SIAM Rev.
Vol. 40, No. 2, pp. 228–264, June 1998

OPTIMIZATION PROBLEMS WITH PERTURBATIONS: A GUIDED TOUR*

J. FRÉDÉRIC BONNANS† AND ALEXANDER SHAPIRO‡

Abstract. This paper presents an overview of some recent, and significant, progress in the theory of optimization problems with perturbations. We put the emphasis on methods based on upper and lower estimates of the objective function of the perturbed problems. These methods allow one to compute expansions of the optimal value function and approximate optimal solutions in situations where the set of Lagrange multipliers is not a singleton, may be unbounded, or is even empty. We give rather complete results for nonlinear programming problems and describe some extensions of the method to more general problems. We illustrate the results by computing the equilibrium position of a chain that is almost vertical or horizontal.

Key words. sensitivity analysis, parameterized optimization, directional differentiability, quantitative stability, duality, expansion of optimal solutions, semi-infinite programming, semidefinite programming, second-order optimality conditions

AMS subject classifications. 49K40, 90C31, 26B25, 41A58, 65K05, 49M45

PII. S0036144596302644

1. Introduction.

In this paper we present a survey of some results on stability and sensitivity analysis of optimization problems which are subject to perturbations. We consider problems of the form

$$\underset{x \in X}{\text{Min}}\, f(x,u) \quad \text{subject to } x \in \Phi(u). \qquad (P_u) \tag{1.1}$$

In most examples considered in this paper, X is a Banach space, while the perturbation parameter u can be a scalar, a finite-dimensional vector, or even an element of an appropriate normed or metric space U. We investigate continuity and differentiability properties of the optimal value $v(u)$ and the set $S(u)$ of optimal solutions of (P_u) considered as functions of the parameter u.

In order to proceed in our analysis we need a constructive method for describing the feasible set $\Phi(u)$. We say that the feasible set is defined by *abstract constraints* if it is given in the following form:

$$\Phi(u) = \{x : G(x,u) \in K\}, \tag{1.2}$$

where K is a closed convex subset of a Banach space Y and $G : X \times U \to Y$. This includes the situations where K is a closed convex cone, in which case we speak of *cone constraints*.

An important particular case is when the feasible set is *finitely constrained*, i.e., is of the form

$$\Phi(u) = \{x : g_i(x,u) = 0,\ i = 1,\dots,q;\ \ g_i(x,u) \le 0,\ i = q+1,\dots,p\}\,. \tag{1.3}$$

Then $G(x,u) = (g_1(x,u),\dots,g_p(x,u)) : X \times U \to \mathbb{R}^p$ and $K = \{0\} \times \mathbb{R}^{p-q}_-$. In that case, if in addition X is finite-dimensional, (P_u) becomes a *nonlinear programming*

*Received by the editors April 26, 1996; accepted for publication (in revised form) June 12, 1997. http://www.siam.org/journals/sirev/40-2/30264.html

†INRIA-Rocquencourt, Domaine de Voluceau, B. P. 105, 78153 Rocquencourt, France, and Ecole Polytechnique, France (frederic.bonnans@inria.fr).

‡School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, GA 30332-0205 (ashapiro@isye.gatech.edu).

problem. Another important case is when Y is the linear space of $n \times n$ symmetric matrices and $K \subset Y$ is the cone of positive semidefinite matrices. This example corresponds to the so-called semidefinite programming. Another example where the "cone constraints" formulation appears in a natural way is the example of semi-infinite programming. That is, let T be a compact metric space, for every u, let $G(x, u)$ be a mapping from X into the space $Y = C(T)$ of continuous real-valued functions on T, and let

$$K := \{\phi \in C(T) : \phi(t) \leq 0 \ \ \forall\, t \in T\}$$

be the set of nonpositive valued functions in $C(T)$. For such defined $G(x, u)$ and K, equation (1.2) is equivalent to

$$\Phi(u) = \{x : g_t(x, u) \leq 0, \ t \in T\}, \tag{1.4}$$

where $g_t(x, u) = G(x, u)(t)$. In case the feasible set is defined by an infinite number of constraints, as in (1.4), and the space X is finite-dimensional, the corresponding program (P_u) is called a semi-infinite program.

Investigation of max-min problems goes back to the classical work of Chebyshev on uniform approximations by algebraic polynomials. It was also one of the driving forces behind the systematic development of nonsmooth analysis in the sixties and seventies. There are numerous studies where various aspects of max-min optimization are discussed (see, e.g., [28, 30]). In this paper we survey some recent results on first- and especially second-order analysis of parameterized optimization problems. In a limited survey paper we have to be selective, and several relevant and important areas (e.g., optimal control problems) are left out or only briefly mentioned. We also do not pretend to give a complete account of a historical development of the ideas. Some important work was done in the seventies in the former Soviet Union, in particular by Levitin [54, 55] (see also [29], and [56] for additional references). Unfortunately, at the time, it was only fragmentarily known and had little or no impact on development of the subject in the West.

The idea of studying *first*-order differentiability properties of the optimal value function by deriving upper and lower estimates for the corresponding directional derivatives goes back, at least, to Danskin [28]. Quite surprisingly it was discovered by Gollan [37] and Gauvin and Janin [35] that in some cases first-order differentiability properties (directional derivatives) of the optimal value function are closely related to a second-order analysis of the corresponding optimization program. On the other hand differentiability properties of the optimal solutions were first obtained by applying the classical implicit function theorem to the first-order optimality conditions written in a form of (nonlinear) equations [33]. This approach is discussed at length and relevant references can be found in Fiacco [32]. The implicit function theorem approach was also applied to a situation where optimal solutions are nondifferentiable in [45] and [12]. The equations-based approach was extended and generalized in works of Robinson [74, 75], where sensitivity analysis of variational inequalities (generalized equations in the terminology of Robinson) was developed.

In recent years substantial advances in our understanding of asymptotic behavior of optimal solutions were made by using a somewhat different approach. The main idea of that approach is based on *second*-order upper and lower estimates of the (optimal) value function. When these estimates are sufficiently close to each other, they allow one to derive a second-order expansion of the optimal value function that, in turn, gives a formula for an expansion of the optimal (nearly optimal) solutions.

The obtained first-order expansion of the optimal solutions is stated in terms of an auxiliary optimization problem. This idea, implicit in [37], is already evident in [79] and was developed in [8, 13, 16, 20, 22, 44, 80, 89]. In this paper we mainly follow this approach. In the case of nonlinear, semidefinite, and semi-infinite programming, this approach leads to quite complete results, while partial extensions are available for optimal control problems.

1.1. Basic notation and terminology. $\mathbb{R}^n$: n-dimensional Euclidean space,
$\mathbb{R}^n_+ = \{x \in \mathbb{R}^n : x_i \geq 0,\ i = 1, \ldots, n\}$: nonnegative orthant,
$\mathbb{R}^n_- = -\mathbb{R}^n_+ = \{x \in \mathbb{R}^n : x_i \leq 0,\ i = 1, \ldots, n\}$: nonpositive orthant,
$x \cdot y = \sum_{i=1}^n x_i y_i$: scalar product of $x, y \in \mathbb{R}^n$,
$\nabla g(x)$: gradient of the function $g : \mathbb{R}^n \to \mathbb{R}$ at the point $x \in \mathbb{R}^n$,
X^*: dual of the normed space X,
$\langle \alpha, x \rangle = \alpha(x)$: value of the linear functional $\alpha \in X^*$ on $x \in X$,
$B_X(x, r) = \{x' \in X : \|x' - x\| < r\}$: open ball of radius $r > 0$ centered at x,
$B_X = B_X(0, 1)$: the unit ball in X,
$[x] = \{tx : t \in \mathbb{R}\}$: linear space generated by vector x,
2^X: the set of subsets of X,
$\Psi : X \to 2^Y$: a multifunction (point-to-set mapping), mapping X into the set of subsets of Y,
$\mathrm{dom}\Psi = \{x \in X : \Psi(x) \neq \emptyset\}$: the domain of the multifunction Ψ,
$\mathrm{range}\Psi = \{y \in Y : y \in \Psi(x),\ x \in X\}$: the range of the multifunction Ψ,
$\mathrm{gph}\Psi = \{(x, y) \in X \times Y : y \in \Psi(x),\ x \in X\}$: the graph of the multifunction Ψ,
$\Psi^{-1}(y) = \{x \in X : y \in \Psi(x)\}$: graph inverse of the multifunction Ψ,
$I(x, u) = \{i : g_i(x, u) = 0,\ i = q + 1, \ldots, p\}$: set of active-at-$(x, u)$ inequality constraints,
$L(x, \lambda, u) = f(x, u) + \langle \lambda, G(x, u) \rangle$: the Lagrangian function,
$\Lambda(x, u)$: the set of Lagrange multipliers of (P_u) at the point $x \in S(u)$,
$L^g(x, \lambda_0, \lambda, u) = \lambda_0 f(x, u) + \langle \lambda, G(x, u) \rangle$: generalized Lagrangian function,
$\mathrm{val}(P)$: optimal value of the program (P),
$\mathcal{S}(P)$: set of optimal solutions of the program (P),
$v(u) = \mathrm{val}(P_u) = \inf_{x \in \Phi(u)} f(x, u)$: optimal value function,
$S(u) = \mathcal{S}(P_u) = \arg\min_{x \in \Phi(u)} f(x, u)$: set of optimal solutions,
$\bar{x}(u) \in S(u)$: an optimal solution,
$\mathrm{dist}(x, S) = \inf_{z \in S} \|x - z\|$: distance from the point $x \in X$ to set $S \subset X$,
$Dg(x) \in \mathcal{L}(X, Y)$: derivative (Gâteaux, Hadamard, or Fréchet, depending on the context) of the mapping $g : X \to Y$ at the point $x \in X$,
$D^2 g(x) : X \to \mathcal{L}(X, Y)$: second-order derivative of the mapping g at the point x,
$D^2 g(x)(h, h) = (D^2 g(x) h) h$: quadratic form corresponding to $D^2 g$,
$D_x g(x, u) \in \mathcal{L}(X, Y)$: partial derivative of $g : X \times U \to Y$,
$Dg(x, u)(h, d) = D_x g(x, u) h + D_u g(x, u) d$,
$g'(x, d) = \lim_{t \to 0^+} [g(x + td) - g(x)]/t$: directional derivative of g at x in direction d,
$f'_+(x, d) = \limsup_{t \to 0^+} [f(x + td) - f(x)]/t$: upper directional derivative of $f : X \to \mathbb{R}$,
$\partial f(x) = \{\alpha \in X^* : f(y) - f(x) \geq \langle \alpha, y - x \rangle\ \forall y \in X\}$: subdifferential of the convex function $f : X \to \mathbb{R} \cup \{-\infty\} \cup \{+\infty\}$,
$C^- = \{\alpha \in X^* : \langle \alpha, x \rangle \leq 0\ \forall x \in C\}$: polar (negative dual) of the convex cone $C \subset X$,
$\Sigma^- = \{x \in X : \langle \alpha, x \rangle \leq 0\ \forall \alpha \in \Sigma\}$: polar (negative dual) of the convex cone $\Sigma \subset X^*$,
$a \preceq_C b$: order relation imposed by the cone C, i.e., $b - a \in C$,
$\mathbb{R}_+(S) = \{tx : x \in S,\ t \geq 0\}$: cone generated by the set S,
$S^\perp = \{\alpha \in X^* : \langle \alpha, x \rangle = 0\ \forall x \in S\}$: orthogonal set to $S \subset X$,

$\mathrm{Sp}(S) = \mathbb{R}_+(S - S)$: linear space generated by $S \subset X$,
$T_S(x) = \{h \in X : \mathrm{dist}(x + th, S) = o(t),\ t \geq 0\}$: tangent cone to the convex set S at the point $x \in S$,
$T_S^2(x, h) = \{w \in X : \mathrm{dist}(x + th + \frac{1}{2}t^2 w, S) = o(t^2),\ t \geq 0\}$: second-order tangent set,
$N_S(x) = \{\alpha \in X^* : \langle \alpha, z - x\rangle \leq 0\ \forall z \in S\}$: normal cone to the set $S \subset X$ at $x \in S$,
$N_\Omega(\alpha) = \{x \in X : \langle \omega - \alpha, x\rangle \leq 0\ \forall \omega \in \Omega\}$: normal cone to the set $\Omega \subset X^*$ at $\alpha \in \Omega$,
$\mathrm{int}(S)$: interior of the set S,
$\mathrm{conv}(S) = \{x : x = \sum_{i=1}^m \lambda_i x_i,\ x_i \in S,\ \lambda_i > 0,\ \sum_{i=1}^m \lambda_i = 1\}$: convex hull of the set S,
$\mathrm{core}(S) = \{x \in S : \forall x' \in X,\ \exists \varepsilon > 0,\ \forall t \in [-\varepsilon, \varepsilon],\ x + tx' \in S\}$,
$\sigma(\lambda, S) = \sup_{x \in S}\langle \lambda, x\rangle$: the support function of the set S,
$C(T)$: Banach space of continuous functions $\psi : T \to \mathbb{R}$ defined on the compact metric space T and equipped with the sup-norm $\|\psi\| = \sup_{\tau \in T} |\psi(\tau)|$.

For $\varepsilon \geq 0$ we say that $\bar{x}$ is an ε-optimal solution of (P_u) if $\bar{x} \in \Phi(u)$ and $f(\bar{x}, u) \leq v(u) + \varepsilon$. Consider a mapping $g : X \to Y$. We say that g is directionally differentiable, at x, in the Hadamard sense if the directional derivative $g'(x, d)$ exists for all d and, moreover,

$$g'(x, d) = \lim_{\substack{t \to 0^+ \\ d' \to d}} \frac{g(x + td') - g(x)}{t}. \tag{1.5}$$

Note that if g is directionally differentiable, at x, in the Hadamard sense, then $g'(x, d)$ is continuous in d and if, in addition, X is finite-dimensional, then $g(x + h) = g(x) + g'(x, h) + o(\|h\|)$. For a discussion and comparisons of various concepts of directional differentiability, see [9, 67, 82].

Consider a multifunction $\Psi : X \to 2^Y$. It is said that Ψ is upper Lipschitz at a point $\bar{x} \in X$, if $\Psi(x) \subset \Psi(\bar{x}) + c\|x - \bar{x}\| B_Y$ for some $c > 0$ and all x in a neighborhood of $\bar{x}$. The multifunction Ψ is said to be closed at a point $x \in X$, if $x_n \to x$, $y_n \in \Psi(x_n)$, and $y_n \to y$ imply that $y \in \Psi(x)$. It is said that Ψ is closed if it is closed at every point of X. Note that Ψ is closed iff its graph $\mathrm{gph}\Psi$ is a closed subset of $X \times Y$.

2. The chain problem. Let us now introduce a simple problem in static mechanics, known as the chain or catenary problem, on which some of the theoretical results presented here will be illustrated. We will deal with situations where the classical regularity results do not hold. For instance, in one of them, the set of Lagrange multipliers is empty.

Consider a chain with m rigid links ($m > 1$), each of length 1. Suppose that the two endpoints of the chain are fixed. Assuming that the mass of each link is concentrated in its middle, we consider the problem of computing the equilibrium position of the chain. A formal description of the problem follows. Denote by (y_k, z_k) the increment, horizontally and vertically, along the link. The links are numbered from 1 to m. We assume that one endpoint is fixed at (0,0). The position of the other endpoint is $\sum_{i=1}^m (y_i, z_i)$ and is constrained to be $(y^e(u), z^e(u))$, where $u \in \mathbb{R}$ is viewed as the perturbation parameter.

The potential energy of the chain is given by

$$E(y, z) = \frac{1}{2}\sum_{k=1}^m \left(\sum_{i=1}^k z_i + \sum_{i=1}^{k-1} z_i\right) = \sum_{k=1}^m \alpha_k z_k, \tag{2.1}$$

where $\alpha_k = m - k + 1/2$ for $k = 1, \ldots, m$. The equilibrium position of the chain can be calculated by minimizing the potential energy subject to involved physical constraints. This leads to the following optimization problem:

$$\begin{aligned} &\operatorname{Min}_{y,z} E(y,z) \\ &\text{subject to } \; y_k^2 + z_k^2 - 1 = 0, \; k = 1,\dots,m, \; \textstyle\sum_{k=1}^m (y_k, z_k) = (y^e(u), z^e(u)). \end{aligned} \tag{2.2}$$

This example has been discussed by various authors, e.g., in the context of numerical algorithms [57] or as an illustration of optimality conditions [92] in finite-dimensional optimization. The above two references cover the situation when the solution is stable with respect to a perturbation of the position of the endpoints. Suppose now that for $u = 0$ the distance between the two endpoints is exactly m. Then the unperturbed problem has a unique feasible point (the chain is a straight line) which is therefore the unique optimal solution. However, the corresponding minimum is not always stable under a small perturbation of u (the set of feasible points vanishes when the distance between the two endpoints increases). These limit cases are perhaps those for which the perturbation analysis is the most interesting. For instance, let us consider the two situations in which the chain is either vertical or horizontal in the null position (i.e., when $u = 0$). That is, assume that either $y^e(u) = 0$, $z^e(u) = m - u$, or $y^e(u) = m - u$, $z^e(u) = 0$, and let one of the endpoints make a small step toward the other; i.e., let u increase from zero to a small positive number. Is it possible then to compute an expansion of the position of the chain?

As an application of the results reviewed here, we will show that the answer is positive. Note that, although these two cases may seem very similar, they are of fundamentally different nature. The chain is a model of a mechanical system, in which the force due to the weight of each link should be compensated by the forces applied at the endpoints of each link by the neighboring links. The relation of equilibrium of forces coincides with the first-order optimality system. In the case of the vertical chain, the set of Lagrange multipliers is nonempty but unbounded (and so is the set of mechanical forces for which the equilibrium is satisfied). In the case of the horizontal chain, however, the forces between links are horizontal, while the weight forces are vertical. Therefore, no mechanical equilibrium is possible. Expressed in terms of optimization theory, this means that the set of Lagrange multipliers is empty. The ability to compute an expansion of the optimal solution with an empty set of Lagrange multipliers is a striking example of the strength of the tools presented here.

Related to the lack of regularity of these problems is the fact that in both cases, the variation of the optimal solution is of the order of the square root of the perturbation parameter. This can be seen easily in the case of a two-links chain. Take, for instance, the case of a vertical chain. Let us denote by $\bar{y}_k(u), \bar{z}_k(u)$, $k = 1,\dots,m$, an optimal solution of (2.2), and consider the case of two links, i.e., $m = 2$. Then for $u > 0$ we have that the feasible set of (2.2) consists of two symmetric points, each one giving an optimal solution, and $\bar{y}_1(u) = \sqrt{u - u^2/4}$. We see here that (2.2) has an optimal solution $(\bar{y}(u), \bar{z}(u))$ which is not Lipschitz continuous at $u = 0$ and that this solution is of order $O(\sqrt{u})$. As we shall see later, such behavior of $(\bar{y}(u), \bar{z}(u))$ persists in cases when the chain has more than two links and $(\bar{y}(u), \bar{z}(u))$ can be expanded in terms of $\sqrt{u}$.

3. Lagrange multipliers and constraint qualifications.

3.1. First-order optimality systems.

In this section we assume that the feasible set $\Phi(u)$ is defined by constraints and that $f(x,u)$ and $G(x,u)$ are continuously differentiable. If x is a local solution of (P_u), then certain relations between the differential properties of the objective function f and the boundary of the feasible set must hold. In order to motivate the following development, let us give a mechanical interpretation of problem (P_u), assuming for the moment that $X = \mathbb{R}^n$ and $\Phi(u)$ is of the form (1.3). Let us view the objective function as a potential energy to be minimized. The mechanical force deriving from it is $-\nabla_x f(x,u)$. Each inequality constraint is viewed as the nonpenetration of a rigid body of equation $g_i(x,u) \geq 0$. With this constraint is associated a repulsion force which is normal to the surface of the body. If $\nabla_x g_i(x,u)$ is nonzero, then the outward normal to the body is $-\nabla_x g_i(x,u)$. Taking into account the fact that the force should be null if the point is not in contact with the body, we may write the equilibrium of forces as

$$\left\{ \begin{array}{l} \nabla_x f(x,u) + \sum_{i=1}^p \lambda_i \nabla_x g_i(x,u) = 0, \\ g_i(x,u) = 0,\ i = 1,\dots,q, \\ g_i(x,u) \leq 0,\ \lambda_i \geq 0,\ \lambda_i g_i(x,u) = 0,\ i = q+1,\dots,p. \end{array} \right. \tag{3.1}$$

This is the celebrated *first-order optimality system.* Note that this way of deriving the optimality system is opposite to the classical argument, where one analyses the outward normal directions to the feasible set rather than to the complementary set (both points of view are equivalent only under regularity assumptions).

If (x,λ) is a solution of the first-order optimality system, then λ is called a *Lagrange multiplier.* The set of Lagrange multipliers is a (possibly empty) polyhedron. We say that a local solution of (P_u) is *qualified* whenever the set of associated Lagrange multiplier is not empty. A simple example of a convex problem with a nonqualified solution is $\mathrm{Min}_x\{x : \ x^2 \leq 0\}$. F. John [46] proved that with any local solution of a finitely constrained problem is associated at least one *generalized Lagrange multiplier*, defined as a nonzero vector $(\lambda_0, \lambda_1, \dots, \lambda_p)$ satisfying the *generalized optimality system*

$$\left\{ \begin{array}{l} \lambda_0 \nabla_x f(x,u) + \sum_{i=1}^p \lambda_i \nabla_x g_i(x,u) = 0, \\ g_i(x,u) = 0,\ i = 1,\dots,q, \\ g_i(x,u) \leq 0,\ \lambda_i \geq 0,\ \lambda_i g_i(x,u),\ i = q+1,\dots,p, \ \ \lambda_0 \geq 0. \end{array} \right. \tag{3.2}$$

A generalized multiplier with $\lambda_0 = 0$ is said to be *singular.* With a nonsingular generalized multiplier is associated the Lagrange multiplier $(\lambda_0)^{-1}(\lambda_1,\dots,\lambda_p)$. Therefore, John's result may be restated as follows: with any local solution of a finitely constrained problem is associated either a Lagrange multiplier or a singular multiplier. Equivalently, a sufficient condition for qualification is that the set of singular multipliers is empty. The latter is a condition that depends only on the parameterization of the feasible set and is equivalent to the Mangasarian–Fromovitz constraint qualification condition [61]:

$$\left\{ \begin{array}{ll} \text{the gradients} & \nabla_x g_i(x,u), i = 1,\dots,q, \ \text{ are linearly independent,} \\ \quad \exists h \in \mathbb{R}^n : & h \cdot \nabla_x g_i(x,u) = 0, i = 1,\dots,q, \\ & h \cdot \nabla_x g_i(x,u) < 0 \ \forall i \in I(x,u), \end{array} \right. \tag{3.3}$$

where $I(x,u)$ denotes the set of active inequality constraints at (x,u). If all constraints are of equality (resp., inequality) type, then (3.3) reduces to the more familiar condition that $D_x G(x,u)$ is onto (resp., $\exists h \in \mathbb{R}^n : \ G(x,u) + D_x G(x,u)h < 0$).

We now turn to the problem with abstract constraints. The first-order generalized optimality system

$$D_x L^g(x,\lambda_0,\lambda,u) = 0, \ \ G(x,u) \in K, \ \ \lambda \in N_K(G(x,u)), \ \ \lambda_0 \geq 0, \tag{3.4}$$

where $L^g(x, \lambda_0, \lambda, u) = \lambda_0 f(x, u) + \langle \lambda, G(x, u) \rangle$, is a natural generalization of (3.2) to problems with abstract constraints. If $\lambda_0 \neq 0$, then one can take $\lambda_0 = 1$ and hence (3.4) becomes the following first-order optimality system:

$$D_x L(x, \lambda, u) = 0, \quad G(x, u) \in K, \quad \lambda \in N_K(G(x, u)). \tag{3.5}$$

Note that if K is a convex closed *cone* and $y \in K$, then $N_K(y) = K^- \cap \{y\}^\perp$. Therefore, in that case the optimality condition $\lambda \in N_K(G(x, u))$ is equivalent to $\lambda \in K^-$ and $\langle \lambda, G(x, u) \rangle = 0$.

We now discuss some qualification conditions that apply to problems with abstract constraints.

3.2. A general qualification condition. The following *regularity condition* is due to Robinson [72]:

$$0 \in \text{int}\ \{G(x, u) + D_x G(x, u) X - K\}. \tag{3.6}$$

If the feasible set is defined by a finite number of constraints in the form (1.3), then regularity conditions (3.3) and (3.6) are equivalent. The above condition recovers many special cases of interest. When K is a singleton (i.e., in the case of equality constraints), (3.6) reduces to the condition that $D_x G(x, u)$ is onto. When K has a nonempty interior, (3.6) reduces (see, e.g., [68, Prop. 3.9(f)]) to the condition

$$\exists h \in X: \quad G(x, u) + D_x G(x, u) h \in \text{int}(K).$$

The latter applies in particular to semi-infinite programming problems, with $\Phi(u)$ defined in the form (1.4) (see [84]). These two characterizations are very similar to those given before for (3.3) when all constraints are of equality (resp., inequality) type.

An interesting case is when $\Phi(u)$ is in the form (1.2) with Y being the Cartesian product of two Banach spaces Y_1 and Y_2, and $K = \{0\} \times K_2 \subset Y_1 \times Y_2$, where K_2 is a closed convex subset of Y_2 with nonempty interior. Then $G(x, u) = (G_1(x, u), G_2(x, u))$, with $G_i(x, u) \in Y_i$, $i = 1, 2$. In this case (3.6) is equivalent to the condition

$$\begin{aligned} &D_x G_1(x, u) \text{ is onto;} \\ &\exists h \in X : D_x G_1(x, u) h = 0, \quad G_2(x, u) + D_x G_2(x, u) h \in \text{int}(K_2). \end{aligned} \tag{3.7}$$

Yet another particular form of the constraints is useful when dealing with optimal control problems. Consider the *abstract optimal control problem*

$$\operatorname*{Min}_{x \in X} f(x, u) \quad \text{subject to } x \in \mathcal{C} \quad \text{and} \quad \mathcal{G}(x, u) \in \mathcal{K}, \tag{3.8}$$

where X is a space of controls, $\mathcal{C}$ is a closed convex subset of X, representing control constraints, and $\mathcal{K}$ is a closed convex subset of Banach space $\mathcal{Y}$; the relation $\mathcal{G}(x, u) \in \mathcal{K}$ might represent state constraints. This is a particular form of (1.1) with the feasible set given in the form (1.2), as is seen by setting $Y := X \times \mathcal{Y}$, $G(x, u) := (x, \mathcal{G}(x, u))$, and $K := \mathcal{C} \times \mathcal{K}$. The regularity condition in this case turns out to be equivalent to

$$0 \in \text{int}\ \{\mathcal{G}(x, u) + D_x \mathcal{G}(x, u)(\mathcal{C} - x) - \mathcal{K}\}.$$

If in addition $\mathcal{K}$ has a nonempty interior, then regularity is equivalent to a classical condition (e.g., [15])

$$\exists x' \in \mathcal{C} : \mathcal{G}(x, u) + D_x \mathcal{G}(x, u)(x' - x) \in \text{int}(\mathcal{K}).$$

Let us mention some more abstract conditions equivalent to (3.6) that may be found in the literature. Because the core of a convex set includes its interior, (3.6) implies

$$0 \in \text{core}\,\{G(x,u) + D_x G(x,u)X - K\}. \tag{3.9}$$

It is not difficult to see that (3.9) is equivalent to the relation $\mathcal{A}(x,u) = Y$, where $\mathcal{A}(x,u)$ is defined by

$$\mathcal{A}(x,u) := \mathbb{R}_+[G(x,u) + D_x G(x,u)X - K]. \tag{3.10}$$

It happens that condition $\mathcal{A}(x,u) = Y$ implies (3.6). Therefore, conditions (3.6), (3.9) and $\mathcal{A}(x,u) = Y$, in fact, are equivalent [95]. The above implication may be proved by considering the multifunction $\mathcal{M} : \mathbb{R} \times X \to 2^Y$ defined by $\mathcal{M}(a,h) = D_x G(x,u)h + a(G(x,u) - K)$, if $a \geq 0$, and $\mathcal{M}(a,h) = \emptyset$, if $a < 0$. Condition $\mathcal{A}(x,u) = Y$ means that $\mathcal{M}$ is onto, i.e., range$(\mathcal{M}) = Y$. It follows then from the generalized open mapping theorem (see Theorem 3.1 below) that $0 \in Y$ is an interior point of the set $\mathcal{M}([0,1] \times X)$ and hence (3.6) follows.

THEOREM 3.1 (generalized open mapping theorem [70], [73, Thm. 1]). *Let X and Y be Banach spaces, and let $\Psi : X \to 2^Y$ be a closed convex multifunction (i.e.,* gphΨ *is a closed convex subset of $X \times Y$). Let $y \in$* int(range Ψ)*. Then for every $x \in \Psi^{-1}(y)$ and all $r > 0$, it follows that $y \in \text{int}\,\Psi\,(B_X(x,r))$.*

The following combination of results by [20, Part I, Prop. 4.1], [64, 71], [95, Thm. 4.1] states that *regularity is a qualification condition.*

PROPOSITION 3.2. *Let x be a locally optimal solution of (P_u) and $\mathcal{A}(x,u)$ be given in* (3.10)*. Then* (i) *if $\mathcal{A}(x,u) = Y$ (i.e.,* (3.6) *holds), then the set of Lagrange multipliers is nonempty and bounded, and there exists no singular multiplier;* (ii) *if $\mathcal{A}(x,u)$ is not dense in Y, then there exists a singular multiplier.*

Note that we do not have an alternative as before: it may happen that there exist no generalized Lagrange multipliers. This can occur only if $\mathcal{A}(x,u)$ is dense in Y but is not equal to Y.

Regularity is also related to stability with respect to perturbations of the feasible set, as it follows from the next result which is a consequence of the Robinson–Ursescu [73], [91] stability theorem (cf. [72, Corollary 1]).

PROPOSITION 3.3. *Suppose that the constraint qualification* (3.6) *holds at (x^0, u^0). Then for all (x,u) in a neighborhood of (x^0,u^0), one has*

$$\text{dist}(x, \Phi(u)) = O\big(\text{dist}(G(x,u), K)\big). \tag{3.11}$$

Note that it follows from (3.11) that for all u in a neighborhood of u^0, dist$(x^0, \Phi(u)) < +\infty$ and hence $\Phi(u)$ is nonempty; i.e., the system $G(x,u) \in K$ has a solution. Consequently, if the constraint mapping G is sufficiently parameterized, i.e., of the form $G(x,u_1) - u_2$, where $u_2 \in Y$, then the converse of the above proposition is also true. That is, in that case (3.11) implies Robinson's constraint qualification (3.6), [26].

The stability property (3.11), called metric regularity, is particularly useful in deriving upper estimates for the optimal value function. Indeed, for a direction $d \in U$ let us consider a path $u(t) := u^0 + td + o(t)$ in the parameter space U. Suppose that we want to obtain upper estimates of the optimal value function $v(u(t))$ by constructing a path $x(t) \in X$ that is *feasible* with respect to the considered perturbations, i.e., $x(t) \in \Phi(u(t))$. We say that $h \in X$ is a (first-order) *feasible direction*, relative to the direction $d \in U$, if there exists $r(t) = o(t)$ such that $x^0 + th + r(t) \in \Phi(u(t))$.

It is easily checked that a necessary condition for h to be a feasible direction is

$$DG(x^0, u^0)(h, d) \in T_K(G(x^0, u^0)). \tag{3.12}$$

This condition is obviously sufficient if (3.11) holds. Therefore, by Proposition 3.3, (3.12) characterizes feasible directions whenever the constraints are regular.

3.3. Directional regularity. We continue the analysis of the set of directions that are feasible with respect to perturbations of the form $u(t) = u^0 + td + o(t)$, with $d \in U$. Whereas regularity is a condition that does not depend on the perturbation, we may search conditions depending on d, under which (3.12) characterizes the set of feasible directions.

The following condition, due to Gollan [37], applies to finitely constrained problems in $\mathbb{R}^n$,

$$\left\{ \begin{array}{lll} \text{the gradients} & \nabla g_i(x^0, u^0), i = 1, \ldots, q, & \text{are linearly independent,} \\ \exists h \in \mathbb{R}^n : & (h, d) \cdot \nabla g_i(x^0, u^0) = 0, i = 1, \ldots, q, & \\ & (h, d) \cdot \nabla g_i(x^0, u^0) < 0 \; \forall i \in I(x^0, u^0). & \end{array} \right. \tag{3.13}$$

A natural extension that applies to general constraints is [20, Part I]:

$$0 \in \operatorname{int} \{G(x^0, u^0) + DG(x^0, u^0)(X \times \mathbb{R}_+(d)) - K\}. \tag{3.14}$$

These conditions may be interpreted as regularity conditions with respect to the set $\{(x, t) : G(x, u^0 + td) \in K, \ t \geq 0\}$. It is clear that they explicitly depend on the chosen direction d, while the regularity conditions (3.3) and (3.6) (applied at (x^0, u^0)) do not. We refer to (3.13) and (3.14) as *directional regularity* condition because of the statement (c) in the following proposition.

PROPOSITION 3.4 (see [20, Part I]). (i) *Regularity implies directional regularity (i.e.,* (3.6) *implies* (3.14)*).* (ii) *The latter is equivalent to* (3.13) *in the case of finitely many constraints in* $\mathbb{R}^n$*.* (iii) *If directional regularity holds, then* h *is a first-order feasible direction iff it satisfies* (3.12).

Note that directional regularity is not a qualification condition (it does not imply existence of a Lagrange multiplier). It has useful characterizations similar to those of (3.6) when the constraints have particular forms. In the case of a product form discussed just before (3.7), such a characterization is

$$\begin{array}{ll} D_x G_1(x^0, u^0) \text{ is onto;} & \exists h \in X, \ \exists \varepsilon > 0 : DG_1(x^0, u^0)(h, d) = 0, \\ & G_2(x^0, u^0) + DG_2(x^0, u^0)(h, \varepsilon d) \in \operatorname{int}(K_2). \end{array} \tag{3.15}$$

In the case of the abstract optimal control problem when $\mathcal{K}$ has a nonempty interior, we get the characterization

$$\exists x' \in \mathcal{C}, \ \exists \varepsilon > 0 : \mathcal{G}(x^0, u^0) + D\mathcal{G}(x^0, u^0)(x' - x^0, \varepsilon d) \in \operatorname{int}(\mathcal{K}).$$

4. Analysis of the optimal value function.

4.1. Problems with unperturbed feasible sets. In this section we discuss first-order differentiability properties of the optimal value function $v(u)$. We start our analysis by considering the case when the feasible set is unperturbed.

EXAMPLE 4.1. *Let* $f(x, u) = xu$*, with* $x, u \in \mathbb{R}$*, and* $\Phi(u) = [-1, 1]$ *for all* u*. It is easily seen that* $v(u) = -|u|$*. Although the feasible set here is unperturbed and compact, the optimal value function is not differentiable at* $u = 0$*.*

This example shows how easily, starting with a smooth data, we ended up with the nondifferentiable optimal value function. Although $v(u)$ is nondifferentiable at zero, it nevertheless possesses directional derivatives in both positive and negative directions. Such *directionally differentiable* behavior is typical for the optimal value function. Quite a complete characterization of first-order differentiability properties of the optimal value function can be given when the feasible set is unperturbed, i.e., $\Phi(u) = \Phi$ for all $u \in U$, as shown in the following result, essentially due to Danskin [28].

THEOREM 4.1. *Let X be a metric space and U be a normed space. Suppose that for all $x \in X$ the function $f(x, \cdot)$ is differentiable and that $f(x, u)$ and $D_u f(x, u)$ are continuous on $X \times U$, and let Φ be a compact subset of X. Then the optimal value function $v(u) := \inf_{x\in\Phi} f(x, u)$ is Hadamard directionally differentiable and*

$$v'(u, d) = \operatorname*{Min}_{x\in S(u)} D_u f(x, u)d. \tag{4.1}$$

It follows from (4.1) that $v'(u^0, d)$ is linear in $d \in U$ if $f(\cdot, u^0)$ has a unique minimizer x^0 over Φ. In that case the optimal value function is Hadamard differentiable at u^0 and $Dv(u^0) = D_u f(x^0, u^0)$.

Application to semi-infinite programming. As an example let us briefly outline how the above result can be used to derive first-order optimality conditions in semi-infinite optimization (cf. [30, 69]). Consider the problem

$$\operatorname*{Min}_{x\in\mathbb{R}^n} f(x) \quad \text{subject to} \;\; g_t(x) \le 0, \; t \in T,$$

where $f(\cdot)$ is a continuously differentiable function, T is a compact metric space, the function $g_t(\cdot)$ is differentiable for every $t \in T$, and $g_t(x)$ and $\nabla g_t(x)$ are continuous jointly in x and t.

Let x^0 be a locally optimal solution of the above semi-infinite problem. Let $T(x^0) = \{t \in T : g_t(x^0) = 0\}$ be the set of active-at-x_0 constraints. We assume that $T(x^0) \neq \emptyset$; otherwise the problem, at least locally, is unconstrained. Note that then $T(x^0) = \arg\max_{t\in T} g_t(x^0)$. Let us observe that if x^0 is a locally optimal solution of the above problem, then the max-function

$$h(x) := \max\left\{ f(x) - f(x^0), \; \max_{t\in T} g_t(x) \right\} \tag{4.2}$$

attains its local minimum (equal 0) at x^0. Setting $g_0(x) := f(x) - f(x^0)$, we may write $h(x) = \max_{t\in T'} g_t(x)$, where $T' = T \cup \{0\}$. This max-function $h(x)$ can be considered as an optimal value function with t being the optimization variable. By Danskin's theorem, $h(x)$ is directionally differentiable and

$$h'(x^0, d) = \max\left\{ d \cdot \nabla f(x^0), \; \max_{t\in T(x^0)} d \cdot \nabla g_t(x^0) \right\}. \tag{4.3}$$

Optimality of x^0 implies that the directional derivative $\eta(d) := h'(x^0, d)$ is nonnegative for all $d \in \mathbb{R}^n$. By (4.3) the directional derivative $\eta(d)$ is a maximum of linear functions and hence is convex. Therefore, $\eta(d) \ge 0$ for all d iff $0 \in \partial\eta(0)$. It also follows from (4.3) that $\eta(\cdot)$ is the support function of the set formed by vectors $\nabla f(x^0)$ and $\nabla g_t(x^0)$, $t \in T(x^0)$, and hence the subdifferential $\partial\eta(0)$ is given by the convex hull of these vectors. (Note that since $T(x^0)$ is compact, this convex hull is also compact and hence is closed.) Together with the optimality condition $0 \in \partial\eta(0)$ this implies that there exist points $t_i \in T(x^0)$, $i = 1, \ldots, n$, and (generalized Lagrange) multipliers $\lambda_i \geq 0$, $i = 0, 1, \ldots, n$, not all of them zero, such that

$$\lambda_0 \nabla f(x^0) + \sum_{i=1}^{n} \lambda_i \nabla g_{t_i}(x^0) = 0. \tag{4.4}$$

Moreover, under a constraint qualification, the multiplier λ_0 is not zero and can be taken as $\lambda_0 = 1$. Note that Robinson's constraint qualification (3.6) is equivalent here (e.g., [84]) to existence of a vector w such that $w \cdot \nabla g_t(x^0) < 0$ for all $t \in T(x^0)$, which in turn can be considered as an extended Mangasarian–Fromovitz constraint qualification.

It is not true in general that if the max-function $h(x)$ attains its (local) minimum at x_0, then x_0 is a (locally) optimal solution of the corresponding semi-infinite program. Think, for example, about a case where $g_t(x) = 0$ for all $x \in \mathbb{R}^n$ and all $t \in T$. However, it is not difficult to see that if x_0 is a strict local minimizer of $h(x)$, i.e., $h(x) > 0$ for all $x \neq x_0$ in a neighborhood of x_0, then x_0 is a strict local minimizer of the corresponding semi-infinite program. Since the space $X = \mathbb{R}^n$ is finite-dimensional, it follows from Hadamard directional differentiability of $h(x)$ that a sufficient condition for x_0 to be a strict local minimizer of $h(x)$ is that $h'(x_0, d) > 0$ for all $d \neq 0$. This, in turn, is equivalent to the condition $0 \in \text{int}\{\partial\eta(0)\}$. Therefore, we obtain the following first-order sufficient conditions for strict local optimality of x_0 (e.g., [41, Thm. 3.1.16]),

$$0 \in \text{int}\left\{\text{conv}\left(\nabla f(x^0),\ \cup_{t \in T(x^0)} \nabla g_t(x^0)\right)\right\}. \tag{4.5}$$

It is possible to conduct a general perturbation analysis using this idea of reduction to composite unconstrained optimization [42, 43, 44].

4.2. Optimal value functions of convex problems. Suppose now that X and U are Banach spaces and that $\Phi(u)$ is defined by cone constraints in the form (1.2) with K being a closed convex cone in the Banach space Y. We say that a mapping $g : X \to Y$ is convex with respect to a convex closed cone $C \subset Y$ if the inequality

$$g(tx + (1-t)y) \preceq_C tg(x) + (1-t)g(y) \tag{4.6}$$

holds for any $x, y \in X$ and any $t \in [0, 1]$. For example, if $Y = \mathbb{R}^n$, then the mapping $g(x)$ can be written componentwise $g(x) = (g_1(x), \ldots, g_n(x))$ and it is convex with respect to the cone $C = \mathbb{R}^n_+$ iff $g_i(x)$, $i = 1, \ldots, n$, are convex (real-valued) functions. Let us observe that (4.6) means that

$$tg(x) + (1-t)g(y) - g(tx + (1-t)y) \in C,$$

which, by duality, is equivalent to the condition that

$$\langle \lambda, tg(x) + (1-t)g(y) - g(tx + (1-t)y) \rangle \leq 0$$

for any $\lambda \in C^-$. This in turn is equivalent to convexity of $\langle -\lambda, g(\cdot) \rangle$ for any $\lambda \in C^-$.

Consider now the unperturbed program

$$\operatorname*{Min}_{x \in X} f(x) \quad \text{subject to } G(x) \in K, \qquad (P_0) \tag{4.7}$$

with $f(x) = f(x, u^0)$ and $G(x) = G(x, u^0)$. Suppose that the unperturbed program (P_0) is *convex*. That is, the function $f(x)$ is convex and the mapping $G(x)$ is convex with respect to the cone $C := -K$. For example, if $K := \mathbb{R}^p_-$, and hence $C = \mathbb{R}^p_+$, program (P_0) becomes a standard nonlinear convex programming problem. Consider the Lagrangian $L(x, \lambda) = f(x) + \langle \lambda, G(x) \rangle$ of the program (P_0). We have then that $\operatorname{Max}_{\lambda \in K^-} L(x, \lambda)$ is $f(x)$ if $G(x) \in K$, and is $+\infty$ otherwise. Therefore, program (P_0) can be written as the min-max problem

$$\operatorname*{Min}_{x \in X} \operatorname*{Max}_{\lambda \in K^-} L(x, \lambda). \tag{4.8}$$

The (Lagrangian) dual of (P_0) is obtained by reversing the order in which min and max are applied; that is,

$$\operatorname{Max} \left\{ \psi(\lambda) := \inf_{x \in X} L(x, \lambda) \right\} \quad \text{subject to } \lambda \in K^-. \qquad (D_0) \tag{4.9}$$

Note that it is always true that $\operatorname{val}(P_0) \geq \operatorname{val}(D_0)$. It is said that there is no duality gap between (P_0) and (D_0) if $\operatorname{val}(P_0) = \operatorname{val}(D_0)$.

Suppose now that $f(x, u) = f(x)$, i.e., the objective function is independent of u, that $U = Y$, and that the mapping $G(x, u)$ is of the form $G(x, u) = G(x) + u$. In that case the program (P_u) becomes

$$\operatorname*{Min}_{x \in X} f(x) \quad \text{subject to } G(x) + u \in K. \tag{4.10}$$

Properties of the optimal value function $v(u)$ of the program (4.10) were thoroughly investigated in Rockafellar [77, 78]. Since (P_0) is convex, the optimal value function $v(u)$ is also convex. A constraint qualification, ensuring "no duality gap" between the dual problems (P_0) and (D_0), is continuity of $v(u)$ at $u = 0$. It is clear that if $v(0)$ is finite, then a necessary condition for continuity of $v(u)$ at $u = 0$ is that the feasible set of the program (4.10) is nonempty for all u in a neighborhood of $0 \in Y$, which can be written in the form

$$0 \in \operatorname{int}\{G(X) - K\}. \tag{4.11}$$

It is possible to show that the converse of that is also true [73, Corollary 1]. We have then the following result (Rockafellar [78, p. 41]). Let us denote by Λ_0 the set of optimal solutions of the dual program (D_0).

THEOREM 4.2. *Suppose that* (i) *the program* (P_0) *is convex,* (ii) *its optimal value* $v(0)$ *is finite, and* (iii) *the constraint qualification* (4.11) *holds. Then* (a) *there is no duality gap between* (P_0) *and its dual* (D_0)*,* (b) *the set* Λ_0 *of optimal solutions of the dual problem is nonempty and bounded,* (c) $v(u)$ *is continuous at* $u = 0$ *and* $\partial v(0) = \Lambda_0$*,* (d) $v(u)$ *is Hadamard directionally differentiable at* $u = 0$ *and*

$$v'(0, d) = \operatorname*{Max}_{\lambda \in \Lambda_0} \langle \lambda, d \rangle. \tag{4.12}$$

We note that if $G(x)$ is continuously differentiable, then (4.11) is equivalent to Robinson's constraint qualification (3.6). Whenever K is a convex closed cone with a nonempty interior, (4.11) is equivalent to the *Slater* condition, that is, existence of a point $\bar{x} \in X$ such that $G(\bar{x}) \in \operatorname{int}(K)$.

It is important to understand a relation between the set Λ_0 of optimal solutions of (D_0) and the set of Lagrange multipliers of (P_0). Let $x^0 \in X$ and $\lambda^0 \in Y^*$ be such that

$$x^0 \in \arg\min_{x \in X} L(x, \lambda^0), \quad G(x^0) \in K, \quad \langle \lambda^0, G(x^0) \rangle = 0, \quad \lambda^0 \in K^-. \tag{4.13}$$

Then (x^0, λ^0) is a saddle point of the min-max problem (4.8) and hence $\mathrm{val}(P_0) = \mathrm{val}(D_0)$, and x^0 and λ^0 solve (P_0) and (D_0), respectively. Moreover, in that case $\lambda \in \Lambda_0$ iff it satisfies the first, third, and fourth conditions of (4.13). Note also that if $f(x)$ and $G(x)$ are continuously differentiable then, by convexity, the first condition in (4.13) is equivalent to $D_x L(x^0, \lambda^0) = 0$. We obtain an important result that, in the convex case, if the set of Lagrange multipliers is nonempty for an optimal solution of (P_0) then it is the same for any other optimal solution of (P_0) and coincides with the set of optimal solutions of the dual problem (D_0) [38, 77].

EXAMPLE 4.2. *Consider the following optimization problem:*

$$\operatorname*{Min}_{x \in \mathbb{R}^2} x_1 \ \text{ subject to } \ -x_1 + e^{x_2} + u \le 0, \tag{4.14}$$

parameterized by $u \in \mathbb{R}$. This is a convex problem and its dual, for $u = 0$, has only one solution $\lambda = 1$. The Slater condition holds and there is no duality gap between the primal and dual problems here. It can be easily seen that $v(u) = u$ and hence formula (4.12) *holds, although the primal problem* (4.14) *does not possess an optimal solution.*

4.3. A general first-order estimate. Let us come back now to the general problem under abstract constraints. Suppose that $f(x, u)$ and $G(x, u)$ are continuously differentiable and let $x^0 \in S(u^0)$. For a given direction $d \in U$ consider the following linearization of (P_u):

$$(PL_d) \qquad \operatorname*{Min}_{h \in X} Df(x^0, u^0)(h, d) \ \text{ subject to } \ DG(x^0, u^0)(h, d) \in T_K(G(x^0, u^0)).$$

Since $[T_K(G(x^0, u^0))]^- = N_K(G(x^0, u^0))$ it is not difficult to see that the (Lagrangian) dual program of (PL_d) can be written in the form

$$(DL_d) \qquad \operatorname*{Max}_{\lambda \in \Lambda(x^0, u^0)} D_u L(x^0, \lambda, u^0) d.$$

It is possible to show by convex analysis that under the directional constraint qualification (3.14) there is no duality gap between (PL_d) and (DL_d) [20, Part I, Prop. 3.1]. The following upper bound (4.15) is then a direct consequence of Proposition 3.4 (cf. [20, Part I], [53]).

PROPOSITION 4.3. *Suppose that the directional constraint qualification* (3.14) *holds in the direction d. Then* $\mathrm{val}(PL_d) = \mathrm{val}(DL_d)$ *and*

$$v'_+(u^0, d) \le \mathrm{val}(PL_d). \tag{4.15}$$

Moreover, $\mathrm{val}(PL_d) > -\infty$ *if and only if $\Lambda(x^0, u^0) \ne \emptyset$, in which case the set of optimal solutions of (DL_d) is a nonempty weak* compact subset of $\Lambda(x^0, u^0)$.*

A natural question arises of whether the upper bound given in (4.15) is tight in the sense that $v(u)$ is directionally differentiable at u^0 in the direction d and the directional derivative is equal to $\mathrm{val}(DL_d)$. As the following example (due to Gauvin and Tolle [36, pp. 308–309]) shows, however, this may not be the case and the above question would have a negative answer.

EXAMPLE 4.3. *Let $X = \mathbb{R}^2$, $U = \mathbb{R}^2$ and consider the problem of minimization of $f(x) = -x_2$ subject to $x_2 + x_1^2 \le u_1$ and $x_2 - x_1^2 \le u_2$. For $u^0 = (0,0)$ this program has the unique optimal solution $x^0 = (0,0)$. The Mangasarian–Fromovitz constraint qualification holds here and the (bounded) Lagrange multipliers set is given by*

$$\Lambda(x^0, u^0) = \{(\lambda_1, \lambda_2) : \lambda_1 + \lambda_2 = 1,\ \lambda_1 \ge 0,\ \lambda_2 \ge 0\}.$$

Consider the direction $d = (1,0)$. Then for $u = td$, $t \ge 0$, this program has two optimal solutions $((t/2)^{1/2}, t/2)$ and $(-(t/2)^{1/2}, t/2)$ and the optimal value function $v(td) = -t/2$, and hence $v'(u^0, d) = -1/2$. On the other hand,

$$\operatorname{val}(PL_d) = \operatorname{val}(DL_d) = \max_{\lambda \in \Lambda(x^0,u^0)} D_u L(x^0, \lambda, u^0)d = 0.$$

As we shall see later, lower bounds for the directional derivatives of $v(u)$ are based on a second-order analysis. It is still possible, however, to obtain directional derivatives of $v(u)$ by a first-order analysis in the convex case. Suppose now that for $u = u^0$ the corresponding program (P_0) is convex. That is, the function $f(\cdot, u^0)$ is convex and the mapping $G(\cdot, u^0)$ is convex with respect to the cone $C = -K$. Recall that this implies that the Lagrangian $L(\cdot, \lambda, u^0)$ is convex for any $\lambda \in K^-$. As we have seen, in the convex case the set $\Lambda(x^0, u^0)$ is the set of optimal solutions of the dual program of (P_0) and hence is independent of a particular $x^0 \in S(u^0)$. Therefore, we can omit x^0 and denote this set by $\Lambda(u^0)$. We now give an extension of Gol'shtein's theorem [38] to the Banach space setting. This result is easily proved by combining the techniques of [20, Prop. 3.2, Part I] and [85].

THEOREM 4.4. *Suppose that* (i) *the problem (P_0) is convex,* (ii) *the optimal set $S(u^0)$ is nonempty and compact,* (iii) *the directional regularity condition holds for all $x^0 \in S(u^0)$, and* (iv) *for sufficiently small $t > 0$ the program (P_{u^0+td}) possesses an $o(t)$-optimal solution $x(t)$ such that $\operatorname{dist}(x(t), S(u^0)) \to 0$ as $t \to 0^+$. Then the optimal value function is directionally differentiable at u^0 in the direction d and*

$$v'(u^0, d) = \inf_{x \in S(u^0)} \sup_{\lambda \in \Lambda(u^0)} D_u L(x, \lambda, u^0)d. \tag{4.16}$$

Recall that if the Slater condition, for the program (P_0), is satisfied, then the constraint qualification (3.6) (and hence the directional regularity condition (3.14)) holds for all $x^0 \in S(u^0)$, and the set $\Lambda(u^0)$ of Lagrange multipliers is nonempty and bounded. In that case $v(u)$ is directionally differentiable at u^0 (in all directions) in the *Hadamard* sense.

Consider, for example, a parameterized semi-infinite program with $X = \mathbb{R}^n$ and the feasible set $\Phi(u)$ given in the form (1.4). In this case the problem (P_0) is convex if the functions $f(\cdot, u^0)$ and $g_t(\cdot, u^0)$, $t \in T$, are convex. The Slater condition is equivalent here to existence of a point $\bar{x} \in \mathbb{R}^n$ such that $g_t(\bar{x}, u^0) < 0$ for all $t \in T$ and assumptions (ii) and (iv) of Theorem 4.4 can be easily ensured by compactness arguments (cf. [85]).

Consider the linearization (PL_d) of the parameterized (not necessarily convex) problem (P_u). Suppose that the optimal value of (PL_d) is finite and let h be an ε-optimal solution of (PL_d). (Note that it can happen that the linearized program (PL_d) does not possess an optimal solution even if its optimal value is finite.) Then it follows from Proposition 3.4 that there exists a path $x(t) = x^0 + th + o(t) \in \Phi(u^0 + td)$ and hence

$$v(u^0 + td) - v(u^0) \le f(x(t), u^0 + td) - f(x^0, u^0) \le t\,\mathrm{val}(PL_d) + t\varepsilon + o(t).$$

Therefore, equality $v'(u^0, d) = \mathrm{val}(DL_d)$ will imply existence, for any $\varepsilon > 0$, of a $(t\varepsilon)$-optimal solution $x(t)$ of (P_{u^0+td}) such that $\|x(t) - x^0\| = O(t)$. It is possible to show that in a sense the converse of that is also true; i.e., existence of Lipschitz stable at x^0, nearly optimal, solutions of (P_{u^0+td}) implies $v'(u^0, d) = \mathrm{val}(DL_d)$ [53, 54, 55].

Let us finish this section by giving the following (first-order) result which holds for general (not necessarily convex) programs. In a finite-dimensional setting it was given in [36, 34] and extended to Banach spaces in [53].

THEOREM 4.5. *Suppose that* (i) $S(u^0) \neq \emptyset$, (ii) *Robinson's constraint qualification holds at every* $x \in S(u^0)$, *and* (iii) *for every* $\varepsilon > 0$ *there exists a* $t\varepsilon$*-optimal solution of* (P_{u^0+td}) *converging to a point of the optimal set* $S(u^0)$. *Then for every* $d \in U$,

$$\begin{aligned}\inf_{x\in S(u^0)} \inf_{\lambda\in\Lambda(x,u^0)} \quad & D_u L(x, \lambda, u^0)d \le v'_-(u^0, d) \\ & \le v'_+(u^0, d) \le \inf_{x\in S(u^0)} \sup_{\lambda\in\Lambda(x,u^0)} D_u L(x, \lambda, u^0)d.\end{aligned} \tag{4.17}$$

In particular, if in addition $\Lambda(x, u^0) = \{\bar\lambda(x)\}$ *is a singleton for every* $x \in S(u^0)$, *then* $v(u)$ *is directionally differentiable at* u^0 *and*

$$v'(u^0, d) = \inf_{x\in S(u^0)} D_u L(x, \bar\lambda(x), u^0)d. \tag{4.18}$$

Note that if the optimal set $S(u^0)$ is compact, then assumption (iii) in the above theorem can be replaced by the assumption of existence of an $o(t)$-optimal solution $x(t)$ of (P_{u^0+td}) such that $\mathrm{dist}(x(t), S(u^0))$ tends to zero as $t \to 0^+$. A discussion of conditions ensuring uniqueness of Lagrange multipliers under cone constraints is given in [86].

5. Strong stability.

5.1. Strong stability for generalized equations. Consider the parameterized problem (P_u) with the corresponding feasible set defined by abstract constraints in the form (1.2). We assume in this section that X, Y, and U are Banach spaces, that K is a closed convex subset of Y, and that $f(x, u)$ and $G(x, u)$ are twice continuously differentiable (jointly in x and u). We may assume, without loss of generality, that the unperturbed problem is formulated for $u = 0$. The perturbation analysis of this section is based on the expansion of the associated optimality system representing the corresponding first-order necessary conditions

$$D_x L(x, \lambda, u) = 0, \quad \lambda \in N_K(G(x, u)). \tag{5.1}$$

Note that the condition $G(x, u) \in K$ is implicit in (5.1), as the normal cone to a point outside K is empty. It is advantageous at this point to embed the above optimality system into the following abstract problem:

$$F(z, u) \in N(z), \tag{5.2}$$

where $F : Z \times U \to W$, Z and W are Banach spaces, and N is a closed multifunction $N : Z \to 2^W$. Relation (5.2) is called a *generalized equation*. Using N_K^{-1}, the graph inverse of N_K, we may write (5.1) as

$$\big(D_x L(x, \lambda, u), G(x, u)\big) \in \big(0, N_K^{-1}(\lambda)\big), \tag{5.3}$$

which is a particular case of (5.2) with $z = (x, \lambda) \in X \times Y^*$. Note that if K is a convex closed cone, then $N_K(y)$ is given by $\{\alpha \in K^- : \langle \alpha, y \rangle = 0\}$, if $y \in K$, and is the empty set if $y \notin K$, and that

$$N_K^{-1}(\lambda) = N_{K^-}(\lambda) = \begin{cases} y \in K : \langle \lambda, y \rangle = 0, & \text{if } \lambda \in K^-, \\ \emptyset, & \text{otherwise.} \end{cases} \tag{5.4}$$

Suppose that for $u = 0$ the generalized equation (5.2) has a solution z^0. Since we know how to linearize the mapping F, and not the multifunction N, we may try to approximate the solution $\bar{z} = \bar{z}(u)$ of (5.2) by $z^0 + z^1(u)$, where $z^1(u)$ is a solution of the following generalized equation (with unknown ζ), representing partial linearization of (5.2),

$$F(z^0, 0) + DF(z^0, 0)(\zeta, u) \in N(z^0 + \zeta). \tag{5.5}$$

The "first-order term" $z^1(u)$, if it exists, is in general a nonlinear function of u. This is a price to pay for the full generality of the framework (5.2). Therefore, we call $z^0 + z^1(u)$ a *pseudoexpansion* of the solution $\bar{z}(u)$.

DEFINITION 5.1. *Let, for $u = 0$, z^0 be a solution of the generalized equation* (5.2). *We say that the strong stability condition holds (that z^0 is a strongly stable solution) if there exist $\varepsilon > 0$ and $M > 0$ such that for all $w \in B_W(0, \varepsilon)$ the linearized generalized equation (with unknown ζ)*

$$F(z^0, 0) + D_z F(z^0, 0)\zeta + w \in N(z^0 + \zeta) \tag{5.6}$$

has in $B_Z(0, M)$ a unique solution $\hat{z}(w)$, and $\hat{z}(\cdot) : B_W(0, \varepsilon) \to B_Z(0, M)$ is Lipschitz continuous.

This condition essentially states that the linearized generalized equation is locally well posed for arbitrary but small perturbations. The idea of conducting perturbation analysis using the abstract generalized equations framework, based on the above strong stability condition, is due to Robinson [74], who proved the following result. (The result stated there is for the case when the multifunction $N(\cdot)$ is a certain normal cone, but the proof goes without modification for any closed multifunction.)

THEOREM 5.1. *Suppose that, for $u = 0$, z^0 is a solution of the generalized equation* (5.2) *and the strong stability condition holds. Then, for all u in a neighborhood of* 0, *the mappings $\bar{z}(u)$ and $z^1(u)$ are well defined in the vicinity of z^0 and in $B_Z(0, M)$, respectively. In addition, $\bar{z}(u)$ is Lispchitz continuous, $z^1(u) = O(\|u\|)$, and $\bar{z}(u) = z^0 + z^1(u) + o(\|u\|)$.*

The above theorem gives effective means to check well-posedness and Lipschitz stability of $\bar{z}(u)$. It does not give, however, any formula for the directional derivative of $\bar{z}(u)$, as it may happen that $t^{-1}z^1(tu)$ has no limit when $t \downarrow 0$.

If the data have enough regularity, then we may write a higher-order pseudoexpansion along the same lines [19]. Suppose, for instance, that the mapping F is twice continuously differentiable, and z^0 is a strongly stable solution of (5.2). Then $\bar{z}(u) = z^0 + z^1(u) + \frac{1}{2}z^2(u) + o(\|u\|^2)$, where $z^1(u)$ is as above and $z^2(u)$ is a solution of the following linear generalized equation (with unknown ζ) obtained by using a second-order expansion of $F(z(u), u)$:

$$F(z^0, 0) + DF(z^0, 0)(z^1(u), u) + \tfrac{1}{2}\left[D_z F(z^0, 0)\zeta + D^2 F(z^0, 0)(z^1(u), u)(z^1(u), u)\right] \\ \in N(z^0 + z^1(u) + \tfrac{1}{2}\zeta).$$

5.2. Strongly stable solutions of the optimality system. When applying Theorem 5.1 to the optimality system (5.1), two questions arise. The first is to establish a relation between the locally optimal solutions of (P_u) and solutions of the optimality system, under the strong stability condition. The second is to give necessary and/or sufficient conditions for strong regularity.

The result below deals with the first question. The only delicate part of the proof consists in checking that the perturbed optimization problem has a solution. The argument uses Ekeland's variational principle [7]; it deals with the scalar perturbation case, but the argument is valid in the case when U is a Banach space.

PROPOSITION 5.2 (see [19, Thm 4.6]). *Let (x^0, λ^0) be a local solution of (P_0) and an associated Lagrange multiplier. If (x^0, λ^0) is a strongly stable solution of* (5.1), *then there exists a neighborhood N of x^0 such that, for all u sufficiently close to $0 \in U$, the perturbed problem (P_u) has a unique local minimizer $\bar{x}(u) \in N$ and an associated unique Lagrange multiplier $\bar{\lambda}(u)$. It follows that $(\bar{x}(u), \bar{\lambda}(u))$ is the locally unique solution of* (5.1).

We now discuss a relation between strong stability, constraint qualifications, and second-order conditions. A necessary condition for strong stability is that the regularity condition (3.6) holds [19, Lemma 4.3]. Below is given a sufficient condition for strong regularity [19, Thm. 4.9], which reduces in a finite-dimensional setting to Theorem 4.1 in [74].

PROPOSITION 5.3. *Sufficient conditions for strong stability of (x^0, λ^0) for problem (P_0) are:* (i) *$D_x G(x^0, 0) : X \to Y$ is onto, and* (ii) *there exists $\alpha > 0$ such that*

$$D^2_{xx} L(x^0, \lambda^0, 0)(h, h) \geq \alpha \|h\|^2 \quad \forall h \in X \text{ satisfying } D_x G(x^0, 0)h \in \mathrm{Sp}(K). \tag{5.7}$$

The expansion of the solution and multiplier has a nice interpretation. Expanding $\bar{x}(u)$ and $\bar{\lambda}(u)$ as $\bar{x}(u) = x^0 + x^1(u) + o(\|u\|)$ and $\bar{\lambda}(u) = \lambda^0 + \lambda^1(u) + o(\|u\|)$, respectively, and specializing (5.5) to the linearized optimality system of (P_u), we obtain that $x^1(u)$ and $\lambda^1(u)$ are solutions of the generalized equation with unknown $(\chi, \lambda) \in X \times Y^*$:

$$\begin{gathered} D_x f(x^0, 0) + D_{(x,u)}[D_x L(x^0, \lambda^0, 0)](\chi, u) + [D_x G(x^0, 0)]^*(\lambda^0 + \lambda) = 0, \\ G(x^0, 0) + DG(x^0, 0)(\chi, u) \in N_K^{-1}(\lambda^0 + \lambda). \end{gathered}$$

Observe that the above (generalized) equations represent the first-order optimality system for the problem

$$\begin{array}{ll} \mathrm{Min}_{h \in X} & Df(x^0, 0)(h, u) + \frac{1}{2} D^2_{(x,u)^2} L(x^0, \lambda^0, 0)((h, u), (h, u)), \\ \text{subject to} & G(x^0, 0) + DG(x^0, 0)(h, u) \in K. \end{array} \tag{Q_u}$$

It can be easily checked that conditions for strong stability of (P_u) and (Q_u) are identical. Consequently, if (x^0, λ^0) is a strongly stable solution of (P_0), then (locally) the optimality system of (Q_u) has the unique solution $(x^1(u), \lambda^0 + \lambda^1(u))$, and $x^1(u)$ is a local minimizer of (Q_u). A practical consequence of this result is that, whenever u is sufficiently close to 0, we may compute $x^1(u)$ by applying any locally convergent algorithm for solving (Q_u), starting from the point x^0 (or (x^0, λ^0) for a primal-dual algorithm).

5.3. Polyhedral case. In the case of scalar perturbations, i.e., $u(t) = td$, with $d \in U$ and $t \geq 0$, and if $\Phi(u)$ is of the form (1.3), then the pseudoexpansion appears to be a Taylor expansion. That is, the coefficients of the pseudoexpansion of

$(\bar{x}(td), \bar{\lambda}(td))$ of order k are, when t is small enough, of the form $t^k(\hat{x}^k, \hat{\lambda}^k)$ for some $(\hat{x}^k, \hat{\lambda}^k) \in X \times Y^*$. Let us verify this for the first-order term. Recall that $I(x, u)$ denotes the set of active-at-(x, u) inequality constraints. For u small enough, only the inequality constraints in $I(x^0, 0)$ may be active for local solutions of (Q_u). Therefore, by setting $\hat{h} := h/t$, we can write an equivalent of (Q_u) in the form

$$\begin{aligned} \operatorname{Min}_{\hat{h}\in X} \quad & t^{-1}Df(x^0,0)(\hat{h},d) + \tfrac{1}{2}D^2L(x^0,\lambda^0,0)((\hat{h},d),(\hat{h},d)), \\ \text{subject to} \quad & Dg_i(x^0,0)(\hat{h},d) = 0, \quad i = 1,\dots,q, \\ & Dg_i(x^0,0)(\hat{h},d) \le 0, \quad i \in I(x^0,0). \end{aligned} \tag{$\hat{Q}$}$$

Writing the associated optimality system, with multiplier $\lambda = \lambda^0 + t\hat{\lambda}$, and using $D_xL(x^0, \lambda^0, 0) = 0$, we obtain an equivalent system that does not depend on t. Therefore the unique associated solution $x^1(td)/t$ is also independent of t. That is, under the strong stability condition, $\bar{x}(td)$ is directionally differentiable at $t = 0$ and the corresponding directional derivatives $\bar{x}'(0, d)$ and $\bar{\lambda}'(0, d)$ are given as the optimal solution of the program $(\hat{Q})$ and its associated Lagrange multiplier (cf. [76, 79]).

If in addition the data of (P_u) are analytic (i.e., locally the involved function can be expanded in power series), then for $t > 0$ small enough $\bar{x}(td)$ and $\bar{\lambda}(td)$ are themselves analytic functions [19, Theorem 4.13].

In the case of finitely many inequality constraints another interesting result is that we know several characterizations for strong stability [31]. One of them reduces to the sufficient condition of Proposition 5.3, if we observe that (a) only active constraints at x^0 should be taken into account and (b) those inequality constraints associated with a positive multiplier will remain active for the perturbed problem, and we change neither the solution nor the multiplier by setting these as equalities. Define

$$J := \{1, \dots, q\} \cup I(x, u), \quad J^0 := \{1, \dots, q\} \cup \{i = q+1, \dots, p\,;\ \lambda_i^0 > 0\}.$$

By G_J we denote the restriction of G to components in J. We have the following proposition.

PROPOSITION 5.4. *Let $\Phi(u)$ be given by a finite number of constraints in the form* (1.3). *Let (x^0, λ^0) be a local solution of (P_0) and an associated Lagrange multiplier. Then (x^0, λ^0) is a strongly stable solution iff $D_xG_J(x^0, 0)$ is onto, and there exists $\alpha > 0$ such that*

$$D^2_{xx}L(x^0,\lambda^0)(h,h) \ge \alpha\|h\|^2 \ \ \forall h \in X,\ D_xg_i(x^0,0)h = 0,\ i \in J^0. \tag{5.8}$$

The above proposition is a consequence of [47, 51, 74]; see [19] for a simple proof and [31, 49, 52] for related results. The strong regularity approach is applied to the sensitivity analysis of optimal control problems in [58, 59]. The stability analysis of generalized equations is discussed in a more abstract setting in [66].

6. Quantitative stability of optimal solutions. We assume in this section that X and U are Banach spaces and discuss continuity properties of optimal solutions of the problem (P_u) from a quantitative point of view. In particular, we will be interested in Hölder and Lipschitzian behavior of the optimal-solution multifunction $S(\cdot) : U \to 2^X$. We say that $S(u)$ is *Hölder stable* of degree α ($\alpha > 0$), at a point $u^0 \in U$, if there exists a constant $\kappa > 0$ such that

$$S(u) \subset S(u^0) + \kappa\|u - u^0\|^\alpha B_U \tag{6.1}$$

for all u in a neighborhood of u^0. If $S(u)$ is Hölder stable of degree $\alpha = 1$, i.e., $S(u)$ is upper Lipschitzian at u^0, we say that $S(u)$ is *Lipschitz* stable. We also discuss Hölder and Lipschitz stability of ε-optimal solutions of (P_u).

EXAMPLE 6.1. *Consider $f(x,u) = x^{2m}/(2m) - ux$, $u \in \mathbb{R}$, m is a positive integer, and let $\Phi(u) = [-1,1]$ for all $u \in \mathbb{R}$. Then $S(u) = \{u^{1/(2m-1)}\}$ is a singleton for any $u \in [0,1)$. Therefore, $S(u)$ is not Lipschitz stable at $u = 0$ if $m > 1$.*

The above example demonstrates that in order to establish Lipschitzian behavior of optimal solutions we need some additional assumptions beside continuity and first-order differentiability properties of the involved functions. It turns out that quantitative stability of $S(u)$ is closely related to a second-order analysis of (P_u). Let us start by introducing the following second-order abstract condition. Consider the unperturbed program (P_0), and suppose that its optimal solution set $S_0 = S(u^0)$ is not empty. Denote $\Phi_0 = \Phi(u^0)$ and $f_0 = v(u^0) = \inf_{x\in\Phi_0} f(x)$, where $f(x) = f(x,u^0)$.

DEFINITION 6.1. *We say that the growth condition of order $\alpha > 0$ (for the program (P_0), in a neighborhood N of the optimal set S_0) holds, if there exists a constant $c > 0$ such that*

$$f(x) \geq f_0 + c\,[\mathrm{dist}(x, S_0)]^\alpha \qquad \forall x \in \Phi_0 \cap N. \tag{6.2}$$

When $\alpha = 1$ (resp., 2) we speak of the first-order (resp., second-order) growth condition.

If the set S_0 is finite, then the second-order growth condition is strongly connected to certain second-order sufficient conditions discussed below. In case S_0 has nonisolated connected parts, verification of (6.2) may be tricky. If the space X is finite-dimensional and Φ_0 is defined by a finite number of constraints, then we know a characterization of the second-order growth condition in terms of the Hessian of Lagrangian and qualification conditions for convex problems [18], whereas for nonconvex problems, we know only some necessary or sufficient conditions [17].

Under (6.2), it is relatively easy to give quite general sufficient conditions for Lipschitzian stability of $S(u)$ when the feasible set $\Phi(u) \equiv \Phi_0$ is unperturbed, i.e., independent of u.

PROPOSITION 6.1 (see [83]). *Suppose that $\Phi(u) \equiv \Phi_0$, that the second-order growth condition* (6.2) *holds in a neighborhood N of the optimal set S_0, and that the difference function $f(\cdot,u) - f(\cdot,u^0)$ is Lipschitz continuous on N modulus $\ell(u)$, and let $\bar{x} \in N$ be an ε-optimal solution of (P_u). Then*

$$\mathrm{dist}(\bar{x}, S(u^0)) \leq c^{-1}\ell(u) + c^{-1/2}\varepsilon^{1/2}. \tag{6.3}$$

In particular, if, in addition, $f(x,u)$ is continuously differentiable in x with $D_x f(x,\cdot)$ being Lipschitz continuous on N, with the corresponding Lipschitz constant independent of $x \in N$, then $S(u)$ is Lipschitz stable at u^0.

6.1. First- and second-order growth conditions. We now discuss some necessary and/or sufficient conditions for quadratic growth, assuming that the feasible set is defined by constraints of the form (1.2) and that $S(u^0) = \{x^0\}$ is a singleton. A central object is the *critical cone*, defined as follows:

$$C_0 = \{h \in X : D_x G(x^0)h \in T_K(G(x^0)),\ D_x f(x^0)h \leq 0\}. \tag{6.4}$$

The critical cone C_0 represents those directions for which the linearization of cost function and constraints does not provide information about optimality of x^0. We may interpret the "feasibility relation" $h \in C_0$ as the optimization problem $\mathrm{Min}_{h\in X}\, 0$

subject to $h \in C_0$. Computing the dual of this problem, we find the problem of maximizing 0 over the set of generalized Lagrange multipliers. In that sense, the critical cone is dual to the set of generalized Lagrange multipliers. In particular, if h is a critical direction and (λ_0, λ) is a generalized Lagrange multiplier, then (this is also a consequence of (3.4)) we have $\lambda_0 Df(x^0)h = 0$ and $\langle \lambda, DG(x^0)h \rangle = 0$. Therefore, if the set of Lagrange multipliers is nonempty, then every critical direction h satisfies $Df(x^0)h = 0$.

The relation between the critical cone and the set of Lagrange multipliers is closely related to the following linearization of the problem (P_0):

$$\operatorname*{Min}_{h \in X} Df(x^0)h \quad \text{subject to} \quad DG(x^0)h \in T_K(G(x^0)). \tag{6.5}$$

As for any problem of minimization of a linear function over a cone, the optimal value of this problem is either 0 or $-\infty$. The dual (in the sense of convex analysis) of (6.5) consists in maximizing 0 over the (possibly empty) set of Lagrange multipliers. Therefore, the optimal value of the dual problem is also either 0 or $-\infty$ and is equal to 0 iff the set of Lagrange multipliers is nonempty. Consequently, there is no duality gap between these problems if there exist Lagrange multipliers.

The case when $C_0 = \{0\}$ is a very specific situation, strongly related to the first-order growth condition as the following statement shows.

PROPOSITION 6.2. *Suppose that $S(u^0) = \{x^0\}$. Then* (i) *if Robinson's constraint qualification and the first-order growth condition hold at x^0, then $C_0 = \{0\}$,* (ii) *if the feasible set is finitely constrained and $C_0 = \{0\}$, then the first-order growth condition holds at x^0.*

The first-order growth condition for nonisolated minima is discussed in [25, 93].

For a constant $\eta > 0$, consider the cone

$$C_\eta = \left\{ h \in X : \operatorname{dist}\left(DG(x^0)h, T_K(G(x^0))\right) \le \eta \|h\|, \ Df(x^0)h \le \eta \|h\| \right\}, \tag{6.6}$$

which we call the *approximate critical cone*. The notation is consistent in the sense that if $\eta = 0$, then C_η reduces to C_0. It can be easily verified that if $\{x^k\}$ is a sequence of feasible points converging to x^0 such that $f(x^k) \le f(x^0) + o(\|x^k - x^0\|)$, then for any $\eta > 0$ the associated normalized directions $h^k := (x^k - x^0)/\|x^k - x^0\|$ belong to C_η for k large enough. This explains the relevance of the approximate critical cones in second-order analysis.

DEFINITION 6.2. *We say that second-order sufficient conditions hold at x^0 if there exist $\eta > 0$, $\beta > 0$, and a bounded set $\Omega \subset \Lambda^g(x^0)$ (of generalized Lagrange multipliers) such that*

$$\sup_{(\lambda_0, \lambda) \in \Omega} D_{xx} L^g(x^0, \lambda_0, \lambda, u^0)(h, h) \ge \beta \|h\|^2 \quad \forall h \in C_\eta. \tag{6.7}$$

THEOREM 6.3 (see [64], [20, Prop. 6.1, Part I and Prop. 3.6, Part II]). *The second-order sufficient conditions* (6.7) *imply the second-order growth condition for the problem (P_0) at the point x^0. Both conditions are equivalent if the feasible set is finitely constrained and the Mangasarian–Fromovitz constraint qualification holds.*

There are situations where it is possible to state some second-order sufficient optimality conditions that are more specific. Since $\cap_{\eta > 0} C_\eta = C_0$, it is possible to show by compactness arguments that if X is a finite-dimensional space, then the approximate critical cone C_η can be replaced by C_0. The second case is when Robinson's constraint qualification holds. Then it can be shown that every $h \in C_\eta$ is at distance

$O(\eta\|h\|)$ from the smaller cone

$$\hat{C}_\eta = \left\{h \in X : D_x G(x^0)h \in T_K(G(x^0)),\ D_x f(x^0)h \le \eta\|h\|\right\}. \tag{6.8}$$

We may then replace C_η by $\hat{C}_\eta$ and choose $\Omega = \{1\} \times \Lambda(x^0)$, i.e., use the Lagrangian instead of the generalized Lagrangian, since in that case $\Lambda(x^0)$ is bounded.

Note also that $T_K(G(x^0))$ is a subset of the topological closure of the linear space $\mathrm{Sp}(K)$, generated by K, and hence (6.7) is weaker than the second-order conditions (5.7), even when $\Lambda(x^0) = \{\lambda^0\}$ is a singleton.

In the case of an unconstrained optimization problem, the second-order conditions may be motivated as follows. If x^0 is an optimal solution of (P_0), then $D_x f(x^0) = 0$ and $D_{xx} f(x^0)(h,h) \ge 0$ for all $h \in X$. Moreover, if the second-order sufficient condition $D_{xx} f(x^0)(h,h) \ge \beta\|h\|^2$ (for some $\beta > 0$ and all $h \in X$) holds, then $\|h\|_1 := (D_{xx} f(x^0)(h,h))^{1/2}$ defines a norm on X which is equivalent to the original norm of X. Therefore, in that case, X is Hilbertisable; i.e., there exists a norm that is equivalent to the one of X and for which X is a Hilbert space. We end up with a surprising consequence. If X is a non-Hilbertisable Banach space, then the second-order growth condition can never hold at a unique point x_0 for an unconstrained optimization of a smooth function! Similarly, if the space X is non-Hilbertisable, then $D^2_{xx} L(x^0, \lambda)(h,h)$ cannot be positive definite for any Lagrange multiplier λ. (This, of course, does not prevent the second-order growth condition from holding for a constrained problem.) This observation was a starting point for development of second-order optimality conditions for optimal control problems on the basis of the so-called "two-norm discrepancy" approach, which in turn have been used to obtain stability results for optimal solutions of perturbed optimal control problems (e.g., [14, 42, 60, 62, 63, 94]).

When the feasible set is defined by the abstract constraints, the second-order sufficient conditions of Definition 6.2 are rather crude since they do not take into account the curvature of the boundary of K. Indeed, by Theorem 6.3, they imply that quadratic growth holds at x^0 for the problem in which we replace constraints $G(x) \in K$ by constraints $G(x) \in G(x^0) + T_K(G(x^0))$, where the set K is "linearized" at the point $G(x^0)$. Second-order necessary conditions, that take into account the curvature of K, have been suggested in [26] based on an approach in [48]. These conditions are based on verification of (local) optimality of x^0 along parabolic curves of the form $x^0 + th + \frac{1}{2}t^2 w + o(t^2)$ (cf. [11]). There is no reason, a priori, why one should verify (local) optimality of x^0 along such parabolic curves only, and in general one can expect a gap between such necessary and the corresponding sufficient second-order optimality conditions. It turns out, however, that in case the space X is finite-dimensional and if a certain assumption (of so-called second-order regularity of the set K at the point $G(x_0)$) holds, then there is no gap between such necessary and sufficient second-order optimality conditions. The concept of second-order regularity was suggested in [20] and developed in [21]. For example, in the case of semidefinite programming, second-order regularity of the cone of positive semidefinite matrices always holds [21].

6.2. Hölder and Lipschitz estimates. Example 4.3 shows that under the hypotheses of regularity of constraints and some second-order sufficient conditions, one may hope to establish Hölder stability of optimal solutions of degree $\alpha = 1/2$ at most. Such Hölder stability, in various frameworks of generality, was established in a number of publications [1, 2, 3, 16, 20, 35, 83].

PROPOSITION 6.4 (see [83]). *Suppose that* (i) *the second-order growth condition holds in a neighborhood N of x^0,* (ii) *$f(\cdot, u)$ is Lipschitz continuous on N with the Lipschitz constant independent of u for all u in a neighborhood of u^0,* (iii) *the difference function $f(\cdot, u) - f(\cdot, u^0)$ is Lipschitz continuous on N modulus $\ell(u) = O(\|u - u^0\|)$,* (iv) *the multifunction $u \to \Phi(u) \cap N$ is upper Lipschitzian at u^0 and* $\operatorname{dist}(x^0, \Phi(u)) = O(\|u - u^0\|)$. *Then for any $\varepsilon(u)$-optimal solution $\bar{x}(u) \in N$, with $\varepsilon(u) = O(\|u - u^0\|)$, it follows that*

$$\|\bar{x}(u) - x^0\| = O(\|u - u^0\|^{1/2}). \tag{6.9}$$

In case $f(x, u)$ is sufficiently smooth, assumptions (ii) and (iii) in the above proposition can be easily ensured by appropriate bounds on the corresponding derivatives of f. If $\Phi(u)$ is defined by abstract constraints, assumption (iv) holds under Robinson's constraint qualification, as it follows from the Robinson–Ursescu stability theorem (Proposition 3.3).

Now, for a given $d \in U$, consider the set $\Lambda_1(x^0, d)$ of optimal solutions of the problem (DL_d), defined in section 4.3. Let us consider the following strengthened form of second-order sufficient conditions: there exist constants $\beta > 0$ and $\eta > 0$ such that

$$\operatorname*{Max}_{\lambda \in \Lambda_1(x^0, d)} D^2_{xx} L(x^0, \lambda)(h, h) \geq \beta \|h\|^2 \qquad \forall\, h \in C_\eta. \tag{6.10}$$

Since $\Lambda_1(x^0, d) \subset \Lambda(x^0)$, we have that condition (6.10) is stronger then (6.7). Also, condition (6.10) depends on the direction $d \in U$. In a finite-dimensional setting, condition (6.10) was introduced in [80] and then extended to a more general setting in [20, 22, 89, 84].

PROPOSITION 6.5. *Suppose that the feasible set $\Phi(u)$ is defined by a finite number of constraints in the form* (1.3) *and that, for a given $d \in U$, directional regularity condition* (3.14) *is satisfied. Let $\bar{x}(t)$, $t \geq 0$, be an $\varepsilon(t)$-optimal solution of (P_{u^0+td}) with $\varepsilon(t) = O(t^2)$ and such that $\bar{x}(t) \to x^0$ as $t \to 0^+$. Then* (i) *under the second-order sufficient conditions* (6.7), *Hölder stability $\|\bar{x}(t) - x^0\| = O(t^{1/2})$, $t \geq 0$, holds, and* (ii) *under the second-order sufficient conditions* (6.10), *Lipschitz stability $\|\bar{x}(t) - x^0\| = O(t)$, $t \geq 0$, holds.*

As we mentioned earlier, Hölder stability (of degree 1/2) of optimal solutions was derived in a number of publications. Lipschitz stability was obtained, for example, in [8, 16, 80]. For an extension to the case of nonisolated optima, see [18, 23, 50, 81].

It turns out that in the general case, where the set K is not polyhedral, regularity condition (3.6) and second-order sufficient conditions alone do not guarantee Lipschitz stability of optimal solutions.

EXAMPLE 6.2 (see [87]). *Consider the linear space $Y = \mathcal{S}_2$ of 2×2 symmetric matrices and the cone $K \subset \mathcal{S}_2$ of 2×2 positive semidefinite matrices. Consider the linear mapping $G : \mathbb{R}^2 \times \mathcal{S}_2 \to \mathcal{S}_2$ given by $G(x_1, x_2, A) := \operatorname{diag}(x_1, x_2) + A$, where $\operatorname{diag}(x_1, x_2)$ denotes the diagonal matrix with diagonal elements x_1 and x_2, and the parameterized problem*

$$\min_{x \in \mathbb{R}^2} \; x_1 + x_1^2 + x_2^2 \quad \text{subject to } G(x_1, x_2, A) \in K. \qquad (P_A) \tag{6.11}$$

For $A = 0$ the feasible set of the corresponding (unperturbed) problem is $\mathbb{R}^2_+$ and its optimal solution is $x_0 = (0, 0)$. Also, we have that the Slater condition holds here and the unperturbed problem has unique Lagrange multiplier $\lambda_0 = \operatorname{diag}(-1, 0)$. Moreover, *the problem* (6.11) *is convex and the Hessian of the Lagrangian is positive definite on the whole space* $\mathbb{R}^2$*; i.e., the strongest form of second-order sufficient conditions holds. However, it is not difficult to verify here that for any nondiagonal matrix* $A \in \mathcal{S}_2$*, the problem* (P_{tA})*,* $t \geq 0$*, has a unique optimal solution* $\bar{x}(t) = (\bar{x}_1(t), \bar{x}_2(t))$ *with* $\bar{x}_2(t)$ *of order* $t^{2/3}$ *as* $t \to 0^+$*. Therefore, the optimal solution* $\bar{x}(t)$ *is not Lipschitz stable along any nondiagonal matrix direction* A*.*

Various additional conditions, ensuring Lipschitz stability of optimal solutions, were suggested in the literature (see [87] for a survey). The following condition was suggested in [84].

Assumption A. The linearized problem (PL_d) has an optimal solution $\hat{h} = \hat{h}(d)$ such that for $t > 0$,

$$\operatorname{dist}\Big(G(x^0, u^0) + tDG(x^0, u^0)(\hat{h}, d), K\Big) = O(t^2). \tag{6.12}$$

Note that the constraint $DG(x^0, u^0)(h, d) \in T_K(G(x^0, u^0))$ of the linearized problem (PL_d) means that the left-hand side of equation (6.12) is of order $o(t)$. Assumption A is stronger and requires it to be of order $O(t^2)$.

Under the directional regularity condition (3.14) and Assumption A, the following upper bound for the optimal value function holds [20, 84]:

$$v(u^0 + td) \leq v(u^0) + t\,\operatorname{val}(PL_d) + O(t^2). \tag{6.13}$$

Note that the above inequality is stronger than the upper bound (4.15) of Proposition 4.3. Together with the second-order conditions (6.10), the upper bound (6.13) implies (directional) Lipschitz stability of the optimal solutions.

THEOREM 6.6 (see [20, 84, 89]). *Suppose that the directional regularity condition* (3.14)*, Assumption A, and second-order conditions* (6.10) *hold, and let* $\bar{x}(t)$*,* $t \geq 0$*, be an* $\varepsilon(t)$*-optimal solution of* (P_{u^0+td})*, with* $\varepsilon(t) = O(t^2)$*, converging to* x^0 *as* $t \to 0^+$*. Then* $\|\bar{x}(t) - x^0\| = O(t)$*.*

It is possible to show that if the space X is reflexive, then existence of an optimal solution of the linearized problem (PL_d) is a *necessary* condition for directional Lipschitz stability of optimal solutions of (P_u) (see [20, Proposition 3.3], [87]). It is not difficult to see that this condition does not hold in Example 6.2 for any nondiagonal matrix direction A, which explains non-Lipschitzian behavior of optimal solutions in that example.

The above result does not say anything about continuity properties of the associated Lagrange multipliers. In fact, it may happen that under the assumptions of Theorem 6.6 the corresponding Lagrange multipliers are not continuous at u^0 even if $\Lambda(x^0, u^0)$ is a singleton (this may happen, for example, in semi-infinite optimization). Therefore, in order to establish Lipschitz stability of $(\bar{x}(u), \bar{\lambda}(u))$, a regularity condition, stronger than Assumption A, is required. Let K be a closed convex cone. Remember that $N_K^{-1}(\lambda) = \{y \in K : \langle \lambda, y \rangle = 0\}$ when $\lambda \in K^-$.

Assumption B. Let K be a closed convex cone and $\lambda^0 \in \Lambda(x^0, u^0)$. Suppose that the point x^0 is regular with respect to the cone $K_0 := N_K^{-1}(\lambda^0)$, i.e.,

$$0 \in \operatorname{int}\{G(x^0) + DG(x^0)X - K_0\}. \tag{6.14}$$

Since $K_0 \subset K$, Assumption B implies the regularity condition (3.6). Moreover, Assumption B implies uniqueness of the Lagrange multiplier λ^0, and if the cone K is polyhedral, then λ^0 is unique iff Assumption B holds [83, Lemma 4.3]. It is possible to show that Assumption B is stronger than Assumption A. Another property that follows from Assumption B is that the Lagrange multipliers have the same order of perturbation as the optimal solutions [83, Lemma 4.4]. Therefore, under Assumption B, Lipschitz stability of Lagrange multipliers follows from Lipschitz stability of optimal solutions.

7. Second-order analysis of the optimal value function. In this section we investigate directional differentiability of $v(u)$ by employing second-order expansions. We assume that x^0 is the unique solution of (P_{u^0}) and that $d \in U$ satisfies the directional regularity condition. Then we deal with three basic situations. The first is when the set of optimal solutions is Lipschitz stable, which is the case if some strong second-order condition holds. Then the upper bound (4.15) is exact. A second-order expansion allows us to state a second-order upper estimate of the optimal value function that, for nonlinear, semidefinite, and some cases of semi-infinite programming problems, happens to be exact and allows us to compute an expansion of $o(t^2)$-optimal paths.

The second situation is when estimate (4.15) is not tight, as is the case in Example 4.3, but the set of solutions is Hölder stable of degree 1/2. Then an expansion of paths that vary like $O(\sqrt{t})$ allows us to obtain a sharper upper bound of the optimal value function that, for nonlinear programming problems, happens to be exact and allows us to compute an expansion of $o(t)$-optimal paths.

The third situation is when the set of Lagrange multipliers is empty and the set of solutions is Hölder stable of degree 1/2 again, which is as before the consequence of some second-order conditions. Then the solution varies in general as $\sqrt{t}$. An expansion of paths that vary like $O(\sqrt{t})$ allows us to estimate an upper limit of $t^{-1/2}(\text{val}(u^0 + td) - \text{val}(u^0))$. For nonlinear programming problems, this estimate is exact and allows us to compute an expansion of $o(\sqrt{t})$-optimal paths.

7.1. Second-order expansion of the optimal value function and first-order expansion of optimal solutions. Let x^0 be an optimal solution of the problem (P_{u^0}) and d be a direction of perturbations for which a directional regularity condition holds. We have established in section 4 that if $h \in \mathcal{S}(PL_d)$ (i.e., h is an optimal solution of the linearized problem (PL_d)) and $x(t)$ is a feasible path of the form $x(t) = x^0 + th + o(t)$, then $x(t)$ is an $o(t)$-optimal solution of (P_{u^0+td}). However, the set $\mathcal{S}(PL_d)$ is usually far too large to give useful information about the variation of solutions. Take, for instance, the case of unconstrained minimization problems; then $\mathcal{S}(PL_d)$ is the whole space X. We obtain more precise upper estimates of the optimal value function, that give some insight into the variation of solutions, by considering feasible paths of the form $x(t) = x^0 + th + \frac{1}{2}t^2 w + o(t^2)$.

Define the second-order tangent set, to K at a point y in a direction z, as

$$T_K^2(y, z) := \left\{ w \in Y : \ \text{dist}(y + tz + \tfrac{1}{2}t^2 w, K) = o(t^2) \right\}. \tag{7.1}$$

Note that $T_K^2(y, z)$ is empty if $y \notin K$ or $z \notin T_K(y)$. The expression of the second-order tangent sets to various cones of nonnegative functions may be found in [27] (see a related result in [43]). For the sake of brevity denote

$$T_K^2(h, d) := T_K^2(G(x^0, u^0), DG(x^0, u^0)(h, d)). \tag{7.2}$$

Feasibility of the path $x(t)$ implies that

$$DG(x^0, u^0)(h, d) \in T_K G((x^0, u^0)), \tag{7.3}$$

$$D_x G(x^0, u^0)w + D^2 G(x^0, u^0)((h,d),(h,d)) \in T_K^2(h,d). \tag{7.4}$$

We assume that $h \in \mathcal{S}(PL_d)$. The expansion (7.4) suggests that in order to derive a tight upper estimate for the directional variation of the optimal value function, the best choice for w is to minimize the second-order term of the Taylor expansion of the objective function, while satisfying the constraint (7.4). That is, to take w as an optimal solution of

$$(\mathcal{PQ}_{d,h}) \qquad \begin{array}{ll} \operatorname{Min}_w & D_x f(x^0, u^0)w + D^2 f(x^0, u^0)((h,d),(h,d)) \\ \text{subject to} & D_x G(x^0, u^0)w + D^2 G(x^0, u^0)((h,d),(h,d)) \in T_K^2(h,d). \end{array}$$

Then the best choice for h is to take it as an optimal solution of

$$(\mathcal{PQ}_d) \qquad \operatorname*{Min}_{h \in \mathcal{S}(PL_d)} \operatorname{val}(\mathcal{PQ}_{d,h}).$$

THEOREM 7.1. *Let x^0 be an optimal solution of the problem (P_{u^0}), with which is associated at least one Lagrange multiplier. Let d be a direction of perturbations for which the directional regularity condition holds. Then*

$$\limsup_{t \downarrow 0} \frac{v(u^0 + td) - v(u^0) - t\operatorname{val}(PL_d)}{\frac{1}{2}t^2} \le \operatorname{val}(\mathcal{PQ}_d). \tag{7.5}$$

Note that if $\operatorname{val}(\mathcal{PQ}_d) > -\infty$, then (7.5) is equivalent to

$$v(u^0 + td) \le v(u^0) + t\operatorname{val}(PL_d) + \tfrac{1}{2}t^2 \operatorname{val}(\mathcal{PQ}_d) + o(t^2). \tag{7.6}$$

Note that $\mathcal{S}(PL_d)$ may be empty even if (PL_d) has a finite value. Also, it may happen that $T_K^2(h,d)$ is empty for all $h \in \mathcal{S}(PL_d)$. In that case the above inequality reduces to the trivial estimate $v(u^0 + td) \le +\infty$.

Under the assumptions of Theorem 7.1, it can be proved that, for any $h \in \mathcal{S}(PL_d)$, the optimal value of $(\mathcal{PQ}_{d,h})$ is equal to the optimal value of the problem

$$(\mathcal{DQ}_{d,h}) \qquad \operatorname*{Max}_{\lambda \in \Lambda_1(x^0,d)} \left\{ D^2 L(x^0, \lambda, u^0)((h,d),(h,d)) - \sigma(\lambda, T_K^2(h,d)) \right\},$$

where $\Lambda_1(x^0, d) := \mathcal{S}(DL_d)$ and $\sigma(\cdot, T_K^2(h,d))$ denotes the support function of the set $T_K^2(h,d)$.

THEOREM 7.2. *Let (P_u) be a (finite-dimensional, finitely constrained) nonlinear programming problem and x^0 be the unique solution of (P_{u^0}), and assume that $S(u^0 + td)$ is nonempty and uniformly bounded for small enough $t > 0$. If d is a direction of perturbations for which both the directional regularity condition and the directional second-order sufficient conditions* (6.10), *hold, then*

$$v(u^0 + td) = v(u^0) + t\operatorname{val}(PL_d) + \tfrac{1}{2}t^2 \operatorname{val}(\mathcal{PQ}_d) + o(t^2). \tag{7.7}$$

In addition, the set of optimal solutions of $(\mathcal{PQ}_d)$ coincides with the set

$$\{h \in X : \exists\, x(t) = x^0 + th + o(t);\ x(t) \text{ is an } o(t^2)\text{-optimal solution of } (P_{u^0+td})\}.$$

Note that in the case of finitely many constraints, the "sigma term" vanishes and the optimal value $\operatorname{val}(\mathcal{PQ}_d)$ in (7.7) is given by the optimal value of the problem

$$\operatorname*{Min}_{h \in \mathcal{S}(PL_d)} \left\{ \psi_d(h) := \operatorname*{Max}_{\lambda \in \Lambda_1(x^0,d)} D^2 L(x^0, \lambda, u^0)((h,d),(h,d)) \right\}. \tag{7.8}$$

It follows that if $\bar{x}(t)$ is an $o(t^2)$-optimal solution of (P_{u^0+td}) converging to x_0, then under the assumptions of Theorem 7.2 every accumulation point of $t^{-1}(\bar{x}(t) - x_0)$ is an optimal solution of the program (7.8). In particular, if the optimization problem (7.8) possesses a unique optimal solution $\bar{h} = \bar{h}(d)$, then the directional derivative $\bar{x}'(u_0, d)$ exists and is equal to $\bar{h}(d)$.

These results were obtained first for nonlinear programming problems, under the assumption of linear independence of the gradients of the constraint functions and the strict complementarity condition, by applying the implicit function theorem to the first-order optimality conditions considered as a system of nonlinear equations [33], [32]. The above approach of investigating directional differentiability of optimal solutions by using a second-order expansion of the corresponding optimal value function was employed in [79] under the assumption of uniqueness of Lagrange multipliers and in [80] under the Mangasarian–Fromovitz constraint qualification. The duality method was suggested in [8], where the results were extended further under the directional regularity condition. The last statement of Theorem 7.2 is due to [16]. The results were extended to the Banach space setting in [20, Part I], taking advantage of the second-order analysis of [26]. It was shown recently [22] that Theorem 7.2 is valid for semidefinite and some cases of semi-infinite programming, under the additional assumption that problem (PL_d) has a nonempty set of solutions. In these cases the sigma term, in general, is nonzero. Its expression in the case of semidefinite programming is given in [88].

EXAMPLE 7.1 (see [80]). *Consider the following problem:*

$$\text{(7.9)} \qquad \underset{(x_1,x_2)\in\mathbb{R}^2}{\text{Min}} \ \tfrac{1}{2}(x_1-1)^2 + \tfrac{1}{2}x_2^2 \quad \textit{subject to} \ \ x_1 \le 0, \ x_1 + u_1 x_2 + u_2 \le 0,$$

depending on the parameter vector $(u_1, u_2) \in \mathbb{R}^2$*. For* $u_0 = (0,0)$ *this problem has the unique optimal solution* $x_0 = (0,0)$*. The Mangasarian–Fromovitz constraint qualification holds at* x_0 *and the corresponding set* Λ_0 *of Lagrange multipliers is*

$$\Lambda_0 = \{(\lambda_1, \lambda_2) : \lambda_1 + \lambda_2 = 1, \ \lambda_1 \ge 0, \ \lambda_2 \ge 0\}.$$

Also, the strong second-order sufficient conditions hold and hence the results of Theorem 7.2 apply. We have here

$$D^2 L(x_0, \lambda, u_0)((h,d)(h,d)) = \tfrac{1}{2}h_1^2 + \tfrac{1}{2}h_2^2 + \lambda_2 d_1 h_2,$$

and $\mathcal{S}(PL_d)$ *is the optimal solutions set of the linearized problem*

$$\text{(7.10)} \qquad \text{Min}(-h_1) \quad \text{subject to} \ \ h_1 \le 0, \ h_1 + d_2 \le 0.$$

Moreover, $D_u L(x_0, \lambda, u_0) = (0, \lambda_2)$ *and hence* $\Lambda_1(x_0, d)$ *is formed by the maximizers of* $d_2\lambda_2$ *over* Λ_0*.*

Let $d = (1,0)$*. Then* $\Lambda_1(x_0,d) = \Lambda_0$*,*

$$\psi_d(h) = \max\{\tfrac{1}{2}h_1^2 + \tfrac{1}{2}h_2^2, \tfrac{1}{2}h_1^2 + \tfrac{1}{2}h_2^2 + h_2\},$$

and $\mathcal{S}(PL_d) = \{(h_1,h_2) : h_1 = 0\}$*. It follows that in that case* $\bar{h} = (0,0)$ *and hence* $\bar{x}'(0,d) = (0,0)$*. Now let* $d = (1,\gamma)$ *for some* $\gamma > 0$*. Then* $\Lambda_1(x_0,d) = \{(0,1)\}$*,* $\psi_d(h)(h) = \frac{1}{2}h_1^2 + \frac{1}{2}h_2^2 + h_2$*, and* $\mathcal{S}(PL_d) = \{(h_1,h_2) : h_1 = -\gamma\}$*. It follows that in that case* $\bar{h} = (-\gamma, -1)$*, and hence* $\bar{x}'(0,d) = (-\gamma,-1)$*. We see that in this example the optimal solution* $\bar{x}(u)$ *is directionally differentiable at* $u = 0$*, but the directional* *derivative $\bar{x}'(0,d)$ is a discontinuous function of the direction d. It follows that $\bar{x}(u)$ is not directionally differentiable in the Hadamard sense and is not Lipschitz continuous in any neighborhood of $u = 0$. A reason for such discontinuous behavior of $\bar{x}'(0,\cdot)$ is that the optimal set $\Lambda_1(x_0, d)$ depends on d in a discontinuous way.*

Let us make the following observations. If the Lagrange multiplier is unique, then the function $\psi_d(\cdot)$ is quadratic and (7.8) becomes a quadratic programming problem. If the assumptions of Proposition 5.4 hold, then the results of this section are a consequence of those proved for strongly regular solutions. If the strict complementarity condition holds, then we recover the result of Fiacco and McCormick [33], [32].

7.2. First-order differentiability of the optimal value function and Hölder expansion of optimal solutions by the second-order analysis. Let x^0 be a solution of problem (P_{u^0}) and d be a direction of perturbation for which directional regularity holds. As was observed in Example 4.3, it may happen that the upper bound (4.15), for the first-order directional behavior of the optimal value function $v(u)$, is not tight. In this section we derive a sharper estimate by employing feasible paths of the form $x(t) = x^0 + t^{1/2}h + tw + o(t)$. By expanding $G(x(t), u + td)$ we obtain that a necessary condition for this path to be feasible is that

$$\begin{gathered} D_x G(x^0, u^0)h \in T_K(G(x^0, u^0)), \\ 2DG(x^0, u^0)(w, d) + D^2_{xx}G(x^0, u^0)(h, h) \in \mathcal{T}^2_K(h), \end{gathered} \tag{7.11}$$

where we have set

$$\mathcal{T}^2_K(h) := T^2_K(G(x^0, u^0), D_x G(x^0, u^0)h).$$

(Note that $\mathcal{T}^2_K(h)$ should not be confused with the set $T^2_K(h, d)$ introduced earlier.) In order to deal with this type of path, we need the following condition. A direction of perturbations $d \in U$ is said to satisfy the *strong directional regularity* condition [20] if it satisfies the directional regularity condition and, in addition, whenever (h, w) is such that (7.11) holds, then for $\gamma < 1$ and arbitrarily close to 1, one can find $w_\gamma \in X$ with $w_\gamma \to w$ and feasible paths of the form $x^\gamma(t) = x_0 + \gamma t^{1/2}h + tw_\gamma + o(t)$. This condition is equivalent to directional regularity whenever $K = \{0\} \times K_2$, with K_2 being a closed convex cone with nonempty interior. In particular, for nonlinear programming problems, this is equivalent to Gollan's condition (3.13). It is not known whether both conditions are equivalent in general.

The above expansion suggests that in order to derive an upper estimate for the directional behavior of the optimal value function the best choice is to take h in the critical cone C_0, and then for a given h, to take w that minimizes the second-order term of the expansion of the objective function, while satisfying the above constraint. That is, to take w as an optimal solution of

$$(\mathcal{PL}_{d,h}) \qquad \begin{array}{ll} \operatorname{Min}_w & 2Df(x^0, u^0)(w, d) + D^2_{xx}f(x^0, u^0)(h, h) \\ \text{subject to} & 2DG(x^0, u^0)(w, d) + D^2_{xx}G(x^0, u^0)(h, h) \in \mathcal{T}^2_K(h), \end{array}$$

and then the best choice for h is to take it as an optimal solution of

$$(\mathcal{PL}_d) \qquad \operatorname*{Min}_{h \in C_0} \operatorname{val}(\mathcal{PL}_{d,h}).$$

THEOREM 7.3. *Let x^0 be an optimal solution of the problem (P_{u^0}), with which is associated at least one Lagrange multiplier. Let d be a direction of perturbations for* *which strong directional regularity holds. Then*

$$\limsup_{t \downarrow 0} \frac{v(u^0 + td) - v(u^0)}{t} \le \frac{1}{2} \operatorname{val}(\mathcal{PL}_d).$$

It may happen that $\mathcal{T}_K^2(h)$ is empty for certain critical directions h. However, for $h = 0$ we note that problem $(\mathcal{PL}_{d,0})$ coincides with the linearized problem (PL_d). Therefore, we have that $\operatorname{val}(\mathcal{PL}_d) \le \operatorname{val}(\mathcal{PL}_{d,0}) = 2\operatorname{val}(PL_d)$, the latter being less than $+\infty$ under the directional qualification condition.

The dual of problem $(\mathcal{PL}_{d,h})$ is

$$(\mathcal{DL}_{d,h}) \qquad \operatorname*{Max}_{\lambda \in \Lambda(x^0,u^0)} \left\{ 2D_u L(x^0, \lambda, u^0)d + D^2_{xx} L(x^0, \lambda, u^0)(h,h) - \sigma(\lambda, \mathcal{T}_K^2(h)) \right\}.$$

It is not known whether $\operatorname{val}(\mathcal{PL}_{d,h})$ is equal to $\operatorname{val}(\mathcal{DL}_{d,h})$ for any $h \in C_0$. It may be proved, however, that under the hypotheses of Theorem 7.3, we have

$$\operatorname{val}(\mathcal{PL}_d) = \inf_{h \in C_0} \operatorname{val}(\mathcal{DL}_{d,h}).$$

THEOREM 7.4. *Let (P_u) be a (finite-dimensional, finitely constrained) nonlinear programming problem and x^0 be the unique solution of (P_{u^0}), with which is associated at least one Lagrange multiplier, and suppose that $S(u^0 + td)$ is nonempty and uniformly bounded for small enough $t > 0$. If d is a direction of perturbation for which the strong directional regularity and the second-order conditions (Definition* 6.2*) both hold, then*

$$v(u^0 + td) = v(u^0) + \tfrac{1}{2} t \operatorname{val}(\mathcal{PL}_d) + o(t).$$

In addition, the set of solutions of $(\mathcal{PL}_d)$ coincides with the set

$$\{h \in X :\ \exists\, x(t) = x^0 + t^{1/2}h + o(t^{1/2});\ x(t) \text{ is an } o(t)\text{-optimal solution of } (P_{u^0+td})\}.$$

The above results were obtained first in the setting of nonlinear programming problems. The expansion of the optimal value function (under a different form) is due to [35, 37], while in [16] the expansion of approximate optimal solutions was obtained. Part II of [20] extends the result to the Banach space setting.

7.3. Square root expansion of the optimal value function and Hölder expansion of optimal solutions. We now consider the case when directional regularity condition holds, but the set of Lagrange multipliers is *empty*. By Proposition 4.3, we already know that in that case the directional derivative $v'(u_0, d) = -\infty$. We show in this section that a more precise result holds. Namely, variations of the optimal value of (P_{u^0+td}) are of the order $t^{1/2}$, and $(v(u^0+td) - v(u^0))/t^{1/2}$ has a limit in the case of nonlinear programming problems, assuming the second-order conditions (Definition 6.2) hold.

Again our analysis is based on considering feasible paths of the form $x(t) = x^0 + t^{1/2}h + tw + o(t)$. Then, as before, (7.11) holds. We wish to minimize the first term of the expansion of the objective function, i.e., $D_x f(x^0, u^0)h$. Since we want this amount to be negative, we assume that $h \in C_0$. We may write the corresponding problem as

$$(\mathcal{PS}_{d,h}) \qquad \begin{array}{l} \operatorname{Min}_w D_x f(x^0, 0)h \\ \text{subject to } 2DG(x^0, u^0)(w, d) + D^2_{xx} G(x^0, u^0)(h,h) \in \mathcal{T}_K^2(h). \end{array}$$

For a given h, this is in fact a feasibility problem (since the objective function does not depend on w). Minimizing over h we obtain the problem

$$(\mathcal{PS}_d) \qquad \operatorname*{Min}_{h \in C_0} \operatorname{val}(\mathcal{PS}_{d,h}).$$

THEOREM 7.5. *Let x^0 be an optimal solution of the problem (P_{u^0}), with which is associated at least one singular multiplier, but no Lagrange multipliers. Let d be a direction of perturbations for which strong directional regularity condition holds. Then* $\operatorname{val}(\mathcal{PS}_d) < 0$ *and*

$$\limsup_{t \downarrow 0} \frac{v(u^0 + td) - v(u^0)}{t^{1/2}} \leq \operatorname{val}(\mathcal{PS}_d).$$

Under the assumptions of Theorem 7.5, it can be shown that the optimal value of $(\mathcal{PS}_d)$ is equal to the optimal value of the following problem (with $\lambda_0 = 0$):

$$(\mathcal{DS}_{d,h})$$

$$\operatorname*{Min}_{h \in C_0} D_x f(x^0, 0)h + \sup_{(\lambda_0, \lambda) \in \Lambda^g(x^0)} \left\{ \begin{array}{l} 2D_u L^g(x^0, \lambda_0, \lambda, u^0)d \\ + D^2_{xx} L^g(x^0, \lambda_0, \lambda, u^0)(h,h) - \sigma(\lambda, \mathcal{T}^2_K(h)) \end{array} \right\}.$$

THEOREM 7.6. *Let (P_u) be a (finite-dimensional, finitely constrained) nonlinear programming problem and x^0 be the unique optimal solution of (P_{u^0}), and suppose that $S(u^0 + td)$ is nonempty and uniformly bounded for small enough $t > 0$. If d is a direction of perturbations for which the strong directional regularity condition holds, no Lagrange multiplier is associated with x^0, and the second-order condition (Definition* 6.2*) is satisfied, then*

$$v(u^0 + td) = v(u^0) + t^{1/2} \operatorname{val}(\mathcal{PS}_d) + o(t^{1/2}).$$

In addition, the set of solutions of $(\mathcal{PS}_d)$ coincides with the set

$$\left\{ h \in X : \begin{array}{l} \exists\, x(t) = x^0 + t^{1/2} h + o(t^{1/2}), \\ x(t) \textit{ is an } o(t^{1/2})\textit{-optimal solution of } (P_{u^0 + td}) \end{array} \right\}.$$

These results were obtained in [13], in the nonlinear programming setting and generalized in [20, Part II]. A major difference with the previously studied situations is that the variation of the cost is of order $t^{1/2}$.

8. Some illustrations.

We apply some of the previous results to the computation of the expansion of the chain when the unperturbed position is either vertical or horizontal.

8.1. Application to the perturbation of the vertical chain.

Equivalence to a directionally qualified problem. It is easy to show that the position of the chain is the same as if we change the last equality into an inequality, obtaining the following optimization problem:

$$(8.1) \qquad \begin{array}{l} \operatorname{Min}_{y,z} E(y,z) \\ \text{subject to } \ \tfrac{1}{2} y_k^2 + \tfrac{1}{2} z_k^2 - \tfrac{1}{2} = 0, \ k = 1, \ldots, m, \ \ \sum_{k=1}^m y_k = 0, \\ m - u - \sum_{k=1}^m z_k \leq 0. \end{array}$$

As u is scalar, we may take $d = 1$, and therefore here $t = u$. It appears that the new problem (8.1) satisfies the directional qualification condition, as we now check, although problem (2.2) does not. Indeed, the linearized equality constraints are onto, as solving

$$h_k^z = \gamma_k, \quad k = 1, \ldots, m, \quad \sum_{k=1}^{m} h_k^y = \beta, \ \gamma \in \mathbb{R}^m, \ \beta \in \mathbb{R} \tag{8.2}$$

obviously has a solution for any (γ, β). If $(\gamma, \beta) = 0$, then $h_k^z = 0$. It follows that a displacement in the kernel of linearized equality constraints cannot strictly satisfy the linearized active inequality constraint $-\sum_k h_k^z \leq 0$. Therefore, qualification does not hold. On the other hand, a directional qualification condition holds as (3.3) is satisfied with $h = (h^y, h^z) = 0$.

Lagrange multipliers, critical cone, and second-order conditions. The expression for the Lagrangian function $L(y, z, \lambda, u)$ is

$$\sum_{k=1}^{m} \alpha_k z_k + \sum_{k=1}^{m} \lambda_k \left(\frac{1}{2} y_k^2 + \frac{1}{2} z_k^2 - \frac{1}{2} \right) + \lambda_{m+1} \sum_{k=1}^{m} y_k + \lambda_{m+2} \left(m - u - \sum_{k=1}^{m} z_k \right).$$

Its derivatives, with respect to (y, z), are null whenever

$$\lambda_{m+1} = 0 \ \text{ and } \ \alpha_k + \lambda_k - \lambda_{m+2} = 0, \quad k = 1, \ldots, m.$$

Therefore, the set of Lagrange multipliers is

$$\Lambda = \{\lambda \in \mathbb{R}^{m+2} : \lambda_{m+1} = 0, \ \lambda_{m+2} \geq 0, \ \lambda_k = \lambda_{m+2} - \alpha_k, \quad k = 1, \ldots, m\}.$$

The equations of the critical cone are

$$\sum_{k=1}^{m} \alpha_k h_k^z \leq 0, \quad h_k^z = 0, \ k = 1, \ldots, m, \quad \sum_{k=1}^{m} h_k^y = 0, \quad \sum_{k=1}^{m} h_k^z \leq 0,$$

i.e.,

$$C_0 = \left\{ (h^y, h^z) : \ h^z = 0 \text{ and } \sum_{k=1}^{m} h_k^y = 0 \right\}.$$

Let us now check the second-order sufficient condition (Definition 6.2). Given a critical direction h, using $h^z = 0$, we compute

$$D_{xx}^2 L(x^0, \lambda, 0)(h, h) = \sum_{k=1}^{m} \lambda_k (h_k^y)^2.$$

If λ_{m+2} is large enough, then $\lambda_k \geq 1$, $k = 1, \ldots, m$. In that case the second-order variation of the Lagrangian is nonzero when h is a nonzero critical direction; i.e., the second-order sufficient condition holds.

It may be checked that the strong directional second-order sufficient condition does not hold.

Study of an auxiliary subproblem. We just checked that the hypotheses of Theorem 7.4 are satisfied. Let us make its result explicit by computing the value and set of solutions of $(\mathcal{PL}_d)$ or, equivalently, of $(\mathcal{DL}_d)$. The expression of the latter is

$$\operatorname*{Min}_{h \in C_0} \sup_{\lambda \in \Lambda(x^0)} \Delta(h, \lambda),$$

where

$$\Delta(h,\lambda) := -2\lambda_{m+2} + \sum_{k=1}^{m} \lambda_k (h_k^y)^2 = \lambda_{m+2}\left(\sum_{k=1}^{m}(h_k^y)^2 - 2\right) - \sum_{k=1}^{m} \alpha_k (h_k^y)^2.$$

If $\sum_k (h_k^y)^2 > 2$, then $\sup\{\Delta(h,\lambda)\ ;\ \lambda \in \Lambda(x^0)\} = +\infty$. Otherwise, the sup is attained for $\lambda_{m+2} = 0$. Therefore, the subproblem has the same solutions as

$$\operatorname*{Min}_{h\in C_0} -\frac{1}{2}\sum_{k=1}^{m} \alpha_k (h_k^y)^2, \qquad \sum_{k=1}^{m}(h_k^y)^2 \le 2.$$

Using the expression of C_0, we formulate this problem as

$$\operatorname*{Min}_{h^y} -\frac{1}{2}\sum_{k=1}^{m} \alpha_k (h_k^y)^2, \quad \frac{1}{2}\sum_{k=1}^{m}(h_k^y)^2 - 1 \le 0, \quad \sum_{k=1}^{m} h_k^y = 0.$$

The problem has solutions (compact nonempty feasible set), a concave objective function, and convex constraints, and the constraints are regular at every feasible point. With a solution h^y are associated multipliers (η,ν) with $\eta \ge 0$ such that

$$-\alpha_k h_k^y + \eta h_k^y + \nu = 0, \qquad k = 1,\dots,m. \tag{8.3}$$

Let us prove that $\nu \ne 0$. If $\nu = 0$, then $(\eta - \alpha_k)h_k^y = 0, k = 1,\dots,m$. All components of α_k having different values, we deduce that $h^y = 0$ except for at most one component. From the equality constraint it follows that $h^y = 0$. Then $\eta = 0$ (as the inequality constraint is not active). However, the second-order necessary condition does not hold at $h^y = 0$. It follows that ν cannot be null, as had to be proved. As $\nu \ne 0$, we deduce from (8.3) that

$$h_k^y = \nu/(\alpha_k - \eta), \qquad k = 1,\dots,m. \tag{8.4}$$

Summing over k, we get

$$0 = \sum_{k=1}^{m} h_k^y = \nu \sum_{k=1}^{m} \frac{1}{\alpha_k - \eta}. \tag{8.5}$$

As $\nu \ne 0$, this gives a scalar equation with the (one) scalar unknown η

$$\sum_{k=1}^{m} \frac{1}{\alpha_k - \eta} = 0.$$

Let us study this equation, remembering that α_k is a strictly decreasing sequence. If $\eta < \alpha_m$ or $\eta > \alpha_m$ all terms of the sum are of the same sign. Here a solution must be in (α_m, α_1) and cannot be equal to some α_k. Let us fix k in $\{1,\dots,m-1\}$. We study the solution for $\eta \in (\alpha_{k+1}, \alpha_k)$. The left-hand side of the equation is strictly increasing from $-\infty$ to $+\infty$. It follows that (8.5) has a unique solution in each interval (α_{k+1}, α_k). This solution will be said to be the kth solution.

We determine the value of ν by plugging the expression of h^y given by (8.4) into the active inequality constraint. It follows that

$$\nu = {}^{+}_{-}\sqrt{2/\sum_{k=1}^{m}(\alpha_k - \eta)^{-2}}.$$

By (8.4) the two values of ν are associated with opposite values of h^y of equal cost. This is natural due to the symmetry in the problem. The components of h^y corresponding to the kth solution have the same sign from 1 to k, and then the opposite sign.

It is intuitively clear that the best solution is the first one. This is easy to prove if we observe, from the data of the subproblem, that $k \to |h_k^y|$ must be decreasing. Otherwise, applying a permutation to the components of h^y that makes them decreasing with k, we get another feasible point associated with a strictly inferior value of the cost. In summary, h_1^y is, say, negative, and all other components of h^y are positive and, by (8.4), strictly decreasing.

As the subproblem has two solutions, application of Theorem 7.4 gives that any $o(u)$-optimal path is, for $u \geq 0$, of the form $x^0 + s(u)\sqrt{u}h + o(\sqrt{u})$, where $s(u)$ has values in $\{-1, 1\}$.

8.2. Application to the perturbation of the horizontal chain.

Equivalence to a directionally qualified problem. It is easily shown that an equivalent problem is obtained by changing the one before the last equality into an inequality, i.e., optimization problem:

$$\begin{aligned} &\mathrm{Min}_{y,z}\, E(y,z) \\ &\text{subject to } \ \tfrac{1}{2}y_k^2 + \tfrac{1}{2}z_k^2 - \tfrac{1}{2} = 0,\ k = 1, \dots, m,\ \ \textstyle\sum_{k=1}^m z_k = 0, \\ &m - u - \textstyle\sum_{k=1}^m y_k \leq 0. \end{aligned} \tag{8.6}$$

Again we may take $d = 1$ and $t = u$. By arguments similar to those used in the case of the vertical chain, we may check that the new problem (8.1) satisfies the directional qualification hypothesis.

Lagrange multipliers, critical cone, and second-order conditions. The expression of the generalized Lagrangian function $L(y, z, \lambda_0, \lambda, u)$ is

$$\lambda_0 \sum_{k=1}^m \alpha_k z_k + \sum_{k=1}^m \lambda_k \left(\frac{1}{2}y_k^2 + \frac{1}{2}z_k^2 - \frac{1}{2} \right) + \lambda_{m+1} \sum_{k=1}^m z_k + \lambda_{m+2} \left(m - u - \sum_{k=1}^m y_k \right).$$

Its derivatives, with respect to (y, z), are null whenever

$$\lambda_k - \lambda_{m+2} = 0, \quad \lambda_0 \alpha_k + \lambda_{m+1} = 0, \quad k = 1, \dots, m.$$

Therefore, the set of Lagrange multipliers is empty, while the set of generalized Lagrange multipliers is

$$\Lambda^g = \{(\lambda_0, \lambda) \in \mathbb{R}^{m+2} : \lambda_0 = \lambda_{m+1} = 0;\ \lambda_{m+2} \geq 0;\ \lambda_k = \lambda_{m+2}, \quad k = 1, \dots, m\}.$$

The critical cone is

$$C_0 = \left\{ (h^y, h^z)\ ;\ h^y = 0;\ \sum_{k=1}^m \alpha_k h_k^z \leq 0;\ \sum_{k=1}^m h_k^z = 0 \right\}.$$

Let us now check the second-order sufficient condition (Definition 6.2). Given a critical direction h, using $h^y = 0$, we have

$$D_{xx}^2 L(x^0, \lambda, 0)(h, h) = \sum_{k=1}^m \lambda_k (h_k^z)^2.$$

Whenever $\lambda_{m+2} > 0$, as $\lambda_k = \lambda_{m+2}, k = 1, \ldots, m$, this is nonzero when h is a nonzero critical direction; i.e., the (singular) second-order sufficient condition holds.

Study of an auxiliary subproblem. We have shown that the hypotheses of Theorem 7.6 are satisfied. Let us make its result explicit by computing the value and set of solutions of $(\mathcal{PS}_d)$ or, equivalently, of $(\mathcal{DS}_d)$. The expression of the latter is

$$\operatorname*{Min}_{h \in C_0} \sup_{\lambda \in \Lambda^g(x^0)} \Delta(h, \lambda),$$

where

$$\Delta(h, \lambda) := \sum_{k=1}^{m} \alpha_k h_k^z + \lambda_{m+2} \left(\sum_{k=1}^{m} (h_k^z)^2 - 2 \right).$$

If $\sum_k (h_k^z)^2 > 2$, then $\sup\{\Delta(h, \lambda) \ ; \ \lambda \in \Lambda(x^0)\} = +\infty$. Otherwise, the sup is attained for $\lambda_{m+2} = 0$. Therefore, the subproblem has the same solutions as

$$\operatorname*{Min}_{h \in C_0} \sum_{k=1}^{m} \alpha_k h_k^z, \qquad \sum_{k=1}^{m} (h_k^z)^2 \leq 2.$$

Using the expression of C_0, we formulate this problem as

$$\operatorname*{Min}_{h^z} \sum_{k=1}^{m} \alpha_k h_k^z, \quad \frac{1}{2} \sum_{k=1}^{m} (h_k^z)^2 - 1 \leq 0, \quad \sum_{k=1}^{m} h_k^z = 0.$$

This is a qualified convex problem. The solution is better expressed in terms of the variation of altitude of the endpoint of the links. The altitude of the end of the kth link is $\sum_{i=1}^{k} z_i$. The associated displacement $\bar{h}_k := \sum_{i=1}^{k} h_i^z$ is a solution of the problem

$$\operatorname*{Min}_{\bar{h}} \sum_{k=1}^{m} \bar{h}_k, \qquad \frac{1}{2} \sum_{k=1}^{m} (\bar{h}_k - \bar{h}_{k-1})^2 - 1 \leq 0,$$

where we have set $\bar{h}_0 := 0$. Let α be the Lagrange multiplier associated with the constraint. Then $\alpha > 0$ and $\bar{h}$ is solution of

$$\operatorname{Min} \sum_{k=1}^{m} \bar{h}_k + \frac{\alpha}{2} \sum_{k=1}^{m} (\bar{h}_k - \bar{h}_{k-1})^2.$$

Therefore, we may compute $\bar{h}$ as follows: let $\hat{h}$ be the solution of the linear system

$$\hat{h}_{k+1} - 2\hat{h}_k + \hat{h}_{k-1} = 1, \qquad \hat{h}_0 = \hat{h}_m = 0.$$

Note that this linear system is similar to the one obtained when discretizing the one-dimensional Poisson equation with Dirichlet boundary conditions and constant right-hand side. Compute the associated energy $\eta := \frac{1}{2} \sum_{k=1}^{m} (\hat{h}_k - \hat{h}_{k-1})^2$. Then $\bar{h} := \eta^{-1/2} \hat{h}$ is the solution of the above linear system. Let h^z be the corresponding displacement on z. Then, by Theorem 7.6, we have $x(u) = x(0) + \sqrt{u}(0, h^z) + o(\sqrt{u})$.

9. Comments and conclusion. Sensitivity analysis for optimization problems is a very active field, and while writing this article we had to be selective. Among important contributions that were not quoted yet, let us mention the work based on epigraphical analysis [4, 5, 6, 90] and applications of polyhedricity theory [65, 39] and its generalizations [10, 20].

We have carefully reviewed an approach to sensitivity analysis that uses lower and upper estimates of the optimal value function. For nonlinear programming problems, this gives strong results, and the theory seems to be more or less complete now. These results are extended to semidefinite programming problems in [22, 88], and to semi-infinite programming in [20, Part III], [22, 84]. At the same time, Example 6.2 suggests that Hölder expansion of order different from $1/2$ should be considered. The fact that many examples of semidefinite programs involve linear objective functions and constraints [24] is another challenge, since the set of optimal solutions of the unperturbed program might typically be nonisolated. Yet it is possible to derive Lipschitz stability of optimal solutions of semidefinite programs under the second-order growth condition and a nondegeneracy assumption [23].

REFERENCES

[1] W. Alt, *Lipschitzian perturbations of infinite optimization problems*, in Mathematical Programming with Data Perturbations, A. V. Fiacco, ed., Marcel Dekker, New York, 1983, pp. 7–21.

[2] W. Alt, *Stability of solutions for a class of nonlinear cone constrained optimization problems, Part* 1: *Basic theory*, Numer. Funct. Anal. Optim., 10 (1989), pp. 1053–1064.

[3] W. Alt, *Local stability of solutions to differentiable optimization problems in Banach spaces*, J. Optim. Theory Appl., 70 (1991), pp. 443–466.

[4] H. Attouch, *Variational Convergence for Functions and Operators*, Pitman, Boston, MA, 1984.

[5] H. Attouch and R. J.-B. Wets, *Quantitative stability of variational systems:* II. *A framework for nonlinear conditioning*, SIAM J. Optim., 3 (1993), pp. 359–381.

[6] H. Attouch, *Viscosity solution of minimization problems*, SIAM J. Optim., 6 (1996), pp. 769–806.

[7] J. P. Aubin and I. Ekeland, *Applied Nonlinear Analysis*, Wiley, New York, 1984.

[8] A. Auslender and R. Cominetti, *First and second order sensitivity analysis of nonlinear programs under directional constraint qualification conditions*, Optim., 21 (1990), pp. 351–363.

[9] V. I. Averbukh and O. G. Smolyanov, *The various definitions of the derivative in linear topological spaces*, Russian Math. Surveys, 23 (1968), pp. 67–113.

[10] L. Barbet and R. Janin, *Analyse de sensibilité différentielle pour un problème d'optimisation paramétré en dimension infinie*, Comptes Rendus Acad. Sci. Paris Sér. I, 318 (1993), pp. 221–226.

[11] A. Ben-Tal, *Second-order and related extremality conditions in nonlinear programming*, J. Optim. Theory Appl., 31 (1980), pp. 143–165.

[12] J. F. Bonnans, *A semi-strong sufficiency condition for optimality in non convex programming and its connection to the perturbation problem*, J. Optim. Theory Appl., 60 (1989), pp. 7–18.

[13] J. F. Bonnans, *Directional derivatives of optimal solutions in smooth nonlinear programming*, J. Optim. Theory Appl., 73 (1992), pp. 27–45.

[14] J. F. Bonnans, *Second order analysis for control constrained optimal control problems of semilinear elliptic systems*, J. Appl. Math. Optim., to appear.

[15] J. F. Bonnans and E. Casas, *Contrôle de systèmes elliptiques semilinéaires comportant des contraintes sur l'état*, in Nonlinear Partial Differential Equations and Their Applications, Collège de France Seminar Vol. VIII, H. Brézis and J. L. Lions, eds., Pitman Res. Notes Math. Ser. 166, Longman Scientific and Technical, New York, 1988, pp. 69–86.

[16] J. F. Bonnans, A. D. Ioffe, and A. Shapiro, *Expansion of exact and approximate solutions in nonlinear programming*, in Proc. French-German Conference on Optim., D. Pallaschke, ed., Lecture Notes in Economics and Mathematical Systems, Springer-Verlag, Berlin, 1992, pp. 103–117.

[17] J. F. Bonnans and A. D. Ioffe, *Second-order sufficiency and quadratic growth for non isolated minima*, Math. Oper. Res., 20 (1995), pp. 801–817.

[18] J. F. Bonnans and A. D. Ioffe, *Quadratic growth and stability in convex programming problems with multiple solutions*, J. Convex Anal. (Special issue dedicated to R. T. Rockafellar), 2 (1995), pp. 41–57.

[19] J. F. Bonnans and A. Sulem, *Pseudopower expansion of solutions of generalized equations and constrained optimization problems*, Math. Programming Ser. A, 70 (1995), pp. 123–148.

[20] J. F. Bonnans and R. Cominetti, *Perturbed optimization in Banach spaces* I: *A general theory based on a weak directional constraint qualification;* II: *A theory based on a strong directional qualification;* III: *Semi-infinite optimization*, SIAM J. Control Optim., 34 (1996), pp. 1151–1171, 1172–1189, and 1555–1567.

[21] J. F. Bonnans, R. Cominetti, and A. Shapiro, *Second order necessary and sufficient optimality conditions under abstract constraints*, SIAM J. Optim., to appear.

[22] J. F. Bonnans, R. Cominetti, and A. Shapiro, *Sensitivity analysis of optimization problems under second order regular constraints*, Math. Oper. Res., to appear.

[23] J. F. Bonnans and A. Shapiro, *Nondegeneracy and quantitative stability of parameterized optimization problems with multiple solutions*, SIAM J. Optim., to appear.

[24] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan, *Linear matrix inequalities*, Stud. Appl. Math. 15, SIAM, Philadelphia, PA, 1994.

[25] S. J. V. Burke and M. C. Ferris, *Weak sharp minima in mathematical programming*, SIAM J. Control Optim., 5 (1993), pp. 1340–1359.

[26] R. Cominetti, *Metric regularity, tangent sets and second order optimality conditions*, Appl. Math. Optim., 21 (1990), pp. 265–287.

[27] R. Cominetti and J. P. Penot, *Tangent sets to unilateral convex sets*, Comptes Rendus Acad. Sci. Paris Sér. I, 321 (1995), pp. 1631–1636.

[28] J. M. Danskin, *The Theory of Max-Min and Its Applications to Weapons Allocation Problems*, Springer-Verlag, New York, 1967.

[29] V. F. Dem'yanov and A. B. Pevny, *Marginal values of problems of mathematical programming and minimax problems*, Vestnik LGU Mat. Mekh. Astron., 4 (1973), pp. 31–45.

[30] V. F. Dem'yanov and V. N. Malozemov, *Introduction to Minimax*, Wiley, New York, 1974.

[31] A. L. Dontchev and R. T. Rockafellar, *Characterizations of strong regularity for variational inequalities over polyhedral convex sets*, SIAM J. Optim., 6 (1996), pp. 1087–1105.

[32] A. V. Fiacco, *Introduction to Sensitivity and Stability Analysis in Nonlinear Programming*, Academic Press, New York, 1983.

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

[34] J. Gauvin and F. Dubeau, *Differential properties of the marginal function in mathematical programming*, Math. Programming Stud., 19 (1982), pp. 101–119.

[35] J. Gauvin and R. Janin, *Directional behaviour of optimal solutions in nonlinear mathematical programming*, Math. Oper. Res., 13 (1988), pp. 629–649.

[36] J. Gauvin and J. W. Tolle, *Differential stability in nonlinear programming*, SIAM J. Control Optim., 15 (1977), pp. 294–311.

[37] B. Gollan, *On the marginal function in nonlinear programming*, Math. Oper. Res., 9 (1984), pp. 208–221.

[38] E. G. Gol'shtein, *Theory of Convex Programming*, Transl. Math. Monographs 36, American Mathematical Society, Providence, RI, 1972.

[39] A. Haraux, *How to differentiate the projection on a convex set in Hilbert space. Some applications to variational inequalities*, J. Math. Soc. Japan, 29 (1977), pp. 615–631.

[40] R. Hettich and K. O. Kortanek, *Semi-infinite programming: Theory, methods and applications*, SIAM Rev., 35 (1993), pp. 380–429.

[41] R. Hettich and P. Zencke, *Numerische Methoden der Approximation und Semi-infiniten Optimierung*, Teubner, Stuttgart, 1982.

[42] A. D. Ioffe, *Necessary and sufficient conditions for a local minimum* 3: *Second order conditions and augmented duality*, SIAM J. Control Optim., 17 (1979), pp. 266–288.

[43] A. D. Ioffe, *Variational analysis of a composite function: A formula for the lower second order epi-derivative*. J. Math. Anal. Appl., 160 (1991), pp. 379–405.

[44] A. D. Ioffe, *On sensitivity analysis of nonlinear programs in Banach spaces: The approach via composite unconstrained optimization*, SIAM J. Optim., 4 (1994), pp. 1–43.

[45] K. Jittorntrum, *Solution point differentiability without strict complementarity in nonlinear programming*, Math. Programming, 21 (1984), pp. 127–138.

[46] F. John, *Extremum problems with inequalities as subsidiary conditions*, in Studies and Essays, R. Courant Anniversary Volume, Interscience, New York, 1948, pp. 187–204.

[47] H. Th. Jongen, D. Klatte, and K. Tammer, *Implicit functions and sensitivity of stationary points*, Math. Programming, 19 (1990), pp. 123–138.

[48] H. Kawasaki, *An envelope-like effect of infinitely many inequality constraints on second order necessary conditions for minimization problems*, Math. Programming, 41 (1988), pp. 73–96.

[49] D. Klatte and K. Tammer, *Strong stability of stationary solutions and Karush-Kuhn-Tucker points in nonlinear optimization*, Ann. Oper. Res., 27 (1990), pp. 285–308.

[50] D. Klatte, *On quantitative stability for non-isolated minima*, Control Cybernetics, 23 (1994), pp. 183–200.

[51] M. Kojima, *Strongly stable stationary solutions in nonlinear programming*, in Analysis and Computation of Fixed Points, S. M. Robinson, ed., Academic Press, New York, 1988, pp. 93–138.

[52] B. Kummer, *Lipschitzian inverse functions directional derivatives and application in $\mathcal{C}^{1,1}$ optimization*, J. Optim. Theory Appl., 70 (1991), pp. 561–582.

[53] F. Lempio and H. Maurer, *Differential stability in infinite-dimensional nonlinear programming*, Appl. Math. Optim., 6 (1980), pp. 139–152.

[54] E. S. Levitin, *On differential properties of the optimal value of parametric problems of mathematical programming*, Dokl. Akad. Nauk SSSR, 224 (1975), pp. 1354–1358.

[55] E. S. Levitin, *Differentiability with respect to a parameter of the optimal value in parametric problems of mathematical programming*, Kibernetika, 12 (1976), pp. 44–59.

[56] E. S. Levitin, *Perturbation Theory in Mathematical Programming and Its Applications*, Wiley, New York, 1994.

[57] D. G. Luenberger, *Introduction to linear and nonlinear programming*, Addison-Wesley, Reading, MA, 1973.

[58] K. Malanowski, *Second order conditions and constraint qualifications in stability and sensitivity analysis of solutions to optimization problems in Hilbert spaces*, Appl. Math. Optim., 25 (1992), pp. 51–79.

[59] K. Malanowski, *Two-norm approach in stability and sensitivity analysis of optimization and optimal control problems*, Adv. Math. Sci. Appl., 2 (1993), pp. 397–443.

[60] K. Malanowski, *Sufficient optimality conditions for optimal control subject to state constraints*, SIAM J. Control Optim., 35 (1997), pp. 205–227.

[61] O. Mangasarian and S. Fromovitz, *The Fritz-John necessary optimality conditions in the presence of equality and inequality constraints*, J. Math. Anal. Appl., 7 (1967), pp. 37–47.

[62] H. Maurer, *First and second order sufficient optimality conditions in mathematical programming and optimal control*, Math. Programming Stud., 14 (1981), pp. 163–177.

[63] H. Maurer and S. Pickenhain, *Second-order sufficient conditions for control problems with mixed control-state constraints*, J. Optim. Theory Appl., 86 (1996), pp. 649–667.

[64] H. Maurer and J. Zowe, *First and second-order necessary and sufficient optimality conditions for infinite-dimensional programming problems*, Math. Programming, 16 (1979), pp. 98–110.

[65] F. Mignot, *Contrôle dans les inéquations variationnelles elliptiques*, J. Functional Anal., 22 (1976), pp. 25–39.

[66] B. Mordukhovich, *Stability theory for parametric generalized equations and variational inequalities via nonsmooth analysis*, Trans. Amer. Math. Soc., 343 (1994), pp. 606–657.

[67] M. Z. Nashed, *Differentiability and related properties of nonlinear operators: Some aspects of the role of differentials in nonlinear functional analysis*, in Nonlinear Functional Anal. Appl., L. B. Rall, ed., Academic Press, New York, 1971, pp. 103–309.

[68] J. P. Penot, *On regularity conditions in mathematical programming*, Math. Programming Stud., 19 (1982), pp. 167–199.

[69] B. N. Pshenichnyi, *Necessary Conditions for an Extremum*, Marcel Dekker, New York, 1971.

[70] S. M. Robinson, *Normed convex processes*, Trans. Amer. Math. Soc., 174 (1972), pp. 127–140.

[71] S. M. Robinson, *First order conditions for general nonlinear optimization*, SIAM J. Appl. Math., 30 (1976), pp. 597–607.

[72] S. M. Robinson, *Stability theorems for systems of inequalities, Part* II: *Differentiable nonlinear systems*, SIAM J. Numer. Anal., 13 (1976), pp. 497–513.

[73] S. M. Robinson, *Regularity and stability for convex multivalued functions*, Math. Oper. Res., 1 (1976), pp. 130–143.

[74] S. M. Robinson, *Strongly regular generalized equations*, Math. Oper. Res., 5 (1980), pp. 43–62.

[75] S. M. Robinson, *Generalized equations and their solutions, Part* II: *Applications to nonlinear programming*, Math. Programming Stud., 19 (1982), pp. 200–221.

[76] S. M. Robinson, *Implicit B-Differentiability in Generalized Equations*, Technical Report 2854, Mathematics Research Center, University of Wisconsin-Madison, 1985.

[77] R. T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton, NJ, 1970.

[78] R. T. Rockafellar, *Conjugate Duality and Optimization*, Regional Conf. Ser. Appl. Math., SIAM, Philadelphia, PA, 1974.
[79] A. Shapiro, *Second-order sensitivity analysis and asymptotic theory of parametrized, nonlinear programs*, Math. Programming, 33 (1985), pp. 280–299.
[80] A. Shapiro, *Sensitivity analysis of nonlinear programs and differentiability properties of metric projections*, SIAM J. Control Optim., 26 (1988), pp. 628–645.
[81] A. Shapiro, *Perturbation theory of nonlinear programs when the set of optimal solutions is not a singleton*, Appl. Math. Optim., 18 (1988), pp. 215–229.
[82] A. Shapiro, *On concepts of directional differentiability*, J. Optim. Theory Appl., 66 (1990), pp. 477–487.
[83] A. Shapiro, *Perturbation analysis of optimization problems in Banach spaces*, Numer. Funct. Anal. Optim., 13 (1992), pp. 97–116.
[84] A. Shapiro, *On Lipschitzian stability of optimal solutions of parametrized semi-infinite programs*, Math. Oper. Res., 19 (1994), pp. 743–752.
[85] A. Shapiro, *Directional differentiability of the optimal value function in convex semi-infinite programming*, Math. Programming Ser. A, 70 (1995), pp. 149–157.
[86] A. Shapiro, *On uniqueness of Lagrange multipliers in optimization problems subject to cone constraints*, SIAM J. Optim., 7 (1997), pp. 508–518.
[87] A. Shapiro, *A variational principle and its applications*, in Proceedings of the Conference on Parametric Optimization and Related Topics IV, Approximation and Optimization Series, J. Guddat, H. Th. Jongen, F. Nozicka, G. Still, and F. Twilt, eds., Verlag Peter Lang, Frankfurt, 1996, pp. 341–357.
[88] A. Shapiro, *First and second order analysis of nonlinear semidefinite programs*, Math. Programming Ser. B, 77 (1997), pp. 301–320.
[89] A. Shapiro and J. F. Bonnans, *Sensitivity analysis of parametrized programs under cone constraints*, SIAM J. Control Optim., 30 (1992), pp. 1409–1422.
[90] D. Torralba, *Développements asymptotiques pour les méthodes d'approximation par viscosité*, Comptes Rendus Acad. Sci. Paris Sér. I, 322 (1996), pp. 123–128.
[91] C. Ursescu, *Multifunctions with convex closed graph*, Czech. Math. J., 25 (1975), pp. 438–441.
[92] K. Veselić, *Finite catenary and the method of Lagrange*, SIAM Rev., 30 (1995), pp. 224–229.
[93] D. E. Ward, *Characterizations of strict local minima and necessary conditions for weak sharp minima*, J. Optim. Theory Appl., 80 (1994), pp. 551–571.
[94] V. Zeidan, *The Riccati equation for optimal control problems with mixed state-control constraints: Necessity and sufficiency*, SIAM J. Control Optim., 32 (1994), pp. 1297–1321.
[95] J. Zowe and S. Kurcyusz, *Regularity and stability for the mathematical programming problem in Banach spaces*, Appl. Math. Optim., 5 (1979), pp. 49–62.