynomial time algorithm for linear complementarity problems*, Math. Programming, 44 (1989), pp. 1–26.

[KMzN 88a] M. S. KOJIMA, S. MIZUNO, AND T. NOMA, *A new continuation method for complementarity problems with uniform P-functions*, Math. Programming, 43 (1988), pp. 107–113.

[KMzN 88b] ——, *Limiting behavior of trajectories generated by a continuation method for monotone complementarity problems*, Research Report, Department of Information Sciences, Tokyo Institute of Technology, Tokyo, 1988.

[Kr 72] R. M. KARP, *Reducibility among combinatorial problems*, in Complexity of Computer Computations, R. E. Miller and J. W. Thatcher, eds., Plenum Press, New York, 1972.

[Kr 75] ——, *On the computational complexity of combinatorial problems*, Networks, 5 (1975), pp. 45–68.

[KR 91] P. P. KHARGONEKHAR AND M. A. ROTEA, *Mixed $H_2/H_\infty$ control: A convex optimization approach*, IEEE Trans. Automat. Control, 36 (1991), pp. 824–837.

[KV 86] S. KAPOOR AND P. M. VAIDYA, *Fast algorithms for convex quadratic programming and multicommodity flows*, in Proc. of the 18th Annual ACM Symposium on Theory of Computing, 1986.

[KZ 91] K. O. KORTANEK AND J. ZHU, *A polynomial barrier algorithm for linearly constrained convex programming problems*, Revised Working Paper Series No. 90-17, College of Business Administration, The University of Iowa, Iowa City, IA, 1991.

[LMC 89] I. J. LUSTIG, J. M. MULVEY, AND T. J. CARPENTER, *The formulation of stochastic programs for interior methods*, Tech. Report SOR-89-16, Department of Civil Engineering and Operations Research, Princeton University, Princeton, NJ, 1989.

[LMS 89] I. J LUSTIG, R. MARSTEN, AND D. F. SHANNO, *Computational experience with a primal-dual interior point method for linear programming*, Report SOR-89-17, Department of Civil

Engineering and Operations Research, Princeton University, Princeton, NJ, 1989.

[LMS 90] I. J LUSTIG, R. MARSTEN, AND D. F. SHANNO, *On implementation Mehrotra's predictor-corrector interior point method for linear programming*, Report SOR-90-3, Department of Civil Engineering and Operations Research, Princeton University, Princeton, NJ, 1990.

[Lo 79] L. LOVASZ, *On the Shannon capacity of a graph*, IEEE Trans. Inform. Theory, 25 (1979), pp. 355–381.

[LSG 89] I. J. LUSTIG, D. F. SHANNO, AND J. W. GREGORY, *The primal-dual interior point method on the Cray supercomputer*, Department of Civil Engineering and Operations Research, Princeton University, Princeton, NJ, 1989.

[Lu 88] I. J. LUSTIG, *Feasibility issues in an interior point method for linear programming*, Tech. Report SOR 88-9, Department of Civil Engineering and Operations Research, Princeton University, Princeton, NJ, 1989; Math. Programming, 49 (1991), pp. 145–162.

[MA 87a] R. D. C. MONTEIRO AND I. ADLER, *An $O(n^3)$ primal-dual interior point algorithm for linear programming*, Report ORC 87-4, Operations Research Center, Dept. of Operations Research, University of California, Berkeley, CA, 1987.

[MA 87b] ——, *An $O(n^3)$ primal-dual interior point algorithm for convex quadratic programming*, Report ORC 87-15, Operations Research Center, Dept. of Operations Research, University of California, Berkeley, CA, 1987.

[MA 89a] ——, *Interior path-following primal-dual algorithms. Part* I: *Linear programming*, Math. Programming, 44 (1989), pp. 27–42.

[MA 89b] ——, *Interior path-following primal-dual algorithms. Part* II: *Convex quadratic programming*, Math. Programming, 44 (1989), pp. 43–66.

[MA 90] ——, *An extension of Karmarkar-type algorithm to a class of convex separable programming problems with global linear rate of convergence*, Math. Oper. Res., 15 (1990), pp. 408–422.

[MAR 88] R. D. C. MONTEIRO, I. ADLER, AND M. G. C. RESENDE, *A polynomial time primal-dual affine scaling algorithm for linear* *and convex quadratic programming and its power series extension*, Tech. Report ESRC 88-8, Engineering Systems Research Center, University of California, Berkeley, CA, 1988; Math. Oper. Res., 15 (1990), pp. 191–214.

[McSMS 89] K. A. McSHANE, C. L. MONMA, AND D. F. SHANNO, *An implementation of a primal-dual interior point method for linear programming*, ORSA J. Comput., 1 (1989), pp. 70–83.

[Mg 86] N. MEGIDDO, *Pathways to the optimal set in linear programming*, Tech. Report RJ 5295, IBM Almaden Research Center, San Jose, CA, 1986; in Progress in Mathematical Programming: Interior-Point and Related Methods, N. Megiddo, ed., Springer-Verlag, Berlin, New York, 1989, pp. 131–158.

[Mh 89] S. MEHROTRA, *On finding a vertex solution using interior point methods*, Tech. Report 89-17, Department of Industrial Engineering and Management Sciences, Northwestern University, Evanston, IL, 1989.

[Mh 90] ——, *On implementation of a (primal-dual) interior point method*, Tech. Report 90-03, Department of Industrial Engineering and Management Sciences, Northwestern University, Evanston, IL, 1990.

[MM 87] C. L. MONMA AND A. J. MORTON, *Computational experience with a dual affine variant of Karmarkar's method for linear programming*, Tech. Report, Bell Communications Research, 1989, Oper. Res. Lett., 6 (1987), pp. 225–253.

[MS 87] S. MEHROTRA AND J. SUN, *An algorithm for convex quadratic programming that requires $O(n^{3.5}L)$ arithmetic operations*, Tech. Report 87-24, Dept. of Industrial Engineering and Management Sciences, Northwestern University, Evanston, IL, 1987; Math. Oper. Res., 15 (1990), pp. 342–363.

[MS 88a] ——, *A method of analytic centers for quadratically constrained convex quadratic programs*, Tech. Report 88-01, Dept. of Industrial Engineering and Management Sciences, Northwestern University, Evanston, IL, 1988; SIAM J. Numer. Anal., 28 (1991), pp. 529–544.

[MS 88b] ——, *An interior point algorithm for solving smooth convex programs based on Newton's method*, Tech. Report 88-08, Dept. of Industrial Engineering and Management Sciences, Northwestern University, Evanston, IL, 1988.

[MSSPB 90] R. E. MARSTEN, M. J. SALTZMAN, D. F. SHANNO, G. S. PIERCE, AND J. F. BALLINTIJN, *Implementation of a dual affine interior point algorithm for linear programming*, ORSA J. Comput., 1 (1990), pp. 287–297.

[MTY 89] S. MIZUNO, M. J. TODD, AND Y. YE, *Anticipated behavior of path-following algorithms for linear programming*, Tech. Report No. 878, School of Operations Research and Industrial Engineering, Cornell University, Ithaca, NY, 1989.

[MTY 90a] ——, *Anticipated behavior of long-step algorithms for linear programming*, Research Report 24, Department of Industrial Engineering and Management, Tokyo Institute of Technology, Tokyo, 1990.

[MTY 90b] ——, *On adaptive-step primal-dual interior-point algorithms for linear programming*, Tech. Report No. 944, School of Operations Research and Industrial Engineering, Cornell University, Ithaca, NY, 1990.

[Mz 89] S. MIZUNO, *A new polynomial time method for a linear complementarity problem*, Research Report 16, Department of Industrial Engineering and Management, Tokyo Institute of Technology, Tokyo, 1989.

[Na 86] J. L. NAZARETH, *Homotopy techniques in linear programming*, Algorithmica, 1 (1986), pp. 529–535.

[Na 87] ——, *Pricing criteria in linear programming*, Report PAM-382, Center for Pure and Applied Mathematics, University of California, Berkeley, CA, 1987; in Progress in Mathematical Programming: Interior-Point and Related Methods, N. Megiddo, ed., Springer-Verlag, Berlin, New York, 1989, pp. 105–129.

[Nm 81] A. S. NEMIROVSKY, *Effective solution methods for variational inequalities with monotone operators*, Ekonomika i Matem. Metody, 17 (1981), pp. 344–359. (In Russian; English trans. Matekon: Translations of Russian and East European Math. Economics.)

[Nm 87] ——, *On an algorithm of Karmarkar's type*, Izvestija AN SSSR, Tekhnitcheskaya kibernetika, 1987, No. 1, pp. 105–118. (In Russian.)

[NN 88] YU. E. NESTEROV AND A. S. NEMIROVSKY, *Polynomial-time barrier methods in convex programming*, Ekonomika i Matem.

Metody, 24 (1988) No. 7, pp. 1084–1091. (In Russian; English trans. Matekon: Translations of Russian and East European Math. Economics.)

[NN 89] YU. E. NESTEROV AND A. S. NEMIROVSKY, *Self-concordant functions and polynomial time methods in convex programming*, USSR Acad. Sci. Central Economic & Mathematical Institute, Moscow, 1989.

[NN 90a] ——, *Projective method for convex quadratic knapsack problem*, Tech. Report, Dept. of Math. Sci., IBM Thomas J. Watson Research Center, Yorktown Heights, NY, Jan. 1990.

[NN 90b] ——, *Logarithmically homogeneous barriers and polynomial time interior point methods based on Lyapunov's functions*, USSR Acad. Sci. Central Economic & Mathematical Institute, Moscow, 1990, preprint.

[NN 90c] ——, *Optimization over positive semidefinite matrices*, Manual, USSR Acad. Sci. Central Economic & Mathematical Institute, Moscow, 1990.

[NN 90d] ——, *Conic duality and its applications in convex programming*, USSR Acad. Sci. Central Economic & Mathematical Institute, Moscow, 1990, preprint; Optim. Software, 1 (1992), pp. 10–31.

[NN 91a] ——, *Path-following polynomial time algorithm for monotone variational inequalities*, USSR Acad. Sci. Central Economic & Mathematical Institute, Moscow, 1991, preprint.

[NN 91b] ——, *Acceleration and parallelization of the path-following interior point methods for a linearly constrained quad ratic programming problem*, SIAM J. Optim., 1 (1991), pp. 548–564.

[Ns 83] YU. E. NESTEROV, *An $O(1/k^2)$-rate of convergence method for smooth convex functions minimization*, Dokl. Akad. Nauk SSSR, 269 (1983), pp. 543–547. (In Russian; English trans. Soviet Math. Dokl.)

[Ns 88a] ——, *The method for linear programming which requires $O(n^3L)$ operations*, Ekonomika i Matem. Metody, 24 (1988), pp. 174–176. (In Russian; English trans. Matekon: Translations of Russian and East European Math. Economics.)

[Ns 88b] ——, *Polynomial time methods in linear and quadratic programming*, Izvestija AN SSSR, Tekhnitcheskaya kibernetika, No. 3, 1988, pp. 324–326. (In Russian.)

[Ns 88c] YU. E. NESTEROV, *Polynomial time iterative methods in linear and quadratic programming*, Voprosy kibernetiki, Moscow, 1988, pp. 102–125. (In Russian.)

[Ns 88d] ——, *A general approach to the design of optimal methods for smooth convex functions minimization*, Ekonomika i Matem. Metody, 24 (1988), pp. 509–517. (In Russian; English trans. Matekon: Translations of Russian and East European Math. Economics.)

[Ns 89] ——, *Dual polynomial time algorithms for linear programming*, Kibernetika, No. 1, 1989, pp. 34–54. (In Russian.)

[NY 79] A. S. NEMIROVSKY AND D. B. YUDIN, *Informational Complexity and Efficient Methods for Solution of Convex Extremal Problems*, Nauka, 1978. (In Russian; English trans. John Wiley, New York, 1983.)

[OR 70] J. M. ORTEGA AND W. C. RHEINBOLDT, *Iterative Solution of Nonlinear Equations in Several Variables*, Academic Press, New York, 1970.

[Ov 90] M. L. OVERTON, *Large scale optimization of eigenvalues*, NYU Computer Science Department Report, # 505, New York, May 1990.

[OW 91] M. L. OVERTON AND R. S. WOMERSLEY, *Optimality conditions and duality theory for minimizing sums of the largest eigenvalues of symmetric matrices*, Tech. Report 566, Computer Science Department, New York University, New York, 1991.

[Po 73] V. M. POLTEROVICH, *Economical equilibrium and optimum*, Ekonomika i Matem. Metody, 9 (1973), pp. 835–845. (In Russian; English trans. Matekon: Translations of Russian and East European Math. Economics.)

[PM 78] V. M. POLTEROVICH AND I. I. MITJUSHIN, *A criterion of monotonicity of a demand function*, Ekonomika i Matem. Metody, 14 (1978), pp. 122–128. (In Russian; English trans. Matekon: Translations of Russian and East European Math. Economics.)

[PZPLB 91] A. PACKARD, K. ZHOU, P. PANDEY, J. LEONHARDSON, AND G. BALAS, *Optimal, constant I/O similarity scaling for full-information and state-feedback control problems*, Department of Mechanical Engineering, University of California, Berkeley, CA, Oct. 1991, preprint.

[Re 86] J. Renegar, *A polynomial time algorithm, based on Newton's method, for linear programming*, Research Report, School of Operations Research and Industrial Engineering, Cornell University, Ithaca, NY, 1986; Math. Programming, 40 (1988), pp. 59–93.

[RS 88] J. Renegar and M. Shub, *Simplified complexity analysis for Newton* LP *methods*, Report No. 807, School of Operations Research and Industrial Engineering, Cornell University, Ithaca, NY, 1988.

[RV 89] C. Roos and J.-P. Vial, *Long steps with the logarithmic penalty barrier function in linear programming*, Report 89-44, Department of Mathematics and Computer Science, Delft University of Technology, Delft, the Netherlands, 1989.

[Sh 87] N. Z. Shor, *Quadratic optimization problems*, Izvestija AN SSSR, Tekhnitcheskaya kibernetika, No. 1, 1987. (In Russian.)

[ShD 85] N. Z. Shor and A. Davydov, *On a method for bounding in quadratical problems*, Kibernetika, No. 2, 1985, pp. 48–54. (In Russian.)

[So 85] G. Sonnevend, *An analytic center for polyhedrons and new classes of global algorithms for linear (smooth, convex) programming*, in Proc. 12th IFIP Conf. on System Modelling and Optimization, Hungary, Budapest, 1985; Lecture Notes in Control and Inform. Sci., Vol. 84, Springer-Verlag, Berlin, New York, 1985.

[SSZ 89] G. Sonnevend, J. Stoer, and G. Zhao, *On the complexity of following the central path of linear programs by linear extrapolation*, Math. Oper. Res., 63 (1989), pp. 19–31.

[SSZ 90] ——, *On the complexity of following the central path of linear programs by linear extrapolation* II, University of Wurtzburg, 1990, manuscript.

[TB 86] M. J. Todd and B. P. Burrell, *An extension of Karmarkar's algorithm for linear programming using dual variables*, Algorithmica, 1 (1986), pp. 409–424.

[To 88a] M. J. Todd, *On Anstreicher's combined Phase* I–*Phase* II *projective algorithm for linear programming*, Tech. Report No. 776, School of Operations Research and Industrial Engineering, College of Engineering, Cornell University, Ithaca, NY, 1988.

[To 88b] M. J. TODD, *Improved bounds and containing ellipsoids in Karmarkar's linear programming algorithm*, Math. Oper. Res., 13 (1988), pp. 650–659.

[To 89] ——, *Anticipated behavior of Karmarkar's algorithm*, Tech. Report No. 879, School of Operations Research and Industrial Engineering, College of Engineering, Cornell University, Ithaca, NY, 1989.

[To 90] ——, *Projected scaled steepest descent in Kojima–Mizuno–Yoshise's potential reduction algorithm for the linear complementarity problem*, Tech. Report No. 950, School of Operations Research and Industrial Engineering, College of Engineering, Cornell University, Ithaca, NY, 1990.

[TW 89] M. J. TODD AND Y. WANG, *On combined Phase* I–*Phase* II *projective methods for linear programming*, Tech. Report No. 877, School of Operations Research and Industrial Engineering, College of Engineering, Cornell University, Ithaca, NY, 1989.

[TY 87] M. J. TODD AND Y. YE, *A centered projective algorithm for linear programming*, Tech. Report No. 763, School of Operations Research and Industrial Engineering, College of Engineering, Cornell University, Ithaca, NY, 1987; Math. Oper. Res., 15 (1990), pp. 508–529.

[Va 87] P. M. VAIDYA, *An algorithm for linear programming which requires $O(((m+n)n^2+(m+n)^{1.5}n)L)$ arithmetic operations*, Tech. Report, AT&T Bell Laboratories, Murray Hill, NJ, 1987; Math. Programming, 47 (1990), pp. 175–202.

[Va 89] ——, *A new algorithm for minimizing convex functions over convex sets*, Tech. Report, AT&T Bell Laboratories, Murray Hill, NJ, 1989.

[VMF 86] R. J. VANDERBEI, M. S. MEKETON, AND B. A. FREEDMAN, *A modification of Karmarkar's linear programming algorithm*, Algorithmica, 1 (1986), pp. 395–407.

[Ye 87] Y. YE, *Interior Algorithms for Linear, Quadratic and Linearly Constrained Convex Programming*, Ph.D. dissertation, Dept. of Engineering–Economic Systems, Stanford University, Stanford, CA, 1987.

[Ye 88a] ——, *A class of potential functions for linear programming*, Working Paper Series No. 88-13, Dept. of Management Sci., The University of Iowa, Iowa City, IA, September 1988.

[Ye 88b] Y. YE, *A further result on the potential reduction algorithm for the P-matrix linear complementarity problem*, Dept. of Management Sci., The University of Iowa, Iowa City, IA, 1988, in Advances in Optimization and Parallel Computing, P. M. Pardalos, ed., North-Holland, Amsterdam, 1992, pp. 310–316.

[Ye 89] ——, *An $O(n^3L)$ potential reduction algorithm for linear programming*, Tech. Report, Department of Management Sci., The University of Iowa, Iowa City, IA, 1989.

[Ye 90a] ——, *Line search in potential reduction algorithms for linear programming*, College of Business Administration, The University of Iowa, Iowa City, IA, 1990, manuscript.

[Ye 90b] ——, *Anticipated behavior of affine scaling algorithms for linear programming*, Working Paper Series No. 90-1, College of Business Administration, The University of Iowa, Iowa City, IA, January 1990.

[Ye 90c] ——, *Complexity analysis on Karmarkar's algorithm*, Working Paper Series No. 90-5, College of Business Administration, The University of Iowa, Iowa City, IA, March 1990.

[YK 87] Y. YE AND M. KOJIMA, *Recovering optimal dual solutions in Karmarkar's polynomial algorithm for linear programming*, Math. Programming, 39 (1987), pp. 305–317.

[YP 89] Y. YE AND P. PARDALOS, *A class of linear complementarity problems solvable in polynomial time*, Dept. of Management Sci., The University of Iowa, Iowa City, IA, 1988; Linear Algebra Appl., 152 (1991), pp. 1–10.

[YT 89] Y. YE AND E. TSE, *An extension of Karmarkar's projective algorithm for convex quadratic programming*, Math. Programming, 44 (1989), pp. 157–179.

[Z 90] J. ZHU, *A path-following algorithm for a class of convex programming problems*, Working Paper Series No. 90-14, Department of Management Sci., The University of Iowa, Iowa City, IA, 1990.

[ZTD 90] Y. ZHANG, R. A. TAPIA, AND J. E. DENNIS, *On the superlinear and quadratic convergence of primal-dual interior point linear programming algorithms*, Working Paper, Department of Mathematical Sciences, Rice University, Houston, TX, 1990.

# Index