SIAM J. CONTROL AND OPTIMIZATION
Vol. 29, No. 2, pp. 493–497, March 1991

014

CALMNESS AND EXACT PENALIZATION*

J. V. BURKE‡

Abstract. The notion of calmness, which was introduced by Clarke and Rockafellar for constrained optimization, is considered. An equivalence to the technique of exact penalization due to Eremin and Zangwill is established. It is then shown that *calmness* is satisfied on a dense subset of the domain of the optimal value function.

Key words. exact penalization, calmness, constrained optimization

AMS(MOS) subject classifications. 49A42, 49D30, 49D37, 90D30

1. Introduction. The notion of calmness was originally formulated by Rockafellar and first appears in the literature of Clarke [3]. Since its appearance it has been recognized as a fundamental concept in optimization theory and consequently many variations on the original definition have been proposed and studied (e.g., see Rockafellar [6]). In general terms, calmness can be described as a basic regularity condition under which we can study the sensitivity properties of certain variational systems. On the other hand, the term "exact penalization" refers to a reduction principle in constrained optimization wherein we replace a constrained optimization problem by an unconstrained optimization problem whose objective is finite-valued on the domain of the original objective function and which under various conditions possesses a common local minimum. The particular reduction technique for exact penalization discussed herein originates in the papers of Eremin [4] and Zangwill [8] (also see Pietrzykowski [5]). We shall establish an equivalence between the notion of calmness and the viability of the Eremin–Zangwill exact penalization procedure for the constrained optimization problem

$$(\mathcal{P}) \qquad \begin{array}{l} \text{minimize } f(x) \\ \text{subject to } g(x) \in C, \end{array}$$

where f is a mapping from the normed linear space X into $\mathbb{R} \cup \{+\infty\}$, g is a mapping from X into the normed linear space Y, and C is a nonempty closed subset of Y. In order to make this statement precise we give the following definitions.

DEFINITION 1.1. Let f, g, X, Y, and C be as in the statement of $\mathcal{P}$ and consider the perturbed problems

$$(\mathcal{P}_u) \qquad \begin{array}{l} \text{minimize } f(x) \\ \text{subject to } g(x) \in C + u, \end{array}$$

where $u \in Y$. Let $\overline{x} \in X$ and $\overline{u} \in Y$ be such that

$$g(\overline{x}) \in C + \overline{u} \text{ and } \overline{x} \in \operatorname{dom}(f) := \{x \in X : f(x) < +\infty\}.$$

*Received by the editors September 8, 1989; accepted for publication (in revised form) March 26, 1990.

‡Department of Mathematics, GN–50, University of Washington, Seattle, Washington 98195.
This work was supported in part by National Science Foundation grants DMS–8602399 and DMS–8803206 and by Air Force Office of Scientific Research grant ISSA-860080.

The problem $\mathcal{P}_{\overline{u}}$ is said to be *calm* at $\overline{x}$ if there are constants $\overline{\alpha} \geq 0$ and $\varepsilon > 0$ such that for every pair $(x, u) \in X \times Y$ with $\|x - \overline{x}\| \leqq \varepsilon$ and $g(x) \in C + u$, we have

$$f(x) + \overline{\alpha}\|u - \overline{u}\| \geq f(\overline{x}). \tag{1.1}$$

Here we use the notation $\|z\|$ for the norm of z. The constants $\overline{\alpha}$ and ε are called the *modulus* and *radius* of calmness for $\mathcal{P}_{\overline{u}}$ at $\overline{x}$, respectively.

Remarks. (1) Definition 1.1 varies from the definition given by Clarke [2, Def. 6.4.1] in that the variable u is not required to satisfy $\|u - \overline{u}\| \leq \varepsilon$ in order for inequality (1.1) to hold. In §2, we show that the restriction on the perturbation u is redundant.

(2) Observe that if $\mathcal{P}_{\overline{u}}$ is calm at $\overline{x}$, then $\overline{x}$ is necessarily a local solution to $\mathcal{P}_{\overline{u}}$.

DEFINITION 1.2. Let f be a mapping from the normed linear space X into $\mathbb{R} \cup \{+\infty\}$ and let S be a subset of X. Let $\varepsilon > 0$. We say that $\overline{x} \in S$ is an *ε-local minimum* of f on S if

$$f(\overline{x}) \leq f(x)$$

for all $x \in S$ with $x \in \overline{x} + \varepsilon\mathbb{B}$ where $\mathbb{B}$ is the closed unit ball in X, i.e., $\mathbb{B} := \{v \in X : \|v\| \leqq 1\}$, and $\overline{x} + \varepsilon\mathbb{B} := \{\overline{x} + \varepsilon v : v \in \mathbb{B}\}$.

The main result of this paper can now be stated.

THEOREM 1.1. *Let $\overline{x} \in X$ and $\overline{u} \in Y$ be such that*

$$g(\overline{x}) \in C + \overline{u} \text{ and } \overline{x} \in \operatorname{dom}(f).$$

Then $\mathcal{P}_{\overline{u}}$ is calm at $\overline{x}$ with modulus $\overline{\alpha}$ and radius ε if and only if $\overline{x}$ is an ε-local minimum of

$$P_{\overline{u},\overline{\alpha}}(x) := f(x) + \alpha \operatorname{dom}(g(x)|C + \overline{u}),$$

where we define

$$\operatorname{dist}(z|C + \overline{u}) := \inf\{\|y + \overline{u} - z\| : y \in C\}.$$

Proof. ($\Longrightarrow$) Let $\delta > 0$. Given $x \in \overline{x} + \varepsilon\mathbb{B}$, set $u_x := g(x) - y_x$, where $y_x \in C$ satisfies

$$\|y_x + \overline{u} - g(x)\| \leq \operatorname{dist}(g(x)|C + \overline{u}) + \delta.$$

Note that $g(x) \in C + u_x$ and

$$\|\overline{u} - u_x\| \leq \operatorname{dist}(g(x)|C + \overline{u}) + \delta.$$

Thus, if $\alpha \geq \overline{\alpha}$, we obtain from the calmness hypothesis that

$$\begin{aligned} f(\overline{x}) + \alpha \operatorname{dist}(g(\overline{x})|C + \overline{u}) &= f(\overline{x}) \\ &\leq f(x) + \alpha\|\overline{u} - u_x\| \\ &\leq f(x) + \alpha \operatorname{dist}(g(x)|C + \overline{u}) + \alpha\delta. \end{aligned}$$

Since $\delta > 0$ was chosen arbitrarily the implication is established.

($\Longleftarrow$) Let $u \in Y$ and $x \in \overline{x} + \varepsilon\mathbb{B}$ be such that $g(x) \in C + u$ and $x \in \operatorname{dom}(f)$. Then

$$\begin{aligned} f(\overline{x}) &\leqq f(x) + \overline{\alpha} \operatorname{dist}(g(x)|C + \overline{u}) \\ &= f(x) + \overline{\alpha} \inf\{\|y + \overline{u} - g(x)\| : y \in C\} \\ &\leq f(x) + \overline{\alpha} \inf\{\|y + u - g(x)\| + \|u - \overline{u}\| : y \in C\} \\ &= f(x) + \overline{\alpha} \operatorname{dist}(g(x)|C + u) + \overline{\alpha}\|u - \overline{u}\| \\ &= f(x) + \overline{\alpha}\|u - \overline{u}\|. \end{aligned}$$

Hence $\mathcal{P}_{\overline{u}}$ is calm at $\overline{x}$. □

Remark. The function $P_{u,\alpha}$ defined above is a familiar tool in the mathematical programming literature [2],[4],[5],[8]. For example, in the case where $X := \mathbb{R}^n$, $C := \mathbb{R}^m_- \subset \mathbb{R}^m =: Y$, and Y is endowed with the l_1 norm, we have

$$P_{u,\alpha}(x) = f(x) + \alpha \sum_{i=1}^{m} (g_i(x) - u_i)_+,$$

where g_i and u_i are the ith components of g and u, respectively, and $z_+ := \max\{0, z\}$ for every $z \in \mathbb{R}$.

In the case where Y is finite-dimensional and $g(x)$ is Lipschitz continuous, Clarke [2, Prop. 6.4.3] has shown that the calmness of $\mathcal{P}_{\overline{u}}$ at $\overline{x}$ implies the existence of a constant $\overline{m} > 0$ such that $\overline{x}$ is a local minimum for the function $P_{\overline{u},m}(x)$ for all $m \geqq \overline{m}$. However Clarke's result does not reveal the full extent of the relationship between calmness and exact penalization. In particular, it does not describe the relationship between the parameters $\overline{m}$ and $\overline{\alpha}$ as given in Theorem 1.1.

As previously stated calmness is an important tool in the sensitivity analysis for $\mathcal{P}$. In this regard Theorem 1.1 can be used to study the sensitivity of $\mathcal{P}_u$ to changes in u and to establish multiplier rules for $\mathcal{P}$. These results and others are pursued in Burke [1]. In the remainder of this note we briefly explore two topics directly related to the definitions of calmness and exact penalization as they are employed in Theorem 1.1. In §2 we compare Definition 1.1 to the definition for calmness used by Clarke in [2, Def. 6.4.1]. We conclude in §3 by providing a result that is in the spirit of Clarke's generic calmness result [2, Prop. 6.4.5] indicating the robustness of the notion of calmness.

2. Another formulation of calmness.

According to Clarke [2, Def. 6.4.1] in order for $\mathcal{P}_{\overline{u}}$ to be calm at $\overline{x}$ we require that inequality (1.1) be satisfied whenever $\|x - \overline{x}\| \leqq \varepsilon$ and $\|u - \overline{u}\| \leqq \varepsilon$. In the next proposition we show that if g is continuous at $\overline{x}$, then no advantage gained by placing this further restriction on the choice of perturbation u.

PROPOSITION 2.1. *Let f, g, C, X, and Y be as in the statement of problem $\mathcal{P}$. Let $(\overline{x}, \overline{u}) \in X \times Y$ be such that g is continuous at $\overline{x}$ and $g(\overline{x}) \in C + \overline{u}$. If there is an $\overline{\alpha} > 0$ and an $\varepsilon > 0$ such that*

$$f(x) + \overline{\alpha}\|u - \overline{u}\| \geq f(\overline{x})$$

for every pair $(x, u) \in X \times Y$ with $\|u - \overline{u}\| \leq \varepsilon$, $\|x - \overline{x}\| \leq \varepsilon$, and $g(x) \in C + u$, then there is an $\widehat{\varepsilon}$ with $0 < \widehat{\varepsilon} \leq \varepsilon$ such that $\overline{x}$ is an $\widehat{\varepsilon}$-local minimum of $P_{\overline{u},\overline{\alpha}}(x)$, and consequently, $\mathcal{P}_{\overline{u}}$ is calm at $\overline{x}$ with modulus $\overline{\alpha}$ and radius $\widehat{\varepsilon}$.

Proof. Let $\delta \in (0, \frac{1}{2})$ and $\alpha \geq \overline{\alpha}$. Since the function $\varphi(x) := \operatorname{dist}(g(x)|C + \overline{u})$ is continuous at $\overline{x}$, there is an $\widehat{\varepsilon} \in (0, \varepsilon]$ such that $0 \leq \varphi(x) \leq \frac{\varepsilon}{2}$ whenever $\|x - \overline{x}\| \leq \widehat{\varepsilon}$. Now given $x \in \overline{x} + \widehat{\varepsilon}\mathbb{B}$, set $u_x := g(x) - y_x$ where $y_x \in C$ satisfies

$$\|y_x + \overline{u} - g(x)\| \leq \operatorname{dist}(g(x)|C) + \delta\varepsilon.$$

Then $g(x) \in C + u_x$ and $\|u_x - \overline{u}\| \leq \varepsilon$. Hence, by hypothesis,

$$\begin{aligned} \mathcal{P}_{\overline{u},\alpha}(\overline{x}) &= f(\overline{x}) \\ &\leq f(x) + \alpha\|u_x - \overline{u}\| \\ &\leq f(x) + \alpha \operatorname{dist}(g(x)|C) + \alpha\delta\varepsilon. \end{aligned}$$

Taking the limit as $\delta \downarrow 0$ yields the result. $\square$

Calmness can also be defined independent of the existence of a solution to $\mathcal{P}_u$. This is done by considering the value function for $\mathcal{P}_u$, $V : X \to \mathbb{R} \cup \{\pm\infty\}$, given by

$$V(u) := \inf\{f(x) : g(x) \in C + u\}$$

if $\{x : g(x) \in C + u\} \neq \emptyset$ and $+\infty$ otherwise. $\mathcal{P}_u$ is then said to be calm at $\overline{u}$ if

$$\liminf_{u \to \overline{u}} \frac{V(u) - V(\overline{u})}{\|u - \overline{u}\|} > -\infty. \tag{2.1}$$

In this connection we have the following straightforward extension to Clarke [2, Prop. 6.4.2].

PROPOSITION 2.2. *Let $f, g.C, X$, and Y be as in the statement of $\mathcal{P}$, and let $\overline{u} \in Y$. If* (2.1) *holds at $\overline{u}$, then for any solution $\overline{x}$ to $\mathcal{P}_{\overline{u}}, \mathcal{P}_{\overline{u}}$ is calm at $\overline{x}$.*

3. Calmness is a dense property in finite dimensions. In the case where P_u has the representation

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \le u_i, \ i = 1, \cdots, m \\ & \text{and } x \in D, \end{array}$$

where $f_i : \mathbb{R}^n \to \mathbb{R}$ is locally Lipschitz for each $i = 0, 1, \cdots, m$ and $D \subset \mathbb{R}^n$ is a nonempty closed set, we can employ the results of Clarke [2, §6.4] to show that if $\mathcal{P}_u$ has a finite value for every u near some $\overline{u} \in \mathbb{R}^m$, then for almost every u near $\overline{u}$ (in the sense of Lebesgue measure) $\mathcal{P}_u$ is calm at u. In the spirit of this result we give the following proposition.

PROPOSITION 3.1. *Let f, g, C, X, and Y be as in $\mathcal{P}$. Furthermore, assume that Y is finite-dimensional, f is lower semicontinuous, and g is continuous. If $\overline{u} \in Y$ and $\gamma > 0$ are such that V is bounded on $\overline{u} + \gamma\mathbb{B}$, then $\mathcal{P}_u$ is calm on a dense subset of $\overline{u} + \gamma\mathbb{B}$.*

Proof. With no loss in generality we assume that $\overline{u} = 0$. Let $u \in Y$ be an element of the interior of $\gamma\mathbb{B}$ and $\varepsilon > 0$. We must show that there is a $u_0 \in u + \varepsilon\mathbb{B}$ such that $\mathcal{P}_{u_0}$ is calm. Define $\theta : [0, \gamma - \|u\|] \to \mathbb{R}$ by

$$\theta(\rho) := \inf\{f(x) : g(x) \in C + u + \rho\mathbb{B}\}.$$

The boundedness of $\theta(\rho)$ on $[0, \gamma - \|u\|]$ follows from that of V on $\gamma\mathbb{B}$. Since θ is nonincreasing on $[0, \gamma - \|u\|]$, θ is differentiable at almost every $\rho \in [0, \gamma - \|u\|]$ (in the sense of Lebesgue measure) by Ward [7]. Let ρ_0 be a point of differentiability for θ such that $0 < \rho_0 < \min\{\varepsilon, \gamma - \|u\|\}$. From the definition of θ there is for each $n \in \{1, 2, \cdots\}$ a $u_n \in u + \rho_0\mathbb{B}$ such that

$$V(u_n) - \frac{1}{n} \le \theta(\rho_0). \tag{3.1}$$

Let u_0 be a cluster point of the sequence $\{u_n\}$. Now since f is lower semi continuous and g is continuous, we have that V is also lower semicontinuous, hence, by (3.1), it must be that $V(u_0) = \theta(\rho_0)$. We now show that $\mathcal{P}_u$ is calm at u_0.

Since θ is differentiable at ρ_0, there is a $\delta \in (0, \gamma - \|u\| - \rho_0)$ and an $\alpha > 0$ such that

$$\theta(\rho) - \theta(\rho_0) \ge -\alpha|\rho - \rho_0|$$

whenever $|\rho-\rho_0| < \delta$. Let $\varepsilon_0 \in (0, \min\{\delta, \min\{\varepsilon, \gamma - \|u\|\} - \rho_0\})$, and let $w \in u_0 + \varepsilon_0\mathbb{B}$. Then

$$V(w) - V(u_0) \geqq \theta(\|u - w\|) - \theta(\rho_0) \geqq -\alpha|\ \|u - w\| - \rho_0|. \tag{3.2}$$

Now if $\|u - w\| \leq \rho_0$, then $\theta(\|u - w\|) \geq \theta(\rho_0)$ so that

$$V(w) - V(u_0) \geq 0 \geq -\alpha\|w - u_0\|.$$

On the other hand, if $\|u - w\| \geq \rho_0$, then

$$|\ \|u - w\| - \rho_0| \leq \|w - u_0\|,$$

and hence

$$V(w) - V(u_0) \geq -\alpha\|w - u_0\|$$

by (3.2). Therefore

$$V(w) + \alpha\|w - u_0\| \geq V(u_0)$$

for all $w \in u_0 + \varepsilon_0\mathbb{B}$. Consequently, $\mathcal{P}_u$ is calm at u_0. □

The conclusion of Proposition 3.1 is weaker than that of Clarke [2, §6.4] since we do not show that $\mathcal{P}_u$ is calm for almost all u near $\overline{u}$. On the other hand, our result is valid for more general constraints than those considered by Clarke.

Acknowledgment. Just before the final version of this article was sent to the Publisher, we became aware of an important reference on this topic. The reference is Thibault [9]. In this thesis, Thibault establishes certain equivalences between the notions of calmness and exact penalization. These results, although similar to our own, are somewhat different. The results of Thibault are complementary to those presented in this paper.

REFERENCES

[1] J.V. BURKE, *An exact penalization viewpoint of constrained optimization*, Report ANL/MCS-TM-95, Mathematics and Computer Science Division, Argonne National Laboratory, Argonne, Ill. 1987.

[2] F.H. CLARKE, *Optimization and Nonsmooth Analysis,* John Wiley, New York, 1983.

[3] ———, *A new approach to Lagrange multipliers*, Math. Oper. Res., 1 (1976), pp.165–174.

[4] I.I. EREMIN, *The 'penalty' method in convex programming*, Dokl. Akad, Nauk SSSR, in English. 173 (1966), pp.459–462.

[5] T. PIETRZYKOWSKI, *An exact potential method for constrained maxima*, SIAM J. Numer. Anal., 6 (1969), pp.299–304.

[6] R.T. ROCKAFELLAR, *Extensions of subgradient calculus with applications to optimization*, Nonlinear Anal., 9 (1985), pp.665–698.

[7] A.J. WARD, *On the differential structure of real functions*, London Math. Soc., 39 (1934), pp.339–362.

[8] W.I. ZANGWILL, *Nonlinear programming via penalty functions*, Management Sci.13 (1967), pp.344-358.

[9] L. THIBAULT, *Proprietes des Sous-Differentiels de Fonctions Localement Lipschitziennes Definies sur un Espace de Banach Separable. Applications.*, Ph.D Thesis, Department of Mathematics, University of Montpellier, Montpellier, France, 1975.