The Radius of Metric Regularity[1]

A. L. Dontchev
Mathematical Reviews
American Mathematical Society
Ann Arbor, Michigan 48107-8604,
ald@ams.org

A. S. Lewis
Department of Combinatorics and Optimization
University of Waterloo, Waterloo, Ontario, Canada N2L 3G1
aslewis@orion.uwaterloo.ca
http://orion.uwaterloo.ca/∼aslewis

R. T. Rockafellar
Department of Mathematics
University of Washington, Seattle, WA 98195
rtr@math.washington.edu

August 4, 2000

Key words: metric regularity, perturbations, distance to irregularity, distance to infeasibility, Eckart-Young theorem, Lusternik-Graves theorem, Robinson-Ursescu theorem, coderivatives.

AMS 2000 Subject Classification: Primary: 49J53. Secondary: 49J52, 90C31.

Abstract

Metric regularity is a central concept in variational analysis for the study of solution mappings associated with "generalized equations," including variational inequalities and parameterized constraint systems. Here it is employed to characterize the distance to irregularity or infeasibility with respect to perturbations of the system structure. Generalizations of the Eckart-Young theorem in numerical analysis are obtained in particular.

[1]Research partially supported by the NSF Grant DMS–9803098 for the first and the third author, and by the Natural Sciences and Engineering Research Council of Canada for the second author.

1 Introduction

Let X and Y be real Banach spaces, with norms both denoted by $\|\cdot\|$ and closed unit balls B_X and B_Y. Let F be a mapping from X to Y, by which we will generally mean a *set-valued* mapping, indicated by $F : X \rightrightarrows Y$, having inverse $F^{-1} : Y \rightrightarrows X$ with $x \in F^{-1}(y) \;\Leftrightarrow\; y \in F(x)$, and having effective graph, domain and range sets given respectively by

$$\operatorname{gph} F = \{\,(x,y)\,|\, y \in F(x)\}, \qquad \operatorname{dom} F = \{\,x \,|\, F(x) \neq \emptyset\}, \qquad \operatorname{rge} F = \operatorname{dom} F^{-1}.$$

Single-valuedness of F on a subset of X is a special case; when the subset is all of X (implying $\operatorname{dom} F = X$), we say F is a single-valued mapping from X to Y and write $F : X \to Y$.

This terminology is suited for the study of "generalized equations" of the form $F(x) \ni y$ and their solutions x for fixed y as parameter. When F is single-valued, such a relation reduces to a true equation, but more broadly it can express a mixture of inequality and equality conditions on x, interpreted as a constraint system. On the other hand, the relation $F(x) \ni y$ can stand for a variational inequality or a system of optimality conditions. Either way, the set of solutions is $F^{-1}(y)$ and is nonempty if and only if $y \in \operatorname{rge} F$. A central issue is how $F^{-1}(y)$ behaves with respect to perturbations in y as a parameter, or for that matter, perturbations in F itself.

It is easy to recognize in this picture two of the abiding themes in numerical work: bounds on what happens to solutions under perturbations and estimates of how large a perturbation can be before good behavior of a solution mapping breaks down. Our goal is to build on recent work on the first of these topics, using advanced tools in variational analysis and convexity to reach a new level of understanding of the second topic. Far-reaching extensions of the Eckart-Young theorem on matrix perturbations will be obtained along with insights into the distance to infeasibility in the perturbation of constraint systems.

In our framework where neither F nor F^{-1} needs to be single-valued, we generally have to focus on a particular $\bar{y} \in Y$ and solution $\bar{x} \in F^{-1}(\bar{y})$, being content with local analysis around the pair $(\bar{x},\bar{y}) \in \operatorname{gph} F$. A key property of "good behavior" is metric regularity. The mapping F is said to be *metrically regular* at $\bar{x}$ for $\bar{y}$ when there exists $\kappa \in (0,\infty)$ such that

$$d(x, F^{-1}(y)) \;\leq\; \kappa d(y, F(x)) \qquad \text{for all } (x,y) \text{ close to } (\bar{x},\bar{y}). \tag{1.1}$$

Here, d denotes distance: thus $d(y,E) = \inf_{e\in E} \|y - e\|$ for any set $E \subset Y$. (By convention, infima over the empty set are $+\infty$.) This property gives an estimate of how far a candidate x can be from the solution set corresponding to y; the distance is bounded above by a multiple κ of the distance between y and $F(x)$, which is a sort of "residual" that may be easy to compute. The *regularity modulus* of F at $\bar{x}$ for $\bar{y}$ is the value

$$\operatorname{reg} F(\bar{x}\,|\,\bar{y}) \;:=\; \inf\{\,\kappa \in (0,\infty) \,|\, (1.1) \text{ holds}\} \;\in [0,\infty]. \tag{1.2}$$

The case $\operatorname{reg} F(\bar{x}\,|\,\bar{y}) = \infty$ corresponds to the absence of any $\kappa \in (0,\infty)$ satisfying (1.1); thus, in this notation, metric regularity of F at $\bar{x}$ for $\bar{y}$ is signaled by $\operatorname{reg} F(\bar{x}\,|\,\bar{y}) < \infty$.

Our main result, formulated below, concerns the extent to which F can be perturbed before metric regularity is lost. A measure of this, called the *radius* of metric regularity, will be introduced and characterized in terms of $\operatorname{reg} F(\bar{x}\,|\,\bar{y})$. To appreciate what is thereby achieved, it is important to understand the deep connections between metric regularity and other kinds of "good behavior" of solution mappings, and to have some sense of the machinery that is already available for determining $\operatorname{reg} F(\bar{x}\,|\,\bar{y})$.

Metric regularity of F is tied to a fundamental Lipschitz-type property of F^{-1}. Following [23], we say F^{-1} has the *Aubin property* at $\bar{y}$ for $\bar{x}$ when there exists $\kappa \in (0,\infty)$ together with a neighborhood O of $\bar{x}$ such that

$$(1.3) \qquad F^{-1}(y') \cap O \subset F^{-1}(y) + \kappa \|y' - y\| B_Y \quad \text{for all } y, y' \text{ close to } \bar{y}.$$

If F^{-1} is single-valued around $\bar{y}$, this is ordinary Lipschitz continuity with constant κ, whereas more generally, if F^{-1} is locally bounded with closed graph, it is such continuity with respect to the Pompeiu-Hausdorff metric—but (1.3) operates usefully without any boundedness requirement on the sets $F^{-1}(y)$. A celebrated fact is that F^{-1} has the Aubin property at $\bar{y}$ for $\bar{x}$ if and only if F is metrically regular at $\bar{x}$ for $\bar{y}$, and moreover the constants κ agree, so that the regularity modulus $\operatorname{reg} F(\bar{x}, \bar{y})$ also has the formula

$$(1.4) \qquad \operatorname{reg} F(\bar{x}\,|\,\bar{y}) \;=\; \inf\{\, \kappa \in (0,\infty) \,|\, (1.3) \text{ holds} \,\}.$$

A full discussion of this relationship and its history, leading even to formulations in which X and Y are merely metric spaces, has recently been provided by Ioffe [12]; details in the finite-dimensional case, with many references beyond, are also available in [23]. In the notation of [23], the right side of (1.4) is $\operatorname{lip} F^{-1}(\bar{y}\,|\,\bar{x})$ and is called the Lipschitz modulus for F^{-1} at $\bar{y}$ for $\bar{x}$.

Another way of looking at the regularity modulus $\operatorname{reg} F(\bar{x}\,|\,\bar{y})$ is through the property of F being *linearly open*, or *locally surjective* at $\bar{x}$ for $\bar{y}$, which refers to the existence of $\kappa \in (0,\infty)$ and a neighborhood O of $\bar{y}$ such that

$$(1.5) \qquad F(x + \operatorname{int} \kappa r B_X) \supset [F(x) + \operatorname{int} r B_Y] \cap O \quad \text{for all } x \text{ close to } \bar{x} \text{ and } r > 0.$$

This too is equivalent to metric regularity of F at $\bar{x}$ for $\bar{y}$ with the same range of values for κ and thus yields for the regularity modulus a third formula,

$$(1.6) \qquad \operatorname{reg} F(\bar{x}\,|\,\bar{y}) \;=\; \inf\{\, \kappa \in (0,\infty) \,|\, (1.5) \text{ holds} \,\}.$$

Again, details can be found in [12] and [23] and the works they cite.

The meaning of metric regularity is much simpler when F is a single-valued mapping that is linear. Let $L(X,Y)$ denote the space of continuous linear mappings $X \to Y$, and for $F \in L(X,Y)$ let $\|F\|$ be the usual operator norm.

Example 1.1 (metric regularity of linear mappings)**.** *For $F \in L(X,Y)$, $\operatorname{reg} F(\bar{x}\,|\,\bar{y})$ is the same for all $(\bar{x},\bar{y}) \in \operatorname{gph} F$, and that common value, denoted simply by $\operatorname{reg} F$, is given by*

$$(1.7) \qquad \operatorname{reg} F \;=\; \inf\{\, \kappa \in (0,\infty) \,|\, \kappa F(B_X) \supset \operatorname{int} B_Y \} \;=\; \sup\{\, d(0, F^{-1}(y)) \,|\, y \in B_Y \}.$$

Moreover $\operatorname{reg} F < \infty$ if and only if F is surjective. If in that case F^{-1} is single-valued, as must be true when $\dim X = \dim Y < \infty$, then

$$(1.8) \qquad \operatorname{reg} F = \|F^{-1}\|.$$

Thus, metric regularity is equivalent to nonsingularity when $\dim X = \dim Y < \infty$.

Detail. In (1.5), the inclusion can be reduced to $\kappa F(B_X) \supset \operatorname{int} B_Y$ by subtracting $F(x)$ from both sides and then dividing by r. This yields the first equation in (1.7). On the other hand, we can write the inequality for metric regularity in (1.1) as $d(0, F^{-1}(y) - x) \leq \kappa d(y - F(x), 0)$ and, by noting that linearity implies $F^{-1}(y) - x = F^{-1}(y - y')$ for $y' = F(x)$, reformulate it as $d(0, F^{-1}(y - y')) \leq \kappa \|y - y'\|$. To say that this holds for all y sufficiently close to $\bar{y}$ and y' such that $y' = F(x)$ for an x sufficiently close to $\bar{x}$ is to say that $d(0, F^{-1}(z)) \leq \kappa \|z\|$ for all z near 0, and hence then for every $z \in Y$, again by linearity. The regularity modulus thus is also given by the second expression in (1.7). The assertion about surjectivity of F follows from (1.7) and the Banach open mapping principle (cf. [24]); the rest is then immediate. □

For single-valued mappings F that are nonlinear but differentiable, the notion of metric regularity, if not the term itself, goes back to a basic theorem in analysis which is associated with the work of Lusternik [15] and Graves [10]. Here we denote by $DF(\bar{x})$ the derivative mapping in $L(X, Y)$ that is associated with F at $\bar{x}$.

Theorem 1.2 (Lusternik-Graves)**.** *For continuously Fréchet differentiable* $F : X \to Y$ *and any* $(\bar{x}, \bar{y}) \in \operatorname{gph} F$*, one has*

$$\operatorname{reg} F(\bar{x} \,|\, \bar{y}) = \operatorname{reg} DF(\bar{x}), \tag{1.9}$$

Thus, F *is metrically regular at* $\bar{x}$ *for* $\bar{y} = F(\bar{x})$ *if and only if* $DF(\bar{x})$ *is surjective.*

Statements of this theorem have usually revolved around the final assertion, but the fact that $\operatorname{reg} F(\bar{x} \,|\, \bar{y}) = \operatorname{reg} DF(\bar{x})$ can be gleaned from the proof of Graves and the observations in Example 1.1, as transferred to the linear mapping $DF(\bar{x})$. For more newer work on the Lusternik-Graves theorem, see [4], [2], [7], [3], [5], [12].

These facts set the stage for studying perturbations and whether they can cause "irregular behavior." The starting point is the classical Eckart-Young theorem in numerical analysis, which is usually presented in terms of $n \times n$ matrices (cf. [11]) but, for the sake of enhancing comparisons, is formulated here in terms of linear mappings. The issue is the extent to which a nonsingular linear mapping F from $I\!R^n$ to $I\!R^n$ can be perturbed by the addition of a linear mapping G without destroying the nonsingularity.

Theorem 1.3 (Eckart-Young)**.** *For nonsingular* $F \in L(X, Y)$ *when* $\dim X = \dim Y < \infty$,

$$\min_{G \in L(X,Y)} \Big\{ \|G\| \,\Big|\, F + G \text{ singular} \Big\} = 1/\|F^{-1}\|. \tag{1.10}$$

Through Example 1.1, we can identify nonsingularity in this result with metric regularity, and $1/\|F^{-1}\|$ with $1/\operatorname{reg} F$. Thus, the largest radius r such that $F + G$ is nonsingular for all $G \in L(X, Y)$ satisfying $\|G\| < r$ is $r = 1/\operatorname{reg} F$. At the same time, $r = 1/\operatorname{reg} F$ is the largest radius under which metric regularity is preserved.

Our aim is to expand this radius idea beyond linear mappings $F : X \to Y$ to general mappings $F : X \rightrightarrows Y$ in order to measure, with respect to a pair $(\bar{x}, \bar{y})$ where metric regularity holds, how far F can be perturbed before metric regularity may be lost. Perturbation of F to $F + G$ shifts $(\bar{x}, \bar{y})$ to $(\bar{x}, \bar{y} + G(\bar{x}))$, so the question is how big G can be before $F + G$ fails to be metrically regular at $\bar{x}$ for $\bar{y}+G(\bar{x})$. Metric regularity of F at $\bar{x}$ for $\bar{y}$ can be thought of as meaning that the generalized equation $F(x) \ni \bar{y}$ is well behaved at its solution $\bar{x}$, whereas the absence of metric regularity of $F+G$ at $\bar{x}$ for $\bar{y}+G(\bar{x})$ can be interpreted as saying that the generalized equation $(F+G)(x) \ni \bar{y}+G(\bar{x})$ is *irregular* at its solution $\bar{x}$.

Definition 1.4 (radius of metric regularity). *For any mapping $F : X \rightrightarrows Y$ and $(\bar{x},\bar{y}) \in \operatorname{gph} F$, the radius of metric regularity at $\bar{x}$ for $\bar{y}$ is the value*

$$\text{(1.11)} \qquad \operatorname{rad} F(\bar{x}\,|\,\bar{y}) \;:=\; \inf_{G\in L(X,Y)} \Big\{ \|G\| \;\Big|\; F+G \textit{ not metrically regular at } \bar{x} \textit{ for } \bar{y}+G(\bar{x}) \Big\}.$$

The value $\operatorname{rad} F(\bar{x}\,|\,\bar{y})$ could equally well be called the *distance to irregularity* in the sense suggested before the definition, with respect to adding a linear mapping to F. Obviously in the Eckart-Young setting of Theorem 1.3, $\operatorname{rad} F(\bar{x}\,|\,\bar{y})$ is the same for all $(\bar{x},\bar{y}) \in \operatorname{gph} F$ and this common value equals $1/\|F^{-1}\|$, which is $1/\operatorname{reg} F$ by Example 1.1. We aim at capturing the analogous relationship locally between $\operatorname{rad} F(\bar{x},\bar{y})$ and $\operatorname{reg} F(\bar{x},\bar{y})$ in much wider circumstances.

Of course, for general $F : X \rightrightarrows Y$ there is no reason to restrict perturbations to the addition of a mapping $G : X \to Y$ that belongs to $L(X,Y)$. An attractive feature of the radius defined by (1.11) is that it turns out to serve also for a vastly larger class of perturbation mappings $G : X \to Y$, at least in finite dimensions. This result is part of our main theorem, stated next. The theorem refers to the Lipschitz modulus of a single-valued mapping G at a point $\bar{x}$, which is

$$\text{(1.12)} \qquad \operatorname{lip} G(\bar{x}) := \limsup_{\substack{x,x'\to\bar{x} \\ x,x'\neq\bar{x}}} \frac{\|G(x')-G(x)\|}{\|x'-x\|}.$$

Obviously $\operatorname{lip} G(\bar{x}) < \infty$ if and only if G is Lipschitz continuous on some neighborhood of $\bar{x}$. When G is continuously differentiable, $\operatorname{lip} G(\bar{x}) = \|DG(\bar{x})\|$. In the context of the broader notion of the Lipschitz modulus of a set-valued mapping in [23], $\operatorname{lip} G(\bar{x}) = \operatorname{lip} G(\bar{x}\,|\,\bar{y}) = \operatorname{reg} G^{-1}(\bar{y}\,|\,\bar{x})$ with respect to $\bar{y} = G(\bar{x})$.

We also use in the theorem statement the concept of *local closedness* of a set at a point, meaning that some neighborhood of the point has closed intersection with the set.

Theorem 1.5 (characterization of the radius of metric regularity). *For a mapping $F : X \rightrightarrows Y$ and any $(\bar{x},\bar{y}) \in \operatorname{gph} F$ at which $\operatorname{gph} F$ is locally closed, $\operatorname{rad} F(\bar{x}\,|\,\bar{y}) \geq 1/\operatorname{reg} F(\bar{x}\,|\,\bar{y})$. In fact*

$$\text{(1.13)} \qquad \operatorname{rad} F(\bar{x}\,|\,\bar{y}) \;=\; 1/\operatorname{reg} F(\bar{x}\,|\,\bar{y}) \qquad \textit{when } \dim X < \infty,\ \dim Y < \infty.$$

Furthermore, in this case the infimum in the definition of $\operatorname{rad} F(\bar{x}\,|\,\bar{y})$ is unchanged if taken with respect to $G \in L(X,Y)$ of rank 1, but also is unchanged when the space of perturbations G is enlarged from linear mappings to locally Lipschitz continuous mappings:

$$\text{(1.14)} \qquad \operatorname{rad} F(\bar{x}\,|\,\bar{y}) \;=\; \min_{G:X\to Y} \Big\{ \operatorname{lip} G(\bar{x}) \;\Big|\; F+G \textit{ not metrically regular at } \bar{x} \textit{ for } \bar{y}+G(\bar{x}) \Big\}.$$

Thus, although defined for convenience in terms of linear perturbations, $\operatorname{rad} F(\bar{x}\,|\,\bar{y})$ is also, when X and Y are finite-dimensional, the largest radius r such that for any $G : X \to Y$ that is Lipschitz continuous around $\bar{x}$ with constant less than r, the mapping $F+G$ is sure to be metrically regular at $\bar{x}$ for $\bar{y}+G(\bar{x})$. In proving Theorem 1.5, in Section 3, we actually show in more detail that for X and Y of arbitrary dimension one always has

$$(1.15)\qquad \min_{G:X\to Y}\Big\{\ \operatorname{lip} G(\bar{x})\ \Big|\ F+G \text{ not metrically regular at } \bar{x} \text{ for } \bar{y}+G(\bar{x})\Big\} \geq 1/\operatorname{reg} F(\bar{x}\,|\,\bar{y}),$$

(see Corollary 3.4), while on the other hand, for finite-dimensional X and Y one can obtain from the calculus of coderivatives that

$$(1.16)\qquad \inf_{\substack{G\in L(X,Y)\\ G \text{ of rank } 1}}\Big\{\ \|G\|\ \Big|\ F+G \text{ not metrically regular at } \bar{x} \text{ for } \bar{y}+G(\bar{x})\Big\} \leq 1/\operatorname{reg} F(\bar{x}\,|\,\bar{y}),$$

These inequalities, and the fact that $\operatorname{lip} G(\bar{x}) = \|G\|$ when G is linear, furnish the result.

Note that it is not necessary to assume in Theorem 1.5 that F itself is metrically regular at $\bar{x}$ for $\bar{y}$, since if this were not true we would have $\operatorname{reg} F(\bar{x}\,|\,\bar{y}) = \infty$ and the formula would be correct under the convention that $1/\infty = 0$. Likewise, in the case of $\operatorname{reg} F(\bar{x}\,|\,\bar{y}) = 0$ the formula should be interpreted as giving $\operatorname{rad} F(\bar{x}\,|\,\bar{y}) = \infty$.

Groundwork will be laid in Section 2 by studying mappings that are positively homogeneous. Graphical derivative and coderivative mappings are a prime target among positively homogeneous mappings, for use in Section 3 in establishing the inequality in (1.16), but the results in Section 2 also are of interest in themselves. They include a different kind of generalization of the Eckart-Young theorem based on an extended definition of "nonsingularity" (see Theorem 2.6). They also establish that for sublinear mappings with closed graph the equation in Theorem 1.5 holds at the origin even when X and Y are not finite-dimensional (see Theorem 2.9).

In Section 4, we apply our results to convex constraint systems, i.e., to the solution set $F^{-1}(\bar{y})$ when $\operatorname{gph} F$ is convex. The question then is how far both F and $\bar{y}$ can be perturbed without encountering infeasibility in the sense of the solution set becoming empty. Working with systems of certain types where F is positively homogeneous, Renegar [18] introduced a notion of the *distance to infeasibility*. A bridge between that notion and metric regularity is provided by the following result of Robinson [19], [20], and Ursescu [25], which generalizes the Banach open mapping principle in a direction different from that taken by the Lusternik-Graves theorem. (This statement omits, as unneeded for present purposes, some accompanying information about the growth of the metric regularity constant κ as $(\bar{x},\bar{y})$ is varied.)

Theorem 1.6 (Robinson-Ursescu). *For a mapping $F: X \rightrightarrows Y$ and $(\bar{x},\bar{y}) \in \operatorname{gph} F$, if F has closed convex graph, then F is metrically regular at $\bar{x}$ for $\bar{y}$ if and only if $\bar{y} \in \operatorname{int} \operatorname{rge} F$.*

According to Theorem 1.6, metric regularity corresponds to "strict feasibility" of a convex constraint system—where $\bar{y}$, at least, can be perturbed by some amount before feasibility is lost. Making use of this, we prove in Section 4 that Renegar's distance to infeasibility is in fact the radius of metric regularity for a mapping $\bar{F}$ derived from F and $\bar{y}$ by a process of "homogenization." The theory of infeasibility is thereby reconfigured as a branch of the theory of metric regularity and opened up to the many resources in that wider theory, while at the same time it is extended from homogeneous to possibly inhomogeneous systems.

Section 5 specializes the "generalized equation" $F(x) \ni \bar{y}$ to a variational inequality and applies our radius theory to that important context, bringing out a connection with a property of *strong* metric regularity in which the inverse mapping F^{-1} has a single-valued localization.

2 Homogeneous Mappings and Nonsingularity

A mapping $F : X \rightrightarrows Y$ is *positively homogeneous* when $0 \in F(0)$ and $F(\lambda x) \supset \lambda F(x)$ for $\lambda > 0$, or equivalently, when $\operatorname{gph} F$ is a cone in $X \times Y$. It is *sublinear* when, in addition, $F(x + x') \supset F(x) + F(x')$, or equivalently, when $\operatorname{gph} F$ is a *convex* cone in $X \times Y$. Sublinear mappings are also called *convex processes* [22] and have been the subject of extensive development. Linear mappings are sublinear mappings in particular. Graphical derivative or coderivative mappings that are defined by way of tangent or normal cones to the graphs of other mappings are always positively homogeneous, but may or may not be sublinear, depending on the circumstances. Obviously, F^{-1} is positively homogeneous if and only if F is positively homogeneous, and likewise with sublinearity.

For a positively homogeneous mapping $F : X \rightrightarrows Y$, the *outer norm* and the *inner norm* are, respectively,

$$(2.1) \qquad \|F\|^+ = \sup_{x \in B_X} \sup_{y \in F(x)} \|y\| \quad \text{and} \quad \|F\|^- = \sup_{x \in B_X} \inf_{y \in F(x)} \|y\|,$$

either of which might be ∞. When F is single-valued on $\operatorname{dom} F$, its outer and inner norms agree and can be written as $\|F\|$; this reduces to the operator norm of F when $F \in L(X, Y)$. Of course, these expressions are not true "norms" apart from the single-valued case, because general set-valued mappings do not form a vector space.

Inner and outer norms can be applied to F^{-1} as well as to F, and that will dominate our use of them. Indeed, the expression on the right of (1.7), characterizing $\operatorname{reg} F$ for a linear mapping F, can be identified with $\|F^{-1}\|^-$. An analogous characterization is available for sublinear mappings F, where however the focus is on $\operatorname{reg} F(0\,|\,0)$ instead of a regularity modulus that is the same at all points of the graph of F.

Example 2.1 (metric regularity of sublinear mappings)**.** *For a sublinear mapping $F : X \rightrightarrows Y$ and any $(\bar{x}, \bar{y}) \in \operatorname{gph} F$,*

$$(2.2) \qquad \operatorname{reg} F(\bar{x}\,|\,\bar{y}) \;\le\; \operatorname{reg} F(0\,|\,0) \;=\; \|F^{-1}\|^-.$$

Thus, F is metrically regular everywhere when it is metrically regular at 0 for 0. In fact

$$(2.3) \qquad \operatorname{reg} F(0\,|\,0) \;=\; \inf\Big\{\, \kappa \in (0, \infty) \,\Big|\, F(x + \kappa r B_X) \supset F(x) + r B_Y \text{ for all } x \in X,\ r > 0 \Big\}.$$

Moreover $\operatorname{reg} F(0\,|\,0) < \infty$ if and only if F is surjective.

Detail. Recall from the introduction that $\operatorname{reg} F(\bar{x}\,|\,\bar{y})$ can be described as the infimum of all $\kappa \in (0, \infty)$ for which (1.5) holds relative to some neighborhood O of $\bar{y}$. This property of κ holding for $(\bar{x}, \bar{y}) = (0, 0)$ clearly implies $F(\kappa B_X) \supset \operatorname{int} B_Y$; intersection with a neighborhood O is superfluous because of positive homogeneity. On the other hand, just from knowing that $F(\kappa B_X) \supset \operatorname{int} B_Y$ for a certain $\kappa > 0$, we obtain for arbitrary $(x, y) \in \operatorname{gph} F$ and $\varepsilon > 0$ through the sublinearity of F that $F(x + \kappa r B_X) \supset F(x) + r F(\kappa B_X) \supset y + \operatorname{int} r B_Y$. This establishes that $\operatorname{reg} F(\bar{x}\,|\,\bar{y}) \le \operatorname{reg} F(0\,|\,0)$ for all $(\bar{x}, \bar{y}) \in \operatorname{gph} F$ and, appealing again to positive homogeneity, that

$$(2.4) \qquad \operatorname{reg} F(0\,|\,0) = \inf\{\, \kappa \in (0, \infty) \,|\, F(\kappa B_X) \supset B_Y \}.$$

By definition, $\|F^{-1}\|^- = \inf\{\,\kappa \in (0,\infty) \,|\, y \in B_Y \;\Rightarrow\; F^{-1}(y) \cap \kappa B_X \neq \emptyset\}$. Hence $\|F^{-1}\|^-$ equals the right side of (2.4) as well. The fact that the finiteness of the value in (2.4) corresponds to F being surjective is an immediate consequence of Robinson-Ursescu result in Theorem 1.6, but it was actually proved earlier by Robinson in [19]. □

Next, as a basis for generalizing the Eckart-Young theorem, we use outer norms of inverses to extend the definition of nonsingularity from linear mappings $X \to Y$ with $\dim X = \dim Y < \infty$ to general positively homogeneous mappings $X \rightrightarrows Y$ without any dimensionality restrictions. Later, when we return to metric regularity, such nonsingularity will be important as a dual property applied to adjoint mappings or coderivative mappings.

Definition 2.2 (extended nonsingularity). *A positively homogeneous mapping $F : X \rightrightarrows Y$ will be called singular if $\|F^{-1}\|^+ = \infty$, nonsingular if $\|F^{-1}\|^+ < \infty$.*

For linear $F : X \to Y$, nonsingularity in this sense coincides with the traditional notion when $\dim X = \dim Y < \infty$ but means in general that F^{-1} is single-valued and continuous relative to $\operatorname{rge} F$, its domain. For positively homogeneous mappings that are not linear, nonsingularity implies $F^{-1}(0) = \{0\}$ but does not necessitate single-valuedness elsewhere.

Example 2.3 (extended nonsingularity in finite dimensions). *When $\dim X < \infty$, $\dim Y < \infty$, a positively homogeneous $F : X \rightrightarrows Y$ with closed graph is nonsingular if and only if $F^{-1}(0) = \{0\}$.*

Example 2.4 (extended nonsingularity of sublinear mappings). *A mapping $F : X \rightrightarrows Y$ that is sublinear is both nonsingular and surjective if and only if $F = L^{-1}$ for a continuous linear mapping $L : Y \to X$.*

Detail. The assumptions imply that F^{-1} is a sublinear mapping L with $\operatorname{dom} L = Y$ and $L(0)$ containing only 0. Then for any $y \in Y$ we have both $L(y)$ and $L(-y)$ nonempty with $L(0) = L(y-y) \supset L(y) + L(-y)$, so that L must be single-valued everywhere with $L(-y) = L(y)$. In combination with positive homogeneity this yields $L(\lambda y) = \lambda L(y)$ for all real λ. Single-valuedness in the sublinearity rule $L(y+y') \supset L(y) + L(y')$ requires $L(y+y') = L(y) + L(y')$, so L must be linear. Furthermore $\|L\| = \|F^{-1}\|^+ < \infty$. □

The following characterization of $\|F^{-1}\|^+$ will help to elucidate the meaning of extended nonsingularity more generally.

Proposition 2.5 (inverse norm formula). *For a positively homogeneous mapping $F : X \rightrightarrows Y$,*

$$\|F^{-1}\|^+ \;=\; \inf\{\,\kappa \in (0,\infty) \,|\, x \in F^{-1}(y) \;\Rightarrow\; \|x\| \le \kappa\|y\|\,\} \;=\; \sup_{\|x\|=1} \frac{1}{d(0, F(x))}. \tag{2.5}$$

Proof. By the definition of the outer norm in (2.1) as applied to F^{-1}, we know that $\|F^{-1}\|^+$ is the supremum of $\|x\|$ over all pairs $(y,x) \in \operatorname{gph} F^{-1}$ with $\|y\| \le 1$. Thus, it is the infimum of all $\kappa \in (0,\infty)$ such that $F^{-1}(B_Y) \subset \kappa B_X$. That translates through positive homogeneity to the middle expression in (2.5). There we observe that the infimum is unchanged when x is restricted to have $\|x\| = 1$, in which case the expression can be identified with the infimum of all $\kappa \in (0,\infty)$ such that $\kappa \ge 1/\|y\|$ whenever $y \in F(x)$ and $\|x\| = 1$. (It is correct in this to interpret $1/\|y\| = \infty$ if $y = 0$.). This shows that the middle expression in (2.5) agrees with the final expression in (2.5). □

Theorem 2.6 (extended Eckart-Young). *For any $F : X \rightrightarrows Y$ that is positively homogeneous,*

$$\inf_{G \in L(X,Y)} \Big\{ \|G\| \,\Big|\, F + G \text{ singular} \Big\} \;=\; 1/\|F^{-1}\|^{+}. \tag{2.6}$$

Moreover the infimum is the same if restricted to mappings $G \in L(X,Y)$ of rank one. (Indeed, if $X = Z^*$ for a Banach space Z, the additional restriction can be made that G is weak-to-norm continuous.)*

Proof. If F is singular, (2.6) holds with 0 on both sides, so we can assume that $\|F^{-1}\|^{+} < \infty$. We can also assume that $\|F^{-1}\|^{+} > 0$, because having $\|F^{-1}\|^{+} = 0$ corresponds to having $\operatorname{dom} F = \{0\}$, and that implies for any G that $\operatorname{dom}(F+G) = \{0\}$, hence $\|(F+G)^{-1}\|^{+} = 0$. In that case (2.6) holds with ∞ on both sides.

Suppose $G \in L(X,Y)$ and $F + G$ is singular. Then by definition there is a sequence of elements $(x_k, y_k) \in \operatorname{gph}(F+G)$ with $\|y_k\| \le 1$ and $0 < \|x_k\| \to \infty$. From $y_k \in (F+G)(x_k)$ we have $x_k \in F^{-1}(y_k - G(x_k))$, hence $\|x_k\| \le \|F^{-1}\|^{+}\|y_k - G(x_k)\|$ and consequently

$$1/\|F^{-1}\|^{+} \le (\|y_k\| + \|G(x_k)\|)/\|x_k\| \le (1/\|x_k\|) + \|G\|.$$

Taking the limit as $k \to \infty$, we get $1/\|F^{-1}\|^{+} \le \|G\|$. Thus "$\ge$" holds in (2.6).

We have to show now that "$\le$" holds as well. Consider any finite $r > 1/\|F^{-1}\|^{+}$. We have $1/r < \|F^{-1}\|^{+}$, so there must exist $(\hat{x}, \hat{y}) \in \operatorname{gph} F$ with $\|\hat{y}\| = 1$ and $\|\hat{x}\| > 1/r$. In the dual space X^* it is possible to find $\hat{x}^*$ with $\langle \hat{x}, \hat{x}^* \rangle = \|\hat{x}\|$ and $\|\hat{x}^*\| = 1$. Define the rank-one mapping $\hat{G} \in L(X,Y)$ by $\hat{G}(x) = -\|\hat{x}\|^{-1}\langle x, \hat{x}^* \rangle \hat{y}$. Then $\hat{G}(\hat{x}) = -\hat{y}$, so that $(F + \hat{G})(\hat{x}) = F(\hat{x}) + \hat{G}(\hat{x}) = F(\hat{x}) - \hat{y} \ni 0$. Therefore $\hat{x} \in (F + \hat{G})^{-1}(0)$, and $F + \hat{G}$ must be singular. On the other hand, $\|\hat{G}\| = \|\hat{y}\|/\|\hat{x}\| = 1/\|\hat{x}\| < r$. Hence the infimum in (2.6) is less than r. Appealing to the choice of r, we confirm that the infimum in (2.6) cannot be more than $1/\|F^{-1}\|^{+}$.

When $X = Z^*$, the latter argument can be refined by taking $\hat{x}^*$ to be an element in the unit ball B_Z that satisfies $\langle \hat{x}, \hat{x}^* \rangle > 1 - \delta$ for small $\delta > 0$, and the proof goes much as before. □

The infimum in Theorem 2.6 would likewise be unaffected if taken over all positively homogeneous $G : X \to Y$, even set-valued $G : X \rightrightarrows Y$ as long as $\|G\|$ is replaced by $\|G\|^{+}$. Only the first part of the proof requires adjustment, and all that is needed is to replace $G(\bar{x})$ by a general element $\bar{z} \in G(\bar{x})$.

Outer and inner norms are elegantly related by duality under a generalized notion of adjoint mappings, and this leads to further insights into metric regularity. With respect to the spaces X^* and Y^* dual to X and Y, the *upper adjoint* of a positively homogeneous mapping $F : X \rightrightarrows Y$ is the mapping $F^{*+} : Y^* \rightrightarrows X^*$ defined by

$$(y^*, x^*) \in \operatorname{gph} F^{*+} \quad \Leftrightarrow \quad \langle x^*, x \rangle \le \langle y^*, y \rangle \text{ for all } (x,y) \in \operatorname{gph} F, \tag{2.7}$$

whereas the *lower adjoint* is the mapping $F^{*-} : Y^* \rightrightarrows X^*$ defined by

$$(y^*, x^*) \in \operatorname{gph} F^{*-} \quad \Leftrightarrow \quad \langle x^*, x \rangle \ge \langle y^*, y \rangle \text{ for all } (x,y) \in \operatorname{gph} F.$$

The graphs of both F^{*+} and F^{*-} correspond to the closed convex cone in $X^* \times Y^*$ that is polar to $\operatorname{gph} F$, except for permuting (x^*, y^*) to (y^*, x^*) and introducing certain changes of sign. In particular, both F^{*+} and F^{*-} are always sublinear with closed graph; furthermore, $\operatorname{gph} F^{*-} = -\operatorname{gph} F^{*+}$. The interesting fact is that when F itself is sublinear with closed graph, one has (see [1] or [23, 11.29]):

$$(F^{*+})^{*-} = (F^{*-})^{*+} = F, \tag{2.8}$$

and moreover

$$\|F\|^{+} = \|F^{*+}\|^{-} = \|F^{*-}\|^{-} \quad \text{and} \quad \|F\|^{-} = \|F^{*+}\|^{+} = \|F^{*-}\|^{+}. \tag{2.9}$$

Proposition 2.7 (adjoint formula for metric regularity)**.** *For a sublinear mapping $F : X \rightrightarrows Y$ with closed graph, the characterizations of $\operatorname{reg} F(0\,|\,0)$ in Example 2.1 are supplemented by*

$$\operatorname{reg} F(0\,|\,0) = \|(F^{*+})^{-1}\|^{+}. \tag{2.10}$$

Proof. We know from Example 2.1 that $\operatorname{reg} F(0\,|\,0) = \|F^{-1}\|^{-}$. However, $\|F^{-1}\|^{-} = \|(F^{-1})^{*-}\|^{+}$ by (2.9) (as applied to F^{-1} instead of F). Also, $(F^{-1})^{*-} = (F^{*+})^{-1}$ by the adjoint formulas. This leads to (2.10). ◻

The relations between adjoints and their norms also induce a duality between the properties of surjectivity and nonsingularity.

Proposition 2.8 (Borwein [1])**.** *For sublinear $F : X \rightrightarrows Y$ with closed graph, F is surjective if and only if F^{*+} is nonsingular. Likewise, F is nonsingular if and only if F^{*+} is surjective.*

Proof. This is immediate from Example 2.1, Definition 2.2 and the relations in (2.9). ◻

Theorem 2.9 (radius of metric regularity for sublinear mappings)**.** *For a sublinear mapping $F : X \rightrightarrows Y$ with closed graph,*

$$\operatorname{rad} F(0\,|\,0) \;=\; 1/\operatorname{reg} F(0\,|\,0) \;=\; \inf_{G \in L(X,Y)} \Big\{ \|G\| \;\Big|\; F + G \textit{ not surjective} \Big\}. \tag{2.11}$$

Moreover the infimum is the same if restricted to G of rank one.

Proof. For any $G \in L(X, Y)$, the mapping $F + G$ is sublinear with closed graph and has $(F + G)^{*+} = F^{*+} + G^{*}$. Thus by Proposition 2.8, $F + G$ is surjective if and only if $F^{*+} + G^{*}$ is nonsingular. It follows that

$$\inf_{G \in L(X,Y)} \Big\{ \|G\| \;\Big|\; F + G \text{ not surjective} \Big\} \;=\; \inf_{G \in L(X,Y)} \Big\{ \|G^{*}\| \;\Big|\; F^{*+} + G^{*} \text{ singular} \Big\}. \tag{2.12}$$

The right side of (2.12) can be identified through Theorem 2.6 with

$$\inf_{H \in L(Y^{*},X^{*})} \Big\{ \|H\| \;\Big|\; F^{*+} + H \text{ singular} \Big\} = 1/\|(F^{*+})^{-1}\|^{+} \tag{2.13}$$

by the observation that any $H \in L(Y^{*}, X^{*})$ of rank one that is weak*-to-norm continuous has the form G^{*} for some $G \in L(X, Y)$ of rank one. In consequence, the left side of (2.12) equals the right side of (2.13), and from (2.10) we then get the equality between the second and third expressions in (2.11). The third expression in (2.11) equals the first by Definition 1.4 and the assertion at the end of Example 2.1. $\square$

Note that since $\operatorname{reg} F(0\,|\,0) = \|F^{-1}\|^{-}$ (cf. Example 2.1), the second part of (2.12) can be written equivalently as

$$\inf_{G\in L(X,Y)} \left\{ \|G\| \,\middle|\, F+G \text{ not surjective} \right\} = 1/\|F^{-1}\|^{-}.$$

This relation was recently established by Lewis [13, Thm. 4.8], but here it has been deduced from broader facts about generalized nonsingularity, as well as placed in a context of metric regularity.

In line with the observation after Theorem 2.6, the infimum in (2.11) would be unaffected if taken over sublinear $G : X \rightrightarrows Y$ such that $\operatorname{dom} G = X$ and $F + G$ has closed graph, as long as $\|G\|$ is replaced by $\|G\|^{+}$. The proof still works with G^* replaced by G^{*+}.

Remark. Once we have established Theorem 1.5 it will be possible to regard Theorem 2.9 as a special case of that result (by way of Theorem 1.6) when X and Y are *finite-dimensional*, and then the additional information in Theorem 1.5 about nonlinear perturbations could be brought in. Theorem 1.5 does not cover Theorem 2.9 for infinite-dimensional X or Y, however.

Corollary 2.10 (adjoint formula for the radius)**.** *For a sublinear mapping $F : X \rightrightarrows Y$ with closed graph,*

$$\operatorname{rad} F(0\,|\,0) = \inf_{\|y^*\|=1} d(0, F^{*+}(y^*)). \tag{2.14}$$

Proof. The left side of (2.14) is $1/\operatorname{reg} F(0\,|\,0)$ by Theorem 2.9, but the right side is $1/\|(F^{*+})^{-1}\|^{+}$ by Proposition 2.5. The two sides are therefore equal by the observation in Proposition 2.7. $\square$

3 Coderivatives and the Radius Characterization

Graphical differentiation of set-valued mappings provides a powerful tool for dealing with metric regularity. Our use of this tool will be limited to a finite-dimensional context, for reasons to be explained, and we therefore tailor the description of graphical differentiation to such spaces as well. Details beyond what we say here can be found in [23].

Let $\dim X < \infty$ and $\operatorname{dom} Y < \infty$, and consider a mapping $F : X \rightrightarrows Y$. The *graphical derivative* of F at $\bar{x}$ for an element $\bar{y} \in F(\bar{x})$ is the mapping $DF(\bar{x}\,|\,\bar{y})$ having as its graph the tangent cone to $\operatorname{gph} F$ at $(\bar{x}, \bar{y})$:

$$\operatorname{gph} DF(\bar{x}\,|\,\bar{y}) = \limsup_{\lambda \searrow 0} \frac{1}{\lambda}[\operatorname{gph} F - (\bar{x}, \bar{y})].$$

Thus, $DF(\bar{x}\,|\,\bar{y})$ is a positively homogeneous mapping $X \rightrightarrows Y$ with closed graph. On the basis of the definitions reviewed in the preceding section, it has an upper adjoint $DF(\bar{x}\,|\,\bar{y})^{*+} : Y^* \rightrightarrows X^*$ and also a lower adjoint. The *coderivative* of F at $\bar{x}$ for $\bar{y}$ is defined by taking limits of the graphs of the upper adjoints at nearby points (x, y):

$$\operatorname{gph} D^*F(\bar{x}\,|\,\bar{y}) = \limsup_{(x,y)\to(\bar{x},\bar{y})} \operatorname{gph} DF(x\,|\,y)^{*+},$$

where $(x, y) \in \operatorname{gph} F$. Therefore, $D^*F(\bar{x}\,|\,\bar{y})$ is a positively homogeneous mapping $Y^* \rightrightarrows X^*$ with closed graph.

The value of coderivative mappings comes from the extensive calculus that is available for determining them from the structure of given mapping F (cf. Chap. 10 of [23]) and from the following result of Mordukhovich; for the somewhat complicated story of this result, in which Ioffe also had a role, see p. 418 of [23].

Theorem 3.1 (Mordukhovich criterion)**.** *For $F : X \rightrightarrows Y$ and any $(\bar{x}, \bar{y}) \in \operatorname{gph} F$ at which $\operatorname{gph} F$ is locally closed, if $\dim X < \infty$ and $\dim Y < \infty$, then*

$$\operatorname{reg} F(\bar{x}\,|\,\bar{y}) \;=\; \|D^*F(\bar{x}\,|\,\bar{y})^{-1}\|^+. \tag{3.1}$$

This can be translated through our Definition 2.2 into an equivalence between metric regularity and coderivative nonsingularity.

Corollary 3.2 (coderivative nonsingularity)**.** *Under the theorem's assumptions, F is metrically regular at $\bar{x}$ for $\bar{y}$ if and only if the coderivative mapping $D^*F(\bar{x}\,|\,\bar{y})$ is nonsingular.*

Coderivative mappings can be defined in a similar manner when X and Y are infinite-dimensional, but the question of exactly what kind of set limit to take is more subtle and can depend on the "smoothness" of these spaces. Anyway, formula (3.1) can fail in the more general setting. See Mordukhovich [16] for more on this matter, including certain substitutes for (3.1) in terms of limits. The validity of (3.1) itself will be crucial to our argument for establishing equation (1.3) of Theorem 1.5 and is the reason why that equation is asserted only for finite-dimensional spaces.

Another ingredient of our proof of Theorem 1.5 will be the following estimate. This estimate can be deduced in various ways from the literature on the Lusternik-Graves theorem; see for example Theorem 1.4 of Dontchev [4] or Section 1 of the more recent paper by Ioffe [12]. We express it here in the notation of $\operatorname{reg} F$ and $\operatorname{lip} G$ and, for completeness, supply a direct proof which is in line with the original arguments of Lusternik and Graves.

Theorem 3.3 (estimate for Lipschitz perturbations)**.** *Consider a mapping $F : X \rightrightarrows Y$ and any $(\bar{x}, \bar{y}) \in \operatorname{gph} F$ at which $\operatorname{gph} F$ is locally closed. Consider also a mapping $G : X \to Y$. If $\operatorname{reg} F(\bar{x}\,|\,\bar{y}) < \kappa < \infty$ and $\operatorname{lip} G(\bar{x}) < \lambda < \kappa^{-1}$, then*

$$\operatorname{reg}(F + G)(\bar{x}\,|\,\bar{y} + G(\bar{x})) < (\kappa^{-1} - \lambda)^{-1} = \frac{\kappa}{1 - \lambda\kappa}. \tag{3.2}$$

Proof. With the notation $B_a(\bar{x}) = \bar{x} + aB_X$ and $B_a(\bar{y}) = \bar{y} + aB_Y$, let $a > 0$ be small enough that $\operatorname{gph} F$ is closed relative to $B_a(\bar{x}) \times B_a(\bar{y})$, G is Lipschitz with constant λ on $B_a(\bar{x})$, and

$$d(x, F^{-1}(y)) \;\le\; \kappa d(y, F(x)) \quad \text{ for all } \quad (x, y) \in B_a(\bar{x}) \times B_a(\bar{y}).$$

In particular, this implies that $d(\bar{x}, F^{-1}(y)) \le \kappa d(y, F(\bar{x})) \le \kappa\|y - \bar{y}\|$ when $y \in B_a(\bar{y})$ and therefore

$$F^{-1}(y) \neq \emptyset \quad \text{ for all } y \in B_a(\bar{y}). \tag{3.3}$$

Choose α such that

$$0 < \alpha < \frac{1}{4}a(1-\kappa\lambda)\min\{1,\kappa\}\,. \tag{3.4}$$

Let $x \in B_{\alpha/4}(\bar{x})$ and $y \in B_{\alpha/(4\kappa)}(\bar{y})$. Then, from the choice of α in (3.4), we have

$$\|y - G(x) + G(\bar{x}) - \bar{y}\| \;\le\; \lambda\|x-\bar{x}\| + \|y-\bar{y}\| \;\le\; (\lambda\alpha/4) + (\alpha/4\kappa) \le a. \tag{3.5}$$

Fix ε such that

$$0 < \varepsilon < \frac{1}{4}a(1-\kappa\lambda)\min\{1,1/\lambda\}\,. \tag{3.6}$$

Then, through (3.3) and (3.5), there exists $z_1 \in F^{-1}(y - G(x) + G(\bar{x}))$ such that

$$\|z_1 - x\| \le d(x, F^{-1}(y - G(x) + G(\bar{x}))) + \varepsilon. \tag{3.7}$$

We obtain from the metric regularity that

$$\begin{aligned}
\|z_1 - x\| \;&\le\; \|x-\bar{x}\| + d(\bar{x}, F^{-1}(y - G(x) + G(\bar{x}))) + \varepsilon \\
&\le\; \|x-\bar{x}\| + \kappa d(y - G(x) + G(\bar{x}), F(\bar{x})) + \varepsilon \\
&\le\; \|x-\bar{x}\| + \kappa\|y-\bar{y}\| + \kappa\lambda\|x-\bar{x}\| + \varepsilon \\
&\le\; (\alpha/4) + (\kappa\alpha/4\kappa) + (\kappa\lambda\alpha/4) + \varepsilon \;\le\; (3\alpha/4) + \varepsilon.
\end{aligned}$$

Then we have $\|z_1 - \bar{x}\| \;\le\; \|z_1 - x\| + \|x - \bar{x}\| \;\le\; (3\alpha/4) + \varepsilon + (\alpha/4) \;\le\; \alpha + \varepsilon \;\le\; a$. By induction now, we construct an infinite sequence of vectors z_j, $j = 1, 2, \ldots$, such that

$$z_{j+1} \in F^{-1}(y - G(z_j) + G(\bar{x})) \quad \text{and} \quad \|z_{j+1} - z_j\| \le (\kappa\lambda)^j\|z_1 - x\|.$$

Suppose that we have generated such $z_2, \ldots, z_n$ from z_1. Then for $j = 1, 2, \ldots, n-1$ we have

$$\begin{aligned}
\|z_j - \bar{x}\| \;&\le\; \sum_{i=1}^{j-1}\|z_{i+1} - z_i\| + \|z_1 - \bar{x}\| \le \sum_{i=0}^{j-1}(\kappa\lambda)^i\|z_1 - x\| + \|z_1 - \bar{x}\| \\
&\le\; \frac{1}{1-\kappa\lambda}\|z_1 - x\| + \|z_1 - \bar{x}\| \le \frac{1}{1-\kappa\lambda}(3\alpha/4 + \varepsilon) + \alpha + \varepsilon \;\le\; a
\end{aligned}$$

according to the choice of the constants α in (3.4) and ε in (3.6). Also,

$$\|y - G(z_j) + G(\bar{x}) - \bar{y}\| \le \|y - \bar{y}\| + \lambda\|z_j - \bar{x}\| \le \frac{\alpha}{4\kappa} + \frac{\lambda}{1-\kappa\lambda}(3\alpha/4 + \varepsilon) + \lambda(\alpha + \varepsilon) \le a.$$

The metric regularity of F yields the existence of $z_{n+1} \in F^{-1}(y - G(z_n) + G(\bar{x}))$ such that $\|z_{n+1} - z_n\| \le \kappa d(y - G(z_n) + G(\bar{x}), F(z_n))$. Since $z_n \in F^{-1}(y - G(z_{n-1}) + G(\bar{x}))$, this implies $\|z_{n+1} - z_n\| \le \kappa\|G(z_n) - G(z_{n-1})\| \le \kappa\lambda\|z_n - z_{n-1}\|$, and the induction step is complete.

The sequence z_n satisfies the Cauchy condition, hence is convergent to some z, which, from the local closedness of $\operatorname{gph} F$, satisfies $z \in F^{-1}(y - G(z) + G(\bar{x}))$, that is, $z \in (F+G)^{-1}(y + G(\bar{x}))$. Moreover,

$$\begin{aligned}
d(x, (F+G)^{-1}(y + G(\bar{x})) \;&\le\; \|z - x\| \le \lim_{n\to\infty}\sum_{i=1}^{n}\|z_{i+1} - z_i\| + \|z_1 - x\| \\
\le \lim_{n\to\infty}\sum_{i=0}^{n}(\kappa\lambda)^i\|z_1 - x\| \;&\le\; \frac{1}{1-\kappa\lambda}\|z_1 - x\| \le \frac{\kappa}{1-\kappa\lambda}\Big[d(y + G(\bar{x}), (F+G)(x)) + \varepsilon\Big],
\end{aligned}$$

the final inequality being obtained from (3.7). Since the left side does not depend on ε, which can be arbitrary small, this tells us that $F+G$ is metrically regular at $\bar{x}$ for $\bar{y}+G(\bar{x})$. $\square$

Corollary 3.4 (general perturbation inequality)**.** *If the mapping $F: X \rightrightarrows Y$ is locally closed at $(\bar{x},\bar{y}) \in \operatorname{gph} F$, then*

$$(3.8) \qquad \inf_{G:X\to Y}\Big\{ \operatorname{lip} G(\bar{x}) \,\Big|\, F+G \textit{ not metrically regular at } \bar{x} \textit{ for } \bar{y}+G(\bar{x}) \Big\} \;\geq\; 1/\operatorname{reg} F(\bar{x}\,|\,\bar{y}).$$

Proof. If $\operatorname{lip} G(\bar{x}) < 1/\operatorname{reg} F(\bar{x}\,|\,\bar{y})$, there exist $\lambda > \operatorname{lip} G(\bar{x})$ and $\kappa > \operatorname{reg} F(\bar{x}\,|\,\bar{y})$ such that $\lambda < 1/\kappa$, and then $F+G$ must be metrically regular at $\bar{x}$ for $\bar{y}+G(\bar{x})$ by the theorem. $\square$

Proof of Theorem 1.5. It is obvious from the definition of $\operatorname{rad} F(\bar{x}\,|\,\bar{y})$ in (1.11) that

$$(3.9) \qquad \operatorname{rad} F(\bar{x}\,|\,\bar{y}) \;\geq\; \inf_{G:X\to Y}\Big\{ \operatorname{lip} G(\bar{x}) \,\Big|\, F+G \text{ not metrically regular at } \bar{x} \text{ for } \bar{y}+G(\bar{x}) \Big\}.$$

The double inequality obtained by combining (3.8) and (3.9) holds without any restriction on the dimensions of X and Y. It reduces our task to demonstrating that $\operatorname{rad} F(\bar{x}\,|\,\bar{y}) = 1/\operatorname{reg} F(\bar{x}\,|\,\bar{y})$ when $\operatorname{dom} X < \infty$ and $\operatorname{dom} Y < \infty$, along with verifying that the infimum in the definition of $\operatorname{rad} F(\bar{x}\,|\,\bar{y})$ also is unchanged when restricted to mappings $G \in L(X,Y)$ of rank one.

In this finite-dimensional setting we have Theorem 3.1 and Corollary 3.2 at our disposal. With these results we can restate the targeted equation as

$$(3.10) \qquad \inf_{G\in L(X,Y)}\Big\{ \|G\| \,\Big|\, D^*(F+G)(\bar{x}\,|\,\bar{y}+G(\bar{x})) \text{ singular} \Big\} \;=\; 1/\|D^*F(\bar{x}\,|\,\bar{y})^{-1}\|^{+}.$$

It is elementary from the calculus of coderivatives (cf. 10.43 of [23]) that $D^*(F+G)(\bar{x}\,|\,\bar{y}+G(\bar{x})) = D^*F(\bar{x}\,|\,\bar{y}) + G^*$. Also, $\|G^*\| = \|G\| < \infty$. Hence (3.10) can be identified with

$$(3.11) \qquad \inf_{G\in L(X,Y)}\Big\{ \|G^*\| \,\Big|\, D^*F(\bar{x}\,|\,\bar{y}) + G^* \text{ singular} \Big\} \;=\; 1/\|D^*F(\bar{x}\,|\,\bar{y})^{-1}\|^{+}.$$

Furthermore, again by finite dimensionality, every $H \in L(Y^*,X^*)$ is the adjoint G^* of some $G \in L(X,Y)$. Therefore (3.11) is valid as the special case of Theorem 2.6 as applied to the mapping $D^*F(\bar{x}\,|\,\bar{y}) : Y^* \rightrightarrows X^*$, and on the basis of that theorem the infimum is unchanged if taken over $G \in L(X,Y)$ of rank one. $\square$

Theorem 1.5 also gives information on what happens to the radius of metric regularity under perturbations.

Corollary 3.5 (perturbed radius of metric regularity)**.** *For $F: X \rightrightarrows Y$ with $\dim X < \infty$ and $\dim Y < \infty$, and any $(\bar{x},\bar{y}) \in \operatorname{gph} F$ at which $\operatorname{gph} F$ is locally closed, one has*

$$(3.12) \qquad \operatorname{rad}(F+G)(\bar{x}\,|\,\bar{y}+G(\bar{x})) \;\geq\; \operatorname{rad} F(\bar{x}\,|\,\bar{y}) - \operatorname{lip} G(\bar{x}) \qquad \textit{when } \operatorname{lip} G(\bar{x}) < \operatorname{rad} F(\bar{x}\,|\,\bar{y}).$$

Proof. For such F and G, suppose $H: X \to Y$ has $\operatorname{lip} H(\bar{x}) < \operatorname{rad} F(\bar{x}\,|\,\bar{y}) - \operatorname{lip} G(\bar{x})$. Is $(F+G)+H$ sure to be metrically regular at $\bar{x}$ for $(\bar{y}+G(\bar{x}))+H(\bar{x})$? Yes, because in terms of $G' = G+H$ this is the same as $F+G'$ being metrically regular at $\bar{x}$ for $\bar{y}+G'(\bar{x})$, and that is true by (1.14) of Theorem 1.5 since $\operatorname{lip} G'(\bar{x}) \leq \operatorname{lip} G(\bar{x}) + \operatorname{lip} H(\bar{x}) < \operatorname{rad} F(\bar{x},\bar{y})$. $\square$

A further conclusion can be drawn. Recall that a mapping $F': X \rightrightarrows Y$ is said to give a *first-order approximation* to a mapping $F: X \rightrightarrows Y$ around $(\bar{x},\bar{y}) \in \operatorname{gph} F$ if on some neighborhood O of $\bar{x}$ there is a mapping $G: O \to Y$ with $G(\bar{x}) = 0$, $\operatorname{lip} G(\bar{x}) = 0$, and $F' = F+G$.

Corollary 3.6 (radius stability under first-order approximation). *When* $\dim X < \infty$ *and* $\dim Y < \infty$*, if* $F : X \rightrightarrows Y$ *has graph locally closed at* $(\bar{x}, \bar{y}) \in \operatorname{gph} F$*, and* $\Phi : X \rightrightarrows Y$ *is any mapping that furnishes a first-order approximation to* F *around* $(\bar{x}, \bar{y})$*, then*

$$\operatorname{rad} F(\bar{x}\,|\,\bar{y}) \;=\; \operatorname{rad} \Phi(\bar{x}\,|\,\bar{y}). \tag{3.13}$$

Proof. Consider $G : O \to Y$ as in the preceding definition of first-order approximation, and extend G in any way to a mapping $X \to Y$. Since Φ agrees with $F + G$ around $\bar{x}$, and $G(\bar{x}) = 0$, we have $\operatorname{rad} \Phi(\bar{x}\,|\,\bar{y}) = \operatorname{rad}(F + G)(\bar{x}\,|\,\bar{y})$. On the other hand, since $\operatorname{lip} G(\bar{x}) = 0$ we have $\operatorname{rad}(F + G)(\bar{x}\,|\,\bar{y}) \geq \operatorname{rad} F(\bar{x}\,|\,\bar{y})$ by Corollary 3.5. Therefore $\operatorname{rad} \Phi(\bar{x}\,|\,\bar{y}) \geq \operatorname{rad} F(\bar{x}\,|\,\bar{y})$. To say that Φ gives a first-order approximation of F is to say equally, however, that F gives a first-order approximation of Φ; the relationship is symmetric, with $-G$ replacing G. Hence the inequality $\operatorname{rad} F(\bar{x}\,|\,\bar{y}) \geq \operatorname{rad} \Phi(\bar{x}\,|\,\bar{y})$ must hold as well. ◻

An example of a first-order approximation to which Corollary 3.6 can be applied is seen when $F(x) = F_0(x) + f(x)$ for a mapping $f : X \to Y$ that is strictly differentiable at $\bar{x}$, and $\Phi(x) = F_0(x) + g(x)$ for $g(x) = f(\bar{x}) + \langle \nabla f(\bar{x}), x - \bar{x} \rangle$. In this case $\Phi = F + G$ for a mapping G that is strictly differentiable at $\bar{x}$ with $G(\bar{x}) = 0$ and Jacobian $\nabla G(\bar{x}) = 0$ and thus has $\operatorname{lip} G(\bar{x}) = 0$.

4 Application to Constraint Systems

In this section, fully back in the context of possibly infinite-dimensional spaces X and Y, we interpret the "generalized equation" $F(x) \ni \bar{y}$ as a constraint system on x, focusing on the case where F has convex graph. Such a system is called *feasible* if $F^{-1}(\bar{y}) \neq \emptyset$, i.e., $\bar{y} \in \operatorname{rge} F$, and *strictly feasible* if actually $\bar{y} \in \operatorname{int} \operatorname{rge} F$.

Recall that any closed, convex cone $K \subset Y$ induces a partial ordering "$\leq_K$" under the rule that $y_0 \leq_K y_1$ means $y_1 - y_0 \in K$. Correspondingly, $y_0 <_K y_1$ means $y_1 - y_0 \in \operatorname{int} K$.

Example 4.1 (convex constraint systems). *Let* $C \subset X$ *be a closed convex set, let* $K \subset Y$ *be a closed convex cone, and let* $A : C \to Y$ *be a continuous mapping that is convex with respect to the partial ordering in* Y *induced by* K*:*

$$A((1-\theta)x_0 + \theta x_1) \leq_K (1-\theta)A(x_0) + \theta A(x_1) \quad \textit{for } x_0, x_1 \in C \textit{ when } 0 < \theta < 1.$$

Define the mapping $F : X \rightrightarrows Y$ *by*

$$F(x) = \begin{cases} A(x) + K & \textit{if } x \in C, \\ \emptyset & \textit{if } x \notin C. \end{cases}$$

Then F *has closed, convex graph, and feasibility of the system* $F(x) \ni \bar{y}$ *refers to*

$$\exists\, \bar{x} \in C \textit{ such that } A(\bar{x}) \leq_K \bar{y},$$

while, as long as $\operatorname{int} K \neq \emptyset$*, strict feasibility refers to*

$$\exists\, \bar{x} \in C \textit{ such that } A(\bar{x}) <_K \bar{y}.$$

Detail. Only the final claim needs a comment. When $\operatorname{int} K \neq \emptyset$, we have $K = \operatorname{cl} \operatorname{int} K$, so that the convex set $\operatorname{rge} F = A(C) + K$ is the closure of the open set $O := A(C) + \operatorname{int} K$. Also, O is convex. If follows then that $\operatorname{int} \operatorname{rge} F = O$. ◻

Example 4.2 (linear-conic constraint systems). *Add to Example 4.1 the assumption that A is linear and C is a cone, so that the condition $\bar{x} \in C$ can be written equivalently as $\bar{x} \geq_C 0$. Then F is sublinear. Furthermore, its adjoint $F^{*+} : Y^* \rightrightarrows X^*$ is given in terms of the adjoint A^* of A and the dual cones $K^+ = -K^*$ and $C^+ = -C^*$ (where * denotes polar) by*

$$F^{*+}(y^*) = \begin{cases} A^*(y^*) - C^+ & \text{if } y \in K^+, \\ \emptyset & \text{if } y \notin K^+, \end{cases}$$

so that $F^{*+}(y^*) \ni x^*$ if and only if $y^* \geq_{K^+} 0$ and $A^*(y^*) \geq_{C^+} x^*$.

In general for the system $F(x) \ni \bar{y}$, we will be interested in perturbations in which F is replaced by $F + G$ and $\bar{y}$ by $\bar{y} + g$ with $G \in L(X, Y)$ and $g \in Y$. Such a perturbation, the magnitude of which is quantified with

$$\|(G, g)\| = \max\{\, \|G\|, \|g\| \,\}, \tag{4.1}$$

transforms the condition $F(x) \ni \bar{y}$ to $(F + G)(x) \ni \bar{y} + g$ and the solution set $F^{-1}(\bar{y})$ to $(F + G)^{-1}(\bar{y} + g)$, creating infeasibility if $(F + G)^{-1}(\bar{y} + g) = \emptyset$, i.e., if $\bar{y} + g \notin \operatorname{rge}(F + G)$.

Such perturbations in the linear-conic context of Example 4.2 were studied by Renegar [18]. For that case he introduced a measure of how large (G, g) can be before the system $(F+G)(x) \ni \bar{y} + g$ becomes infeasible, and a sizable literature has since grown around this; see e.g. [9], [8], [17]. We extend Renegar's definition to general convex systems as follows.

Definition 4.3 (distance to infeasibility). *For $F : X \rightrightarrows Y$ with convex graph and $\bar{y} \in \operatorname{rge} F$, the **distance to infeasibility** of the system $F(x) \ni \bar{y}$ is the value*

$$\inf_{G \in L(X,Y),\, g \in Y} \Big\{\, \|(G, g)\| \,\Big|\, \bar{y} + g \notin \operatorname{rge}(F + G) \Big\}. \tag{4.2}$$

It will be essential to our approach in analyzing this value that there would be no difference if feasibility were replaced by strict feasibility in the definition.

Lemma 4.4 (infeasibility versus strict infeasibility). *The distance to infeasibility in the sense of Definition 4.3 is the same as the distance to strict infeasibility, namely the value*

$$\inf_{G \in L(X,Y),\, g \in Y} \Big\{\, \|(G, g)\| \,\Big|\, \bar{y} + g \notin \operatorname{int} \operatorname{rge}(F + G) \Big\}. \tag{4.3}$$

Proof. Let S_1 denote the set of (G, g) over which the infimum is taken in (4.2) and let S_2 be the corresponding set in (4.3). Obviously $S_1 \subset S_2$, so the first infimum cannot be less than the second. We must show that it also cannot be greater. This amounts to demonstrating that for any $(G, g) \in S_2$ and any $\varepsilon > 0$ we can find $(G', g') \in S_1$ such that $\|(G', g')\| \leq \|(G, g)\| + \varepsilon$. In fact, we can get this with $G' = G$ simply by noting that when $\bar{y} + g \notin \operatorname{int} \operatorname{rge}(F + G)$ there must exist $g' \in Y$ with $\bar{y} + g' \notin \operatorname{rge}(F + G)$ and $\|g' - g\| \leq \varepsilon$. $\square$

Corollary 4.5 (reduction to metric regularity). *The distance to infeasibility of the system $F(x) \ni \bar{y}$, as in Definition 4.3, is also the value*

$$\inf_{(G,g) \in L(X,Y) \times Y} \Big\{\, \|(g, G)\| \,\Big|\, F + G \text{ not metrically regular at any } \bar{x} \text{ for } \bar{y} + g \Big\}. \tag{4.4}$$

Proof. This is now immediate from Theorem 1.6 (Robinson-Ursescu) and Definition 1.4 of the radius of metric regularity. □

Although Corollary 4.5 expresses the distance to infeasibility in terms of metric regularity, it does not tie in with our theory of the radius of metric regularity. For that, we have to pass from F to a special mapping $\bar{F}$ constructed as a "homogenization" of the convex system $F(x) \ni \bar{y}$. We will then be able to apply to $\bar{F}$ the result on distance to metric regularity in Theorem 2.9.

Following [23], we denote by F^∞ the *horizon mapping* associated with F, the graph of F^∞ in $X \times Y$ being the horizon cone of the graph of F. When F has closed, convex graph, this is the same as $\operatorname{gph} F^\infty$ being the recession cone of $\operatorname{gph} F$ in the sense of convex analysis:

$$(x', y') \in \operatorname{gph} F^\infty \quad \Leftrightarrow \quad \operatorname{gph} F + (x', y') \subset \operatorname{gph} F.$$

Definition 4.6 (homogenization). *For $F : X \rightrightarrows Y$ and $\bar{y} \in \operatorname{rge} F$, the homogenization of the constraint system $F(x) \ni \bar{y}$ is the system $\bar{F}(x,t) \ni 0$, where $\bar{F} : X \times I\!R \rightrightarrows Y$ is defined by*

$$\bar{F}(x,t) = \begin{cases} tF(t^{-1}x) - t\bar{y} & \text{if } t > 0, \\ F^\infty(x) & \text{if } t = 0, \\ \emptyset & \text{if } t < 0, \end{cases} \tag{4.5}$$

and the solution sets to the two systems are related by

$$x \in F^{-1}(\bar{y}) \quad \Leftrightarrow \quad (x,1) \in \bar{F}^{-1}(0). \tag{4.6}$$

Note that if F is positively homogeneous with closed graph, then $tF(t^{-1}x) = F(x) = F^\infty(x)$ for all $t > 0$, so that we simply have $\bar{F}(x,t) = F(x) - t\bar{y}$ for $t \geq 0$ but $\bar{F}(x,t) = \emptyset$ for $t < 0$.

In what follows, we adopt the norm

$$\|(x,t)\| = \|x\| + |t| \quad \text{for } (x,t) \in X \times I\!R. \tag{4.7}$$

The case of the next theorem in which the mapping F is sublinear reduces to the main result of Lewis[14, Thm. 4.5], and further specialization to Example 4.2 then captures Renegar's result [18, Thm. 1.3].

Theorem 4.7 (radius characterization of the distance to infeasibility). *Let $F : X \rightrightarrows Y$ have closed, convex graph and let $\bar{y} \in \operatorname{rge} F$. Then in the homogenized system $\bar{F}(t,x) \ni 0$ the mapping $\bar{F}$ is sublinear with closed graph, and*

$$\bar{y} \in \operatorname{int} \operatorname{rge} F \quad \Leftrightarrow \quad 0 \in \operatorname{int} \operatorname{rge} \bar{F} \quad \Leftrightarrow \quad \bar{F} \text{ is surjective.} \tag{4.8}$$

Furthermore, for the given constraint system $F(x) \ni \bar{y}$ one has

$$\textit{distance to infeasibility} \;=\; \operatorname{rad} \bar{F}(0,0|0) \;=\; 1/\operatorname{reg} \bar{F}(0,0|0). \tag{4.9}$$

Proof. The definition of $\bar{F}$ corresponds to $\operatorname{gph} \bar{F}$ being the closed convex cone in $X \times I\!R \times Y$ that is generated by $\{\, (x, 1, y - \bar{y}) \,|\, (x,y) \in \operatorname{gph} F \}$. Hence $\bar{F}$ is sublinear, and also, $\operatorname{rge} \bar{F}$ is a convex cone. We have $(\operatorname{rge} F) - \bar{y} = F(X) - \bar{y} = \bar{F}(X,1)$, so it is obvious that if $\bar{y} \in \operatorname{int} \operatorname{rge} F$, then $0 \in \operatorname{int} \operatorname{rge} \bar{F}$. Since $\operatorname{rge} \bar{F}$ is a convex cone the latter is equivalent to having $\operatorname{rge} \bar{F} = Y$, i.e., surjectivity.

Conversely now, suppose $\bar{F}$ is surjective; Theorem 1.6 (Robinson-Ursescu) informs us that in this case $0 \in \operatorname{int} \bar{F}(W)$ for every neighborhood W of the origin in $I\!R \times X$. It must be verified, however, that $\bar{y} \in \operatorname{int} \operatorname{rge} F$. In terms of $C(t) = \bar{F}(B_X, t) \subset Y$, it will suffice to show that $0 \in \operatorname{int} C(t)$ for some $t > 0$. Note that the sublinearity of $\bar{F}$ implies that

$$C((1-\theta)t_0 + \theta t_1) \supset (1-\theta)C(t_0) + \theta C(t_1) \quad \text{for } 0 < \theta < 1. \tag{4.10}$$

Our assumption that $\bar{y} \in \operatorname{rge} F$ ensures that $F^{-1}(\bar{y}) \neq \emptyset$. Choose $\tau \in (0, \infty)$ small enough that $1/2\tau > d(0, F^{-1}(\bar{y}))$. Then

$$0 \in C(t) \text{ for all } t \in [0, 2\tau], \tag{4.11}$$

whereas, because $[-2\tau, 2\tau] \times B_X$ is a neighborhood W of the origin in $I\!R \times X$, we have

$$0 \in \operatorname{int} \bar{F}(B_X, [-2\tau, 2\tau]) = \operatorname{int} \bigcup_{0 \le t \le 2\tau} C(t). \tag{4.12}$$

We will use this to show that actually $0 \in \operatorname{int} C(\tau)$. For $y^* \in Y^*$ define

$$\sigma(y^*, t) := \sup_{y \in C(t)} \langle y, y^* \rangle, \qquad \lambda(t) := \inf_{\|y^*\|=1} \sigma(y^*, t).$$

The property in (4.10) makes $\sigma(y^*, t)$ be concave in t, and the same then follows for $\lambda(t)$. As long as $0 \le t \le 2\tau$ we have $\sigma(y^*, t) \ge 0$ and $\lambda(t) \ge 0$ by (4.11). On the other hand, the union in (4.12) includes some ball around the origin, so

$$\exists\, \varepsilon > 0 \text{ such that } \sup_{0 \le t \le 2\tau} \sigma(y^*, t) \ge \varepsilon \text{ for all } y^* \in Y^* \text{ with } \|y^*\| = 1. \tag{4.13}$$

We argue next that $\lambda(\tau) > 0$. If not, then since λ is a nonnegative concave function on $[0, 2\tau]$ we would have to have $\lambda(t) = 0$ for all $t \in [0, 2\tau]$. Supposing that to be the case, choose $\delta \in (0, \varepsilon/2)$ and, in the light of the definition of $\lambda(\tau)$, an element $\hat{y}^*$ with $\sigma(\hat{y}^*, \tau) < \delta$. The nonnegativity and concavity of $\sigma(\hat{y}^*, \cdot)$ on $[0, 2\tau]$ implies then that $\sigma(\hat{y}^*, t) \le (\delta/\tau)t$ when $\tau \le t \le 2\tau$ and $\sigma(\hat{y}^*, t) \le 2\delta - (\delta/\tau)t$ when $0 \le t \le \tau$. But that gives us $\sigma(\hat{y}^*, t) \le 2\delta < \varepsilon$ for all $t \in [0, 2\tau]$ in contradiction to the property of ε in (4.13). Therefore, $\lambda(\tau) > 0$ as claimed.

We have $\sigma(y^*, \tau) \ge \lambda(\tau)$ when $\|y^*\| = 1$, hence by positive homogeneity that $\sigma(y^*, \tau) \ge \lambda(\tau)\|y^*\|$ for all $y^* \in Y^*$. In this inequality, $\sigma(\cdot, \tau)$ is the support function of the convex set $C(\tau)$, or equivalently of $\operatorname{cl} C(\tau)$, whereas $\lambda(\tau)\|\cdot\|$ is the support function of $\lambda(\tau)B_Y$. It follows therefore that $\operatorname{cl} C(\tau) \supset \lambda(\tau)B_Y$, so that at least $0 \in \operatorname{int} \operatorname{cl} C(\tau)$.

We invoke next a result of Robinson [20, Lemma 1.1]: If D is a closed convex subset of $X \times Y$ with projections D_X and D_Y on X and Y, and if D_X is bounded, then $\operatorname{int} \operatorname{cl} D_Y = \operatorname{int} D_Y$. In taking $D = \{\, (x, t, y) \in \operatorname{gph} \bar{F} \mid x \in B_X,\ t = \tau \}$ we have $D_X \subset B_X \times \{\tau\}$ and $D_Y = C(\tau)$, so this allows us to pass from the knowledge that $0 \in \operatorname{int} \operatorname{cl} C(\tau)$ to having $0 \in \operatorname{int} C(\tau)$, as desired.

We have established (4.8) and turn now to (4.9). The first thing to observe is that every $\bar{G} \in L(X \times I\!R, Y)$ can be identified with a pair $(G, g) \in L(X, Y) \times Y$ under the formula $\bar{G}(x, t) =$

$G(x) - tg$. Moreover, under this identification we get $\|\bar{G}\|$ equal to the expression in (4.1), due to the choice of norm in (4.7). The next thing to observe is that

$$(\bar{F}+\bar{G})(x,t) = \begin{cases} t(F+G)(t^{-1}x) - t(\bar{y}+g) & \text{if } t>0,\\ (F+G)^\infty(x) & \text{if } t=0,\\ \emptyset & \text{if } t<0,\end{cases}$$

so that $\bar{F}+\bar{G}$ gives the homogenization of the perturbed system $(F+G)(x) \ni \bar{y}+g$. Therefore, on the basis of what has so far been proved, we have

$$\bar{y}+g \in \operatorname{int}\operatorname{rge}(F+G) \quad\Leftrightarrow\quad \bar{F}+\bar{G} \text{ surjective.}$$

Hence, through Lemma 4.4, the distance to infeasibility for the system $F(x) \ni \bar{y}$ is the infimum of $\|\bar{G}\|$ over all $\bar{G} \in L(X \times I\!R, Y)$ such that $\bar{F}+\bar{G}$ is not surjective. Theorem 2.9 then furnishes the conclusion we wanted in (4.9). ◻

Theorem 4.8 (elaborated distance formula). *Let $F : X \rightrightarrows Y$ have closed, convex graph and let $\bar{y} \in \operatorname{rge} F$. Define the convex, positively homogeneous function $h : X^* \times Y^* \to (-\infty,\infty]$ by*

$$h(x^*,y^*) = \sup\nolimits_{x,y}\{\,\langle x,x^*\rangle - \langle y,y^*\rangle \mid y \in F(x)\}.$$

Then for the system $F(x) \ni \bar{y}$,

$$\text{(4.14)}\qquad \textit{distance to infeasibility} = \inf_{\|y^*\|=1,\ x^*} \max\Big\{\,\|x^*\|,\ h(x^*,y^*) + \langle \bar{y},y^*\rangle\Big\}.$$

Proof. By Theorem 4.7, the distance to infeasibility is $1/\operatorname{reg}\bar{F}(0,0\,|\,0)$. On the other hand, $\operatorname{reg}\bar{F}(0,0\,|\,0) = \|(\bar{F}^{*+})^{-1}\|^+$ for the adjoint mapping $\bar{F}^{*+} : Y^* \rightrightarrows X^* \times I\!R$; cf. Proposition 2.7. By definition, $(x^*,s) \in \bar{F}^{*+}(y^*)$ if and only if $(x^*,s,-y^*)$ belongs to the polar cone $(\operatorname{gph}\bar{F})^*$. Because $\operatorname{gph}\bar{F}$ is the closed convex cone generated by $\{\,(x,1,y-\bar{y}) \mid (x,y) \in \operatorname{gph} F\}$, this condition is the same as

$$s + \langle x,x^*\rangle - \langle y-\bar{y},y^*\rangle \le 0 \ \text{ for all } \ (x,y) \in \operatorname{gph} F$$

and can be expressed as $s + h(x^*,y^*) + \langle \bar{y},y^*\rangle \le 0$. Hence

$$\text{(4.15)}\qquad \|(F^{*+})^{-1}\|^+ = \sup\Big\{\,\|y^*\| \,\Big|\, \|(x^*,s)\| \le 1,\ s + h(x^*,y^*) + \langle \bar{y},y^*\rangle \le 0\Big\},$$

where the norm $\|(x^*,s)\|$ on $X^* \times I\!R$ dual to the one in (4.7) is $\|(x^*,s)\| = \max\{\,\|x^*\|,|s|\}$. The distance to infeasibility, being the reciprocal of the quantity in (4.15), can be expressed therefore (through the positive homogeneity of h) as

$$\text{(4.16)}\qquad \inf_{\|y^*\|=1,\ x^*,\,s}\Big\{\max\{\,\|x^*\|,|s|\}\,\Big|\, s + h(x^*,y^*) + \langle \bar{y},y^*\rangle \le 0\Big\}.$$

(In converting from (4.15) to an infimum restricted to $\|y^*\|=1$ in (4.16), we need to be cautious about the possibility that there might be no elements $(x^*,s,y^*) \in \operatorname{gph}(\bar{F}^{*+})^{-1}$ with $y^* \ne 0$, in which case the infimum in (4.16) is ∞. Is it correct then that the expression in (4.15) is 0? Yes.) Observe next that, in the infimum in (4.16), s will be taken to be as near to 0 as possible while maintaining $-s \ge h(x^*,y^*) + \langle \bar{y},y^*\rangle$. Thus, $|s|$ will be the max of 0 and $h(x^*,y^*) + \langle \bar{y},y^*\rangle$, and $\max\{\,\|x\|,|s|\}$ will be the max of these two quantities and $\|x\|$—but then the 0 is superfluous, and we end up with (4.16) equaling the expression on the right side of (4.14). ◻

Corollary 4.9 (homogeneous systems)**.** *Let $F : X \rightrightarrows Y$ be sublinear with closed graph and let $\bar{y} \in \operatorname{rge} F$. Then for the system $F(x) \ni \bar{y}$,*

$$\textit{distance to infeasibility} \;=\; \inf_{\|y^*\|=1} \max\Big\{\, d(0, F^{*+}(y^*)),\, \langle \bar{y}, y^* \rangle \Big\}.$$

Proof. In this case the function h in Theorem 4.8 has $h(x^*, y^*) = 0$ when $x^* \in F^{*+}(y^*)$, but $h(x^*, y^*) = \infty$ otherwise. □

Example 4.10 (linear-conic case)**.** *For a constraint system of type $x \geq_C 0$, $A(x) \leq_K \bar{y}$, with respect to a continuous linear mapping $A : X \to Y$ and closed, convex cones $C \subset X$ and $K \subset Y$,*

$$\textit{distance to infeasibility} \;=\; \inf_{y^* \in K^+,\, \|y^*\|=1} \max\Big\{\, d(A^*(y^*), C^+), \langle \bar{y}, y^* \rangle \Big\}.$$

Detail. This specializes to the mapping F in Example 4.2, using the formula for F^{*+} furnished there. We have $d(0, A^*(y^*) - C^+) = d(A^*(y^*), C^+)$. □

5 Application to Variational Inequalities

In this section we let $Y = X^*$ and consider a continuous mapping $f : X \to X^*$ along with a nonempty, closed, convex set $C \subset X$. We take

$$(5.1) \qquad F(x) := f(x) + N_C(x), \quad \text{where } N_C(x) = \begin{cases} \text{normal cone to } C \text{ at } x & \text{if } x \in C, \\ \emptyset & \text{if } x \notin C. \end{cases}$$

The normal cone is the usual one of convex analysis, so that

$$(5.2) \qquad \bar{x} \in F^{-1}(\bar{y}) \quad \Leftrightarrow \quad \bar{x} \in C \text{ and } \langle x - \bar{x},\, f(\bar{x}) - \bar{y} \rangle \geq 0 \text{ for all } x \in C.$$

In this case, therefore, solving $F(x) \ni \bar{y}$ amounts to solving the *variational inequality* for C and f with "forcing term" $\bar{y}$.

The concepts and results in Sections 1–3 can be applied in this framework of variational inequalities to gain information on how the solution set $F^{-1}(\bar{y})$ behaves with respect to perturbations in $\bar{y}$, or perturbations that replace F by $F + G$, which amount to replacing f by $f + G$. Note that F is positively homogeneous when f is positively homogeneous and C is a cone, but F is sublinear only when f is linear and C is a subspace of X.

If actually $C = X$, the variational inequality turns into the equation $f(x) = \bar{y}$ and we are back in the context of the Lusternik-Graves theorem, where the derivative mapping $Df(\bar{x}) : X \to X^*$ enters. Our aim now is to develop results about metric regularity for more general C in which $Df(\bar{x})$ and its adjoint $Df(\bar{x})^*$ likewise have a role. Our technique will be to rely on coderivatives, utilizing the Mordukhovich criterion in Theorem 3.1, so it will be necessary to restrict our analysis to finite-dimensional X.

Coderivatives of the normal cone mapping $N_C : x \mapsto N_C(x)$ will be needed. Clearly, when $(\bar{x}, \bar{y}) \in \operatorname{gph} F$ we have $\bar{y} - f(\bar{x}) \in N_C(\bar{x})$ and can work with the coderivative mapping $D^*N_C(\bar{x} \,|\, \bar{y} - f(\bar{x})) : X \rightrightarrows X^*$.

Theorem 5.1 (regularity radius for variational inequalities). *Suppose* $\dim X < \infty$ *and let* f *be continuously differentiable. Then for the mapping* F *in (5.1) and any* $(\bar{x}, \bar{y}) \in \operatorname{gph} F$,

$$\text{(5.3)} \qquad \operatorname{rad} F(\bar{x}\,|\,\bar{y}) \;=\; \inf_{\|u\|=1} d(-\,Df(\bar{x})^*(u),\, D^*N_C(\bar{x}\,|\,\bar{v})(u)\,), \quad \textit{where } \bar{v} = \bar{y} - f(\bar{x}).$$

Proof. Because $F = f + N_C$, we have $D^*F(\bar{x}\,|\,\bar{y}) = Df(\bar{x})^* + D^*N_C(\bar{x}\,|\,\bar{v})$ for $\bar{v} = \bar{y} - f(\bar{x})$ through the coderivative calculus in [23, 10.43(b)]. Hence by Theorem 3.1, $\operatorname{reg} F(\bar{x}\,|\,\bar{y})$ equals $\|(Df(\bar{x})^* + D^*N_C(\bar{x}\,|\,\bar{v}))^{-1}\|^+$. Proposition 2.5 then gives us

$$\operatorname{reg} F(\bar{x}\,|\,\bar{y})^{-1} \;=\; \inf_{\|u\|=1} d(0,\, Df(\bar{x})^*(u) + D^*N_C(\bar{x}\,|\,\bar{v})(u)).$$

Since $d(0,\, Df(\bar{x})^*(u) + D^*N_C(\bar{x}\,|\,\bar{v})(u)\,) = d(-\,Df(\bar{x})^*(u),\, D^*N_C(\bar{x}\,|\,\bar{v})(u)\,)$ and $\operatorname{rad} F(\bar{x}\,|\,\bar{y}) = 1/\operatorname{reg} F(\bar{x}\,|\,\bar{y})$ by Theorem 1.5, we get (5.3). □

Corollary 5.2 (effect of first-order approximations). *Under the assumptions in the theorem about* $F = f + N_C$, *one has*

$$\operatorname{rad} F(\bar{x}\,|\,\bar{y}) = \operatorname{rad} \Phi(\bar{x}\,|\,\bar{y}) \quad \textit{for } \Phi = \varphi + N_C$$

whenever $\varphi : I\!R^n \to I\!R^n$ *is continuously differentiable with* $\varphi(\bar{x}) = f(\bar{x})$ *and* $D\Phi(\bar{x}) = Df(\bar{x})$. *In particular this holds for the linearization* $\varphi(x) = f(\bar{x}) + Df(\bar{x})(x - \bar{x})$.

Proof. The radius formula in the theorem depends on f only through $Df(\bar{x})$. □

Especially worthy of attention here is a connection with the following property, which we define in general although planning for now to use it only for F of the type in (5.1).

Definition 5.3 (strong metric regularity). *A mapping* $F : X \rightrightarrows Y$ *will be called strongly metrically regular at* $\bar{x}$ *for* $\bar{y}$, *where* $(\bar{x}, \bar{y}) \in \operatorname{gph} F$, *if it is metrically regular and neighborhoods* X_0 *of* $\bar{x}$ *and* Y_0 *of* $\bar{y}$ *exist such that for each* $y \in Y_0$ *there is only one* $x \in X_0$ *in* $F^{-1}(y)$.

Through the characterization in (1.4) of metric regularity by way of the Aubin property, this means equivalently that F^{-1} has a single-valued localization around $(\bar{y}, \bar{x})$ that is Lipschitz continuous. (To say that a mapping has a *single-valued localization* around a point in its graph is to say that the intersection of the graph with some neighborhood of the point is the graph of a single-valued mapping defined on a neighborhood of the domain-projection of that point.)

The next theorem underscores the importance of this concept for variational inequalities

$$\text{(5.4)} \qquad g(w, x) + N_C(x) \ni y$$

having not only y and possibly an additional element w as parameters. The proof relies on results about polyhedral variational inequalities in the Dontchev-Rockafellar paper [6].

Theorem 5.4 (strong metric regularity from polyhedral convexity). *Let* $\dim X < \infty$, *and let* F *be of form (5.1) with* f *continuously differentiable and* C *polyhedral. Then at any* $(\bar{x}, \bar{y}) \in \operatorname{gph} F$, *metric regularity implies strong metric regularity. Moreover,* $\operatorname{rad} F(\bar{x}\,|\,\bar{y})$ *is in this case the largest radius* $r > 0$ *such that for any continuously differentiable mapping* $g : I\!R^d \times I\!R^n \to I\!R^n$ *(any* d*) and any* $\bar{w} \in I\!R^d$ *satisfying*

$$\text{(5.5)} \qquad g(\bar{w}, \bar{x}) = f(\bar{x}), \qquad \|D_x g(\bar{w}, \bar{x}) - Df(\bar{x})\| < r,$$

the solution mapping for the parameterized variational inequality in (5.4), namely

$$S : (y, w) \mapsto \{\, x \mid g(w, x) + N_C(x) \ni y \},\tag{5.6}$$

has a single-valued localization at $(\bar y, \bar w, \bar x)$ *that is Lipschitz continuous.*

Proof. For $g : I\!\!R^d \times I\!\!R^n \to I\!\!R^n$ continuously differentiable and $\bar w \in I\!\!R^d$, consider the partial linearization

$$f_g(x) \;=\; g(\bar w, \bar x) + D_x g(\bar w, \bar x)(x - \bar x).\tag{5.7}$$

By [6, Theorems 1 and 3], the solution mapping S in (5.6) has a single-valued localization at $(\bar y, \bar w, \bar x)$ that is Lipschitz continuous if and only if the mapping $y \mapsto \{\, x \mid f_g(x) + N_C(x) \ni y\}$ has the Aubin property at $\bar y$ for $\bar x$, or in other words, $F_g = f_g + N_C$ is metrically regular at $\bar x$ for $\bar y$. Here $F_g = F + G_g$ for $G_g = [f_g - f]$ and $\operatorname{lip} G_g(\bar x) = \|D_x g(\bar w, \bar x) - Df(\bar x)\|$. We know from Theorem 1.5 that $\operatorname{rad} F(\bar x \,|\, \bar y)$ is the largest radius $r > 0$ such that for any $G : X \to Y$ with $\operatorname{lip} G(\bar x) < r$, $F + G$ is metrically regular at $\bar x$ for $\bar y + G(\bar x)$. It follows that $\operatorname{rad} F(\bar x \,|\, \bar y)$ is the largest radius $r > 0$ such the conditions in (5.5) guarantee a single-valued Lipschitzian localization in (5.6).

When $g(w, x) \equiv f(x)$, we have $S(y, w) = F^{-1}(y)$. From this special case we deduce in particular that if F is metrically regular at $\bar x$ for $\bar y$ it is strongly metrically regular there. ◻

The single-valued Lipschitzian localization property of solution mappings S of the kind in (5.6), as implied by the corresponding property of $F_g = f_g + N_C$ with f_g as in (5.7), is what Robinson, that topic's pioneer, called in [21] the strong regularity of the parameterized variational inequality in (5.4). Theorem 5.4 reveals that the notion of strong metric regularity we have introduced in Definition 5.3, although posed in a different context, resonates closely with Robinson's notion.

The results in [6] also furnish for polyhedral C a description of coderivatives of the mapping N_C which can be brought into the formula in Theorem 5.4.

Proposition 5.5 (coderivative formula)**.** *Let* $\dim X < \infty$ *and let* $C \subset X$ *be a nonempty, convex set that is polyhedral. Let* $\bar v \in N_C(\bar x)$ *and let* $K(\bar x, \bar v)$ *be the associated critical cone to* C*, which is the polyhedral convex cone given by* $K(\bar x, \bar v) = \{\, u \in T_C(\bar x) \,|\, \langle u, \bar v\rangle = 0\}$*, where* $T_C(\bar x)$ *is the tangent cone to* C *at* $\bar x$*. Let* $\mathcal{K}$ *be the (finite) collection of all polyhedral (convex) cones* K *such that* $K = K' - K''$ *for closed faces* K' *and* K'' *of* $K(\bar x, \bar v)$ *with* $K' \supset K''$*. Then*

$$w \in D^* N_C(\bar x \,|\, \bar v)(u) \quad \Leftrightarrow \quad \exists\, K \in \mathcal{K} \ \text{ with } \ -u \in K,\, w \in K^*,\tag{5.8}$$

and in particular $\operatorname{dom} D^* N_C(\bar x \,|\, \bar v) = K(\bar x, \bar v) - K(\bar x, \bar v)$*, this being the subspace generated by* $K(\bar x, \bar v)$ *and also the largest of the cones* $K \in \mathcal{K}$*.*

Proof. Although not explicit in [6], this is the chief content of the proof of Theorem 2 there. ◻

Theorem 5.6 (radius formula from polyhedral convexity)**.** *Let* $\dim X < \infty$*, and let* F *be of form (5.1) with* f *continuously differentiable and* C *polyhedral. Let* F *be metrically regular at* $\bar x$ *for* $\bar y$*, and let* $\mathcal{K}$ *be defined as in Proposition 5.5 with respect to* $\bar v = \bar y - f(\bar x)$*. Then*

$$\operatorname{rad} F(\bar x \,|\, \bar y) = \min_{K \in \mathcal{K}} \rho_K(Df(\bar x)^*), \quad \text{where} \ \ \rho_K(A) = \min_{u \in K,\, \|u\|=1} d(A(u), K^*).\tag{5.9}$$

Proof. This simply combines the coderivative description in Proposition 5.5 with the formula in Theorem 5.1. □

The challenges in applying the formula in Theorem 5.6 are to identify the cone collection $\mathcal{K}$ and to compute the quantities $\rho_K(A)$ for $A = Df(\bar{x})^*$. When C is a box (a product of nonempty, closed intervals, not necessarily bounded), the collection $\mathcal{K}$ is easy to determine and the cones K it contains are very simple, being boxes themselves and thus the product of intervals of the form $(-\infty,\infty)$, $[0,\infty)$, $(-\infty,0]$ or $[0,0]$ (singleton); cf. [6]. The nature of $\rho_K(A)$, on the other hand, is illuminated by the following fact.

Proposition 5.7 (cone restrictions). *Let X be a Euclidean space (finite-dimensional with a Euclidean norm), and identify X^* with X. Let K be a closed convex cone in X, and let P_K be the (nearest point) projection mapping from X onto K. Then for any $A \in L(X,X)$, the quantity $\rho_K(A)$ defined in (5.9) has the alternative expression*

$$\rho_K(A) \;=\; 1/\|A_K^{-1}\|^+, \tag{5.10}$$

where A_K is the positively homogeneous mapping defined on X by

$$A_K(u) = \begin{cases} P_K(A(u)) & \text{if } u \in K, \\ \emptyset & \text{if } u \notin K. \end{cases} \tag{5.11}$$

In particular, $\rho_K(A) > 0$ if and only if A_K is nonsingular.

Proof. It is enough to observe that $d(w,K^*) = \|P_K(w)\|$ in this Euclidean setting. The formula then follows from Proposition 2.5 as applied to A_K, with reference also to Definition 2.2. □

Corollary 5.8 (subspace restrictions). *In the case of the proposition where K is a subspace of X, so that the mapping $A_K : K \to K$ is linear,*

$$\rho_K(A) \;=\; \operatorname{rad} A_K \;=\; \begin{cases} 1/\|A_K^{-1}\| & \text{if } A_K \text{ is nonsingular}, \\ 0 & \text{if } A_K \text{ is singular}. \end{cases} \tag{5.12}$$

Example 5.9 (radius in the nondegenerate case). *Suppose in Theorem 5.6 that X is Euclidean and the variational inequality*

$$F(x) \ni \bar{y} \quad \text{for} \quad F(x) = f(x) + N_C(x)$$

is nondegenerate at its solution $\bar{x}$, in the sense that the critical cone $K(\bar{x},\bar{v})$ is a subspace. Then

$$\operatorname{rad} F(\bar{x}\,|\,\bar{y}) \;=\; \operatorname{rad} A_K \qquad \text{for } A = Df(\bar{x})^* \text{ and } K = K(\bar{x},\bar{v}). \tag{5.13}$$

Detail. In this case the collection $\mathcal{K}$ in Proposition 5.5 has only one element, namely $K = K(\bar{x},\bar{v})$. We specialize (5.9) to this case and apply the formula in Corollary 5.8. □

The nondegenerate case in Example 5.9 is generic for polyhedral variational inequalities. All the complications in (5.9) not covered by the simple formula (5.13) are thus associated with various types of degeneracy.

As a more specific illustration of how the formula in Theorem 5.6 might be used to calculate the radius of metric regularity, consider the nonlinear programming problem

$$(5.14)\qquad \text{minimize}\ \ \psi_0(z)\ \ \text{subject to}\ \ \psi_i(z)\begin{cases} = 0 & \text{for } i\in[1,r],\\ \le 0 & \text{for } i\in[r+1,m],\end{cases}$$

where $\psi_i : I\!R^n \to I\!R, i = 0, 1, \dots, m$ are $\mathcal{C}^2$ functions. In terms of the Lagrangian function

$$L(z,\lambda) = \psi_0(z) + \sum_{i=1}^{m} \lambda_i\psi_i(z),$$

the first-order (Karush-Kuhn-Tucker) optimality condition for the problem (5.14) at a point $\bar z$ concerns the existence of a multiplier vector $\bar\lambda$ satisfying

$$\nabla_z L(\bar z,\bar\lambda) = 0, \quad \nabla_\lambda L(\bar z,\bar\lambda) \in N_\Lambda(\bar\lambda), \quad \text{where } \Lambda = I\!R^r \times I\!R^{m-r}_+.$$

This gradient condition is the variational inequality (5.1) with respect to $x = (z, \lambda)$ for

$$f(x) = (\nabla_x L(z,\lambda), -\nabla_\lambda L(z,\lambda)), \quad C = I\!R^n \times \Lambda, \quad \bar x = (\bar z,\bar\lambda), \quad \bar y = (0,0).$$

To compute the radius of the metric regularity of $F := f + N_C$ at $\bar x$ for $\bar y$, we appeal to (5.9). The cone collection $\mathcal{K}$ has been described in [6, Theorem 5] as follows. Define the index sets

$$\begin{aligned} I_1 &= \{i\in[r+1,m] \mid \psi_i(\bar z) = 0,\ \bar\lambda_i > 0\} \cup \{1,\cdots,r\}\\ I_2 &= \{i\in[r+1,m] \mid \psi_i(\bar z) = 0,\ \bar\lambda_i = 0\}\\ I_3 &= \{i\in[r+1,m] \mid \psi_i(\bar z) < 0\}.\end{aligned}$$

Then every cone K in the collection $\mathcal{K}$ is defined by a partition $\mathcal{I}$ of $\{1, 2, \cdots, m\}$ into index sets I_1', I_2', I_3' with the property $I_1 \subset I_1' \subset I_1 \cup I_2$ and $I_3 \subset I_3' \subset I_2 \cup I_3$ in the following way:

$$(z',\lambda') \in K \Longleftrightarrow \begin{cases} z' \text{ free}, & \\ \lambda_i' \text{ free} & \text{for } i\in I_1',\\ \lambda_i' \ge 0 & \text{for } i\in I_2',\\ \lambda_i' = 0 & \text{for } i\in I_3'.\end{cases}$$

Then

$$(z'',\lambda'') \in K^* \Longleftrightarrow \begin{cases} z'' = 0, & \\ \lambda_i'' = 0 & \text{for } i\in I_1',\\ \lambda_i'' \le 0 & \text{for } i\in I_2',\\ \lambda_i'' \text{ free} & \text{for } i\in I_3'.\end{cases}$$

Using this particular form of the cone K and its polar, and taking the norm to be the canonical Euclidean norm, we see that for a given $x' = (z', \lambda') \in K$ a simple rearrangement gives

$$\begin{aligned} d(Df(\bar x)^*(x'),K^*)^2 &= \|\nabla^2_{zz}L(\bar z,\bar\lambda)z' - \textstyle\sum_{i\in I_1'\cup I_2'}\lambda_i'\nabla\psi_i(\bar z)\|^2\\ &\quad + \textstyle\sum_{i\in I_1'}\langle\nabla\psi_i(\bar z),z'\rangle^2 + \sum_{i\in I_2'}\max\{0,\langle\nabla\psi_i(\bar z),z'\rangle\}^2,\end{aligned}$$

This expression can in principle be utilized in (5.9). Further simplifications may be available here and in other special cases, but we leave that for future investigation.

References

[1] J. M. Borwein, Norm duality for convex processes and applications, *Journal of Optimization Theory and Applications* **48** (1986), 53–64.

[2] J. M. Borwein, D. M. Zhuang, Verifiable necessary and sufficient conditions for openness and regularity of set-valued maps, *J. Math. Anal. Appl.* **134** (1988), 441–459.

[3] A. V. Dmitruk, A. A. Milyutin, N. P. Osmolovskiĭ, The Lusternik theorem and the theory of extremum, *Uspekhi Math. Nauk* **35** (1980), 11–46.

[4] A. L. Dontchev, The Graves theorem revisited, *J. Convex Anal.* **3** (1996), 45–53.

[5] A. L. Dontchev, W. W. Hager, An inverse function theorem for set-valued maps, *Proc. Amer. Math. Soc.* **121** (1994), 481–489.

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

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

[8] S. Filipowski, On the complexity of solving sparse symmetric linear approximate data, *Math. Oper. Research* **22** (1997), 769–792.

[9] R. M. Freund, J. R. Vera, Some characterizations and properties of the distance to ill-posedness and the condition number of a conic linear system, *Math. Programming, Ser. A* **86** (1999), 225–260.

[10] L. M. Graves, Some mapping theorems, *Duke Math. J.* **17** (1950), 111-114.

[11] R. A. Horn, C. Johnson, *Matrix Analysis.* Cambridge University Press, Cambridge, U.K., 1985.

[12] A. D. Ioffe, Metric regularity and subdifferential calculus, preprint.

[13] A. S. Lewis, Ill-conditioned convex processes and conic linear systems, *Math. of Oper. Research* **23** (1999), 829–834.

[14] A. S. Lewis, Ill-conditioned inclusions, *Set-Valued Analysis*, submitted.

[15] L. A. Lusternik, On the conditional extrema of functionals, *Math. Sbornik* **41** (1934), 390-401.

[16] B. S. Mordukhovich, Coderivatives of set-valued mappings: calculus and applications, *Nonlin. Anal.* **30** (1997), 3059–3070.

[17] J. Peña, Understanding the geometry of infeasible perturbations of a conic linear system, *SIAM J. Optim.* **10** (2000), 534–550.

[18] J. Renegar, Linear programming, complexity theory and elementary functional analysis, *Math. Programming Ser. A* **70** (1995), 279–351.

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

[20] S. M. Robinson, Regularity and stability for convex multifunctions, *Math. of Oper. Research* **1** (1976), 130–143.

[21] S. M. Robinson, Strongly regular generalized equations, *Math. of Oper. Research* **5** (1980), 43–62.

[22] R. T. Rockafellar, *Convex Analysis.* Princeton University Press, Princeton, N.J., 1970.

[23] R. T. Rockafellar, R. J-B Wets, *Variational Analysis*, Springer-Verlag, Berlin, 1997.

[24] W. Rudin, *Functional Analysis,* Second Ed., McGraw-Hill, New York, 1991.

[25] C. Ursescu, Multifunctions with closed convex graph, *Czechoslovak Mathematical Journal* **25** (1975), 438–441.