Journal of Global Optimization, **29:** 497–509, 2004.

Some Results about Duality and Exact Penalization*

Y.Y. ZHOU[1] and X.Q. YANG[2]
[1]*Department of Mathematics, Suzhou University, Suzhou, Jiangsu 215006, China*
[2]*Department of Applied Mathematics, The Hong Kong Polytechnic University, Kowloon, Hong Kong, P.R. China*

(Received 1 August 2002; accepted 13 August 2003)

Abstract. In this paper, we introduce the concept of the valley at 0 augmenting function and apply it to construct a class of valley at 0 augmented Lagrangian functions. We establish the existence of a path of optimal solutions generated by valley at 0 augmented Lagrangian problems and its convergence toward the optimal set of the original problem and obtain the zero duality gap property between the primal problem and the valley at 0 augmented Lagrangian dual problem. Moreover, we establish the exact penalization representation results in the framework of valley at 0 augmented Lagrangian.

Key words. exact penalty function, optimal path, valley at 0 augmenting function, zero duality gap.

1. Introduction

Let $\bar{R}=R\cup\{+\infty,-\infty\}$ and φ: $R^n\to\bar{R}$ be an extended real-valued function. Consider the primal problem

$$\inf_{x\in R^n}\varphi(x). \tag{1}$$

In the Lagrangian dual approach, appropriate augmenting functions are in demand in defining dual problems for (1) such that the zero duality gap holds. In [8] a convex augmenting function was introduced and the corresponding zero duality property was obtained. A level-bounded augmenting function was given in [3] where the convexity of augmenting functions in [8] is replaced by a level-boundedness condition. Furthermore, a peak at 0 augmenting function was introduced in [9] and applied to establish an equivalence of the zero duality gap properties of a corresponding augmented Lagrangian dual problem and a Lagrangian-type dual problem [10]. It is clear that a level-bounded augmenting function is valley at 0, but not vice versa. It is still unknown whether the valley at 0 function guarantees a zero duality gap. This paper gives a positive answer to this question.

The existence and convergence of an optimal path generated by penalty/dual problems toward the optimal set is important for numerical solution methods,

*This research is supported by the Research Committee of the Hong Kong Polytechnic University and NNSF(10161010) of China.

see [1, 3, 12]. In this paper, we introduce the concept of valley at 0 augmenting functions and apply it to construct a class of valley at 0 augmented Lagrangian. We establish the existence of a path of optimal solutions generated by valley at 0 augmented Lagrangian problems and its convergence toward the optimal set. A corresponding zero duality gap between the primal problem and the augmented Lagrangian dual problem is obtained. We also establish the exact penalization representation results in the framework of new augmented Lagrangian under weaker conditions than the ones in [3]. It is worth noting that, as the valley property is a much weaker condition than the level boundeness, the techniques used in this paper are different from the ones in [3].

The outline of this paper is as follows. In Section 2, we introduce the concepts of a valley at 0 augmenting function and a corresponding augmented Lagrangian. In Section 3, we establish the existence of a path of optimal solutions generated by augmented Lagrangian problems and its convergence to the optimal set for the primal problem. In Section 4, we obtain necessary and sufficient conditions for an exact penalty representation in the framework of new augmented Lagrangians.

2. Augmented Lagrangians

A function $\bar{f}$: $R^n \times R^m \to \bar{R}$ is said to be a dualizing parameterization function for φ if $\varphi(x) = \bar{f}(x, 0), \forall x \in R^n$.

DEFINITION 2.1 [8]. (i) Let $X \subset R^n$ be a closed subset and f: $X \to \bar{R}$ be an extended real-valued function. The function f is said to be level-bounded on X if, for any $\alpha \in R$, the set $\{x \in X\colon f(x) \leqslant \alpha\}$ is bounded. (ii) A function $\bar{f}$: $R^n \times R^m \to \bar{R}$ with value $\bar{f}(x, u)$ is said to be level-bounded x locally uniform in u if, for each $\bar{u} \in R^m$ and $\alpha \in R$, there exists a neighborhood $U(\bar{u})$ of $\bar{u}$ along with a bounded set $D \subset R^n$, such that $\{x \in R^n\colon \bar{f}(x, u) \leqslant \alpha\} \subset D$ for any $u \in U(\bar{u})$.

DEFINITION 2.2. The function f: $X \to \bar{R}$ is said to have a valley at 0 in X if, $f(0) = 0$, $f(x) > 0$, for all $x \neq 0$, and $c_\delta = \inf_{\|x\| \geqslant \delta} f(x) > 0$, for each $\delta > 0$.

It is clear that a continuous function f has a valley at 0 if and only if $-f$ has a peak at 0, see [9].

DEFINITION 2.3. (i) A function σ: $R^m \to R_+ \cup \{+\infty\}$ is said to be a generalized augmenting function if it is proper, lower semicontinuous (lsc, for short), level-bounded on R^m, $\operatorname{argmin}_y \sigma(y) = \{0\}$ and $\sigma(0) = 0$. (ii) A function σ: $R^m \to R_+ \cup \{+\infty\}$ is said to be a valley at 0 augmenting function if it is proper, lower semicontinuous, has a valley at 0 in R^m.

Remark 2.1. It is easy to see if σ is a generalized augmenting function, then it is a valley at 0 augmenting function. In fact, if σ hasn't valley at 0 in R^m, then there exists $\delta > 0$ such that $c_\delta = \inf_{\|x\| \geqslant \delta} \sigma(x) = 0$, hence there exists $\{x_j\} \subseteq R^m$, $\|x_j\| \geqslant \delta$ such that $\sigma(x_j) \to c_\delta$. If $\{x_j\}$ is unbounded, by the definition of level-bounded function, we get $\{\sigma(x_j)\}$ is unbounded. Thus there exists a subsequence of $\{\sigma(x_j)\}$ converging to infinite. This contradicts to $c_\delta = 0$. So $\{x_j\}$ is bounded, there exists x_0, such that $x_j \to x_0$. The lsc of σ and $\operatorname{argmin}_y \sigma(y) = \{0\}$ imply $\sigma(x_0) = 0$, hence $x_0 = 0$, but $\|x_0\| \geqslant \delta$, which implies a contradiction.

If σ is a valley at 0 augmenting function, then it may be isn't a generalized augmenting function. For example, let $u = (u_1, \ldots, u_m) \in R^m$ and $0 < \gamma < 1$,

$$\sigma(u) = \begin{cases} \|u\|_p^\gamma, & \text{if } \|u\|_p \leqslant 1 \\ 1, & \text{if } \|u\|_p > 1 \end{cases},$$

where

$$\|u\|_p = \begin{cases} \left(\sum_{j=1}^m |u_j|^p\right)^{1/p}, & \text{if } 1 \leqslant p < +\infty \\ \max\{|u_1|, \ldots, |u_m|\}, & \text{if } p = +\infty \end{cases}.$$

It is easy to check σ is a valley at 0 augmenting function, but it is not generalized augmenting function.

DEFINITION 2.4. Consider the primal problem (1). Let $\bar{f}$ be any dualizing parameterization function for φ, and σ be a valley at 0 augmenting function.

(i) The valley at 0 augmented Lagrangian (with parameter $r > 0$) $\bar{l}$: $R^n \times R^m \times (0, +\infty) \to \bar{R}$ is defined by

$$\bar{l}(x, y, r) = \inf\{\bar{f}(x, u) - \langle y, u \rangle + r\sigma(u) \colon u \in R^m\},$$
$$x \in R^n, \ y \in R^m, \ r > 0,$$

where $\langle y, u \rangle$ denotes the inner product.

(ii) The valley at 0 augmented Lagrangian dual function is defined by

$$\bar{\psi}(y, r) = \inf\{\bar{l}(x, y, r) \colon x \in R^n\}, \quad y \in R^m, \quad r > 0. \tag{2}$$

(iii) The valley at 0 augmented Lagrangian dual problem is defined as

$$\sup_{(y,r) \in R^m \times (0, +\infty)} \bar{\psi}(y, r). \tag{3}$$

Now we discuss the property of the valley at 0 augmented Lagrangian for the primal problem (1).

Define the perturbation function by

$$p(u) = \inf\{\bar{f}(x,u)\colon x \in R^n\}.$$

Then $p(0)$ is just the optimal value of the problem (1).

The following proposition summarizes some basic properties of the valley at 0 augmented Lagrangian, which will be useful in the sequel. Its proof is elementary and omitted.

PROPOSITION 2.1. *For any dualizing parameterization and any valley at* 0 *augmenting function, we have*

(i) *the valley at* 0 *Lagrangian* $\bar{l}(x,y,r)$ *is concave, upper semicontinuous in* (y,r) *and nondecreasing in* r.
(ii) *weak duality holds*:

$$\bar{\psi}(y,r) \leqslant p(0), \quad \forall (y,r) \in R^m \times (0,+\infty). \tag{4}$$

3. Optimal Paths and Zero Duality Gaps

In this section, we establish the existence of a path of optimal solutions generated by valley at 0 augmented Lagrangian problems and its convergence to the optimal solution set of the primal problem (see Theorem 3.1). Using Theorem 3.1, we get a zero duality gap property between the primal problem (1) and its valley at 0 augmented Lagrangian dual problem (3) (see Theorem 3.2).

Consider the primal problem (1) and its valley at 0 augmented Lagrangian problem:

$$P(y,r) \inf_{(x,u)\in R^n\times R^m} \{\bar{f}(x,u) + r\sigma(u) - \langle y,u\rangle\}.$$

Note that $P(y,r)$ is the same as the problem of evaluating the valley at 0 augmented Lagrangian dual function $\bar{\psi}(y,r)$. Let S and $V(y,r)$ denote the optimal solution sets of the problems (1) and $P(y,r)$, respectively. Recall that $p(0)$ and $\bar{\psi}(y,r)$ are the optimal values of the problems (1) and $P(y,r)$, respectively.

We have the following general results concerning the existence of an optimal path.

If the dualizing function $\bar{f}(x,u)$ is lsc, and level-bounded in x, we see that φ is proper, lsc and level-bounded. It follows that S is nonempty and compact, see [3].

THEOREM 3.1 (Optimal path). *Consider the primal problem* (1) *and its valley at* 0 *augmented Lagrangian problem* $P(y,r)$. *Assume that* φ *is proper, and that its dualizing parameterization function* $\bar{f}(x,u)$ *is proper, lsc, and level-bounded* *in x locally uniform in u. Suppose that there exists $(\bar{y},\bar{r})\in R^m\times(0,+\infty)$ such that*

$$\bar{m}=\inf\{\bar{l}(x,\bar{y},\bar{r}):\ x\in R^n\}>-\infty. \tag{5}$$

Then

(i) *There exists $r_0>\bar{r}$, such that for any $r\geqslant r_0$, $V(\bar{y},r)$ is nonempty and compact.*
(ii) *For each selection $(x(r),u(r))\in V(\bar{y},r)$ with $r\geqslant r_0$, the optimal path $\{(x(r),u(r))\}$ is bounded and its limit points take the form $(x^*,0)$, where $x^*\in S$.*

Proof. (i) Let $m_r=\inf\{\bar{l}(x,\bar{y},r),x\in R^n\}$. From (5), we have $m_r\geqslant\bar{m}$ for all $r>\bar{r}$. By the definition of m_r, there exists $(x_j,u_j)\in R^n\times R^m$, satisfying

$$\bar{f}(x_j,u_j)+r\sigma(u_j)-\langle\bar{y},u_j\rangle\to m_r. \tag{6}$$

Hence, for some $\epsilon_0\geqslant 0$, there exists an integer $N>0$, such that

$$\bar{f}(x_j,u_j)+r\sigma(u_j)-\langle\bar{y},u_j\rangle\leqslant m_r+\epsilon_0,\quad\forall j>N. \tag{7}$$

It is clear that $m_r\leqslant p(0)$. Since σ has a valley at 0, $c_\delta=\inf_{\|x\|\geqslant\delta}\sigma(x)>0$, for each $\delta>0$. Denote $r_1=\frac{p(0)+\epsilon_0-\bar{m}}{c_\delta}+\bar{r}$, we will prove that $V(\bar{y},r)$ is nonempty for each $r>r_1$. From (5) and (7), we have

$$\sigma(u_j)\leqslant\frac{p(0)+\epsilon_0-\bar{m}}{r-\bar{r}}<c_\delta,\quad\forall j>N$$

This implies that $u_j\in\{u\in R^m:\ \|u\|\leqslant\delta\}$, $\forall j>N$. Let $\delta_1=\max\{\|u_1\|,\|u_2\|,\ldots,\|u_N\|\}$. Again from (7), we have

$$\bar{f}(x_j,u_j)\leqslant p(0)+\epsilon_0+\|\bar{y}\|(\delta+\delta_1),\quad\forall j. \tag{8}$$

Because $\bar{f}(x,u)$ is level-bounded in x locally uniform in u, it follows from (8) that $\{x_j\}$ is bounded, so $\{(x_j,u_j)\}$ is bounded. We may assume that $(x_j,u_j)\to(x_0,u_0)$. The lsc of f and σ, together with (6), implies

$$\begin{aligned}\bar{f}(x_0,u_0)+r\sigma(u_0)-\langle\bar{y},u_0\rangle&\leqslant\liminf_{j\to+\infty}\bar{f}(x_j,u_j)+r\liminf_{j\to+\infty}\sigma(u_j)-\lim_{j\to+\infty}\langle\bar{y},u_j\rangle\\&\leqslant\lim_{j\to+\infty}(\bar{f}(x_j,u_j)+r\sigma(u_j)-\langle\bar{y},u_j\rangle)=m_r.\end{aligned} \tag{9}$$

Hence

$$\bar{f}(x_0,u_0)+r\sigma(u_0)-\langle\bar{y},u_0\rangle=m_r.$$

That is $V(\bar{y}, r) \neq \phi$ for each $r > r_1$. The lsc of f and σ implies that $V(\bar{y}, r)$ is closed.

Let $\bar{x} \in R^n$ such that $-\infty < \varphi(\bar{x}) < +\infty$. Let

$$\mathrm{O}(r) = \{(x, u) \in R^n \times R^m \colon \bar{f}(x, u) + r\sigma(u) - \langle \bar{y}, u \rangle \leqslant \varphi(\bar{x})\}.$$

Denote $r_0 = \max\{r_1, \frac{\varphi(\hat{x}) - \bar{m}}{c_\delta} + \bar{r}\}$, we prove that $\mathrm{O}(r)$ is a bounded set for each $r > r_0$. Suppose to the contrary that $\exists (x_j, u_j) \in \mathrm{O}(r)$ such that $\|(x_j, u_j)\| \to +\infty$. Since $(x_j, u_j) \in \mathrm{O}(r)$, we have

$$\bar{f}(x_j, u_j) + r\sigma(u_j) - \langle \bar{y}, u_j \rangle \leqslant \varphi(\bar{x}). \tag{10}$$

This, combined with (5), yields

$$\sigma(u_j) \leqslant \frac{\varphi(\bar{x}) - \bar{m}}{r - \bar{r}} < c_\delta.$$

Then $u_j \in \{u \in R^m, \|u\| \leqslant \delta\}$. From (10), we have

$$\bar{f}(x_j, u_j) \leqslant \varphi(\bar{x}) + \|\bar{y}\|\delta.$$

Hence, it follows from the level-boundedness in x locally uniformness in u of $\bar{f}(x, u)$ that $\{x_j\}$ is bounded, so $\{(x_j, u_j)\}$ is bounded, a contradiction arising. Thus, $\mathrm{O}(r)$ is bounded. Since $\mathrm{O}(r)$ is closed, so $\mathrm{O}(r)$ is compact for $r \geqslant r_0$. This implies the solution set $V(\bar{y}, r) \subset \mathrm{O}(r)$ is nonempty and compact for any $r \geqslant r_0$.

(ii) Let $(x(r), u(r)) \in V(\bar{y}, r)$ with $r \geqslant r_0$, we have $(x(r), u(r)) \in \mathrm{O}(r) \subset \mathrm{O}(r_0)$. Hence $\{(x(r), u(r))\}$ is bounded. Then, there exists $r_0 < r_j \to +\infty$ and $(x^*, u^*) \in R^n \times R^m$ such that $(x(r_j), u(r_j)) \to (x^*, u^*)$.

Arbitrarily fix a $\bar{x} \in R^n$ such that $-\infty < \varphi(\bar{x}) < +\infty$. It is clear that

$$\bar{f}(x(r_j), u(r_j)) + r_j\sigma(u(r_j)) - \langle \bar{y}, u(r_j) \rangle \leqslant \bar{f}(\bar{x}, 0) = \varphi(\bar{x}). \tag{11}$$

Inequality (11), together with (5), gives

$$(r_j - \bar{r})\sigma(u(r_j)) \leqslant \varphi(\bar{x}) - \bar{m}, \quad \forall j.$$

Thus,

$$\sigma(u(r_j)) \leqslant \frac{\varphi(\bar{x}) - \bar{m}}{r_j - \bar{r}}.$$

This inequality, together with the lsc property of σ, gives

$$\sigma(u^*) \leqslant \liminf_{j \to +\infty} \sigma(u(r_j)) = 0.$$

Therefore, $u^* = 0$.

It follows from (11) that

$$\bar{f}(x(r_j), u(r_j)) - \langle \bar{y}, u(r_j) \rangle \leqslant \varphi(\bar{x}). \tag{12}$$

Taking lower limit in (12) as $j \to +\infty$ and use the lsc property of $\bar{f}(x,u)$, we get

$$\varphi(x^*) = f(x^*, 0) \leqslant \liminf_{j\to+\infty} \bar{f}(x(r_j), u(r_j)) - \langle \bar{y}, u(r_j) \rangle \leqslant \varphi(\bar{x}).$$

By the arbitrariness of $\bar{x} \in R^n$, we conclude that $x^* \in S$. So (ii) is proved. □

THEOREM 3.2 (Zero duality gap). *Consider the primal problem* (1) *and valley at* 0 *augmented Lagrangian dual problem* (3). *Assume that* φ *and its dualizing parameterization function* $\bar{f}(x,u)$ *satisfy the same conditions as in Theorem* 3.1. *Then*

(i) $$p(0) = \lim_{r\to+\infty} \bar{\psi}(\bar{y}, r).$$

(ii) *Zero duality gap holds*:

$$p(0) = \sup_{(y,r)\in R^m\times(0,+\infty)} \bar{\psi}(y, r).$$

Proof. (i) We need only to show that, for each sequence $\{r_j\} \to +\infty, \bar{\psi}(\bar{y}, r_j) \to p(0)$. From (i) of Theorem 3.1, suppose that there exists $(x(r_j), u(r_j)) \in V(\bar{y}, r_j)$ such that

$$\begin{aligned}\bar{\psi}(\bar{y}, r_j) &= \bar{f}(x(r_j), u(r_j)) - \langle \bar{y}, u(r_j) \rangle + r_j \sigma(u(r_j)) \\ &= \inf_{(x,u)\in R^n\times R^m} \{\bar{f}(x,u) + r_j\sigma(u) - \langle y, u \rangle\}\end{aligned}$$

Consequently,

$$\bar{\psi}(\bar{y}, r_j) \geqslant \bar{f}(x(r_j), u(r_j)) - \langle \bar{y}, u(r_j) \rangle. \tag{13}$$

From (ii) of Theorem 3.1, without loss of generality, we may assume that

$$(x(r_j), u(r_j)) \to (x^*, 0) \in S \times \{0\}.$$

This, combined with the lsc of $\bar{f}(\cdot,\cdot)$ and (13), yields

$$\begin{aligned}\liminf_{j\to+\infty} \bar{\psi}(\bar{y}, r_j) &\geqslant \liminf_{j\to+\infty} \bar{f}(x(r_j), u(r_j)) - \lim_{j\to+\infty} \langle \bar{y}, u(r_j) \rangle \\ &\geqslant \bar{f}(x^*, 0) = \varphi(x^*) = p(0).\end{aligned} \tag{14}$$

By Proposition 2.1 (ii),

$$\limsup_{j\to+\infty}\bar{\psi}(\bar{y},r_j)\leqslant p(0). \tag{15}$$

(14) together with (15) gives

$$\lim_{j\to+\infty}\bar{\psi}(\bar{y},r_j)=p(0).$$

So (i) has been proved.

(ii) From (i), we deduce

$$\sup_{(y,r)\in R^m\times(0,+\infty)}\bar{\psi}(y,r)\geqslant p(0). \tag{16}$$

This, together with Proposition 2.1 (ii), yields (ii). The proof is complete. □

Remark 3.1. If f and σ satisfy (5), σ has a valley at 0 in X, we can't get that σ is level-bounded on X. For example, let $(x,u)\in R^n\times R^m$,

$$\sigma(u)=\begin{cases}\|u\|_p^\gamma, & \text{if } \|u\|_p\leqslant 1,\\ 1, & \text{if } \|u\|_p>1,\end{cases}\qquad 0<\gamma<1.$$

$$f(x,u)=\|x\|^\alpha+\|u\|^\beta,\quad \alpha>0,\ \beta>1.$$

It is easy to check f and σ satisfy (5), σ has a valley at 0 in R^m, but σ isn't level-bounded on R^m.

It follows from Remark 2.1 and Remark 3.1 that Theorem 3.1 and Theorem 3.2 extend the corresponding results of [3, 8].

4. Exact Penalty Representation

In this section, we establish exact penalization representation results of the valley at 0 augmented Lagrangian scheme.

DEFINITION 4.1. (Exact penalty representation). Consider the problem (1). Let the valley at 0 augmented Lagrangian $\bar{l}$ be defined as in Definition 2.4. A vector $\bar{y}\in R^m$ is said to support an exact penalty representation for the problem (1) if there exists $\bar{r}>0$ such that

$$p(0)=\inf_{x\in R^n}\bar{l}(x,\bar{y},r),\quad \forall r\geqslant\bar{r} \tag{17}$$

and

$$\operatorname{argmin}_x\varphi(x)=\operatorname{argmin}_x\bar{l}(x,\bar{y},r),\quad \forall r\geqslant\bar{r}. \tag{18}$$

We have the following result.

LEMMA 4.1. *Suppose one of following conditions is satisfied*:

(i) $\bar{y}$ *supports an exact penalty representation for the problem* (1).
(ii) $p(0)$ *is finite and there exists* $\bar{r}' > 0$ *such that*

$$\bar{m}' = \inf\{\bar{f}(x,u) - \langle \bar{y}, u\rangle + \bar{r}'\sigma(u)\colon (x,u) \in R^n \times R^m\} > -\infty. \tag{19}$$

Then, we have

$$p(0) = \sup_{r \geqslant r_*} \inf_{x \in R^n} \bar{l}(x, \bar{y}, r) \tag{20}$$

where $r_* = \max\{\bar{r}, \bar{r}'\}$.

Proof. If the condition (i) is satisfied, it is easy to see (20) holds.

If the condition (ii) is satisfied, we prove (20) by contradiction. By the weak duality (4), there exists $\epsilon_0 > 0$ such that

$$p(0) > \sup_{r \geqslant r_*} \inf_{x \in R^n} \bar{l}(x, \bar{y}, r) + \epsilon_0.$$

Then there exist $x^k \in R^n$ and $u^k \in R^m$ such that

$$p(0) \geqslant \bar{f}(x^k, u^k) - \langle \bar{y}, u^k\rangle + r\sigma(u^k) + \epsilon_0, \quad \forall r \geqslant r_*. \tag{21}$$

Since σ has a valley at 0, $c_\delta = \inf_{\|x\| \geqslant \delta} \sigma(x) > 0$, for each $\delta > 0$. Denote $r_0 = \frac{p(0) - \epsilon_0 - \bar{m}'}{c_\delta} + r_*$. From (19) and (21), we have

$$\sigma(u^k) \leqslant \frac{p(0) - \epsilon_0 - \bar{m}'}{r - \bar{r}'} < c_\delta, \quad \forall r > r_0. \tag{22}$$

This implies $u^k \in \{u \in R^m\colon \|u\| \leqslant \delta\}$. Because $\bar{f}(x,u)$ is level-bounded in x locally uniform in u, it follows from (21) that $\{x^k\}$ is bounded. So $\{(x^k, u^k)\}$ is bounded. Assume, without loss of generality, that $(x^k, u^k) \to (\bar{x}, \bar{u})$. Let $r_0 < r^k \to +\infty$, (22) implies $\bar{u} = 0$. The lsc of f and σ combined with (21) yields $p(0) \geqslant p(0) + \epsilon_0$. This is a contradiction. So, (20) holds. □Corresponding to Theorem 11.61 in [8] and Theorem 4.1 in [3], we have:

THEOREM 4.1. *The following statements are true*:

(i) *If* $\bar{y}$ *supports an exact penalty representation for the problem* (1), *then there exist* $\bar{r} > 0$ *and a neighborhood* W *of* $0 \in R^m$ *such that*

$$p(u) \geqslant p(0) + \langle \bar{y}, u\rangle - \bar{r}\sigma(u), \quad \forall u \in W. \tag{23}$$

(ii) *The converse of* (i) *is true if the condition* (ii) *of Lemma* 4.1 *is satisfied.*

Proof. (i) Since $\bar{y}$ supports an exact penalty representation, there exists $\bar{r}>0$ such that (17) holds with $r=\bar{r}$, i.e.,

$$p(0)=\inf\{\bar{l}(x,\bar{y},\bar{r}):\ x\in R^n\}$$
$$=\inf\{\bar{f}(x,u)-\langle\bar{y},u\rangle+\bar{r}\sigma(u):\ (x,u)\in R^n\times R^m\}.$$

Consequently,

$$p(0)\leqslant\bar{f}(x,u)-\langle\bar{y},u\rangle+\bar{r}\sigma(u),\quad\forall x\in R^n,\quad u\in R^m,$$

implying

$$p(0)\leqslant p(u)-\langle\bar{y},u\rangle+\bar{r}\sigma(u),\quad\forall u\in R^m.$$

This proves (i).

(ii) In assuming (23) there is no loss of generality in taking W to be a ball $B_\delta=\{u\in R^m,\|u\|\leqslant\delta\},\delta>0$. First we prove that there exists $r^*>\bar{r}$, such that

$$p(u)\geqslant p(0)+\langle\bar{y},u\rangle-r^*\sigma(u)>0,\quad\forall u\in R^m\setminus B_\delta. \tag{24}$$

It follows from (i) of Theorem 3.2 that

$$p(0)=\lim_{r\to+\infty}\bar{\psi}(\bar{y},r).$$

Therefore, for $\forall\epsilon>0$, there exists $r_1>0$, such that

$$|\bar{\psi}(\bar{y},r_1)-p(0)|<\epsilon.$$

Hence

$$\inf_{(x,u)\in R^n\times R^m}\{\bar{f}(x,u)+r_1\sigma(u)-\langle\bar{y},u\rangle\}\geqslant p(0)-\epsilon.$$

Then

$$p(u)=\inf_{x\in R^n}\{\bar{f}(x,u)+r_1\sigma(u)-\langle\bar{y},u\rangle\}-r_1\sigma(u)+\langle\bar{y},u\rangle$$
$$\geqslant p(0)-\epsilon-r_1\sigma(u)+\langle\bar{y},u\rangle. \tag{25}$$

Since $u\in R^m\setminus B_\delta$ and σ has a valley at 0 in R^m, $c_\delta=\inf_{\|u\|\geqslant\delta}\sigma(u)>0$. Letting $r^*=\max\{\bar{r},\frac{\epsilon}{c_\delta}+r_1+1\}$, we have

$$\sigma(u)\geqslant c_\delta>\frac{\epsilon}{r^*-r_1}.$$

That is

$$\epsilon-(r^*-r_1)\sigma(u)<0.$$

This and (25) imply

$$p(u)+r^*\sigma(u)-\langle\bar{y},u\rangle\geqslant p(0).$$

Thus (24) holds. From (23), we know that the inequality in (24) holding for all $u\in R^m$. Therefore

$$\inf_{x\in R^n}\bar{l}(x,\bar{y},r^*)\geqslant p(0).$$

Then, using Lemma 4.1, we have

$$p(0)=\inf_{x\in R^n}\bar{l}(x,\bar{y},r),\quad \forall r\geqslant r^*. \tag{26}$$

Since $\sigma(0)=0$, (23) and (24) show that

$$\operatorname{arg}\min_{u\in R^m}\{p(u)+r\sigma(u)-\langle\bar{y},u\rangle\}=\{0\}. \tag{27}$$

Fix $r>r^*$, and let

$$\begin{aligned} g(x,u)&:=\bar{f}(x,u)+r\sigma(u)-\langle\bar{y},u\rangle,\\ h(u)&:=\inf_{x\in R^n}g(x,u),\\ k(x)&:=\inf_{u\in R^m}g(x,u). \end{aligned}$$

Obviously $h(u)=p(u)+r\sigma(u)-\langle\bar{y},u\rangle$ and $k(x)=\bar{l}(x,\bar{y},r)$. According to the interchange rule in 1.35 in [8], one has

$$\left.\begin{array}{r}\bar{u}\in\operatorname{arg}\min_u h(u)\\ \bar{x}\in\operatorname{arg}\min_x g(x,\bar{u})\end{array}\right\}\Longleftrightarrow\left\{\begin{array}{l}\bar{x}\in\operatorname{arg}\min_x k(x)\\ \bar{u}\in\operatorname{arg}\min_u g(\bar{x},u).\end{array}\right. \tag{28}$$

Now we prove that

$$\operatorname{arg}\min_x\varphi(x)=\operatorname{arg}\min_x\bar{l}(x,\bar{y},r),\quad\forall r\geqslant r^*. \tag{29}$$

For any $\bar{x}\in\operatorname{arg}\min_x\varphi(x)$, we have

$$\bar{x}\in\operatorname{arg}\min_x g(x,0).$$

It follow from (27) that

$$\operatorname{argmin}_u h(u) = \{0\}.$$

So

$$\bar{u} = 0 \in \operatorname{argmin}_u h(u).$$

Thus from (28),

$$\bar{x} \in \operatorname{argmin}_x k(x) = \operatorname{argmin}_x \bar{l}(x, \bar{y}, r), \quad \forall r \geqslant r^*.$$

We have

$$\operatorname{argmin}_x \varphi(x) \subset \operatorname{argmin}_x \bar{l}(x, \bar{y}, r), \quad \forall r \geqslant r^*. \tag{30}$$

On the other hand, for any $\bar{x} \in \operatorname{argmin}_x \bar{l}(x, \bar{y}, r) = \operatorname{argmin}_x k(x), \forall r \geqslant r^*$, we have

$$\inf_{u \in R^m} g(\bar{x}, u) = \bar{l}(\bar{x}, \bar{y}, r) = \inf_{x \in R^n} \bar{l}(x, \bar{y}, r) = \inf_{u \in R^m} h(u) = h(0) = g(\bar{x}, 0).$$

Thus, $\bar{u} = 0 \in \operatorname{argmin}_u g(\bar{x}, u)$. It follows from (28) that $\bar{x} \in \operatorname{argmin}_x g(x, 0) = \operatorname{argmin}_x \varphi(x)$. So

$$\operatorname{argmin}_x \bar{l}(x, \bar{y}, r) \subset \operatorname{argmin}_x \varphi(x), \quad \forall r \geqslant r^*. \tag{31}$$

Therefore, from (30) and (31) that (29) holds. (26) and (29) imply that $\bar{y}$ supports an exact penalty representation for the problem (1). □

Remark 4.1. In Theorem 4.1, we don't need to assume that there exist $\tau > 0$ and $N > 0$ such that $\sigma(u) \geqslant \tau \|u\|$ when $\|u\| \geqslant N$. But this assumption is necessary in the proof of [8]. In [8], σ was assumed to be proper, lsc, convex and $\operatorname{argmin}_y \sigma(y) = \{0\}$. As noted in [8], σ is level-coercive. This implies the existence of $\tau > 0$ and $N > 0$ satisfying $\sigma(u) \geqslant r\|u\|$ when $\|u\| \geqslant N$. In [3], the corresponding theorem (Theorem 5.1 in [3]) also need this assumption. Thus Theorem 4.1 improves the corresponding results in [3, 8].

For the special case where $\bar{y} = 0$ supports an exact penalty representation for the problem (1), we have the following result.

COROLLARY 4.1. *In the framework of the generalized augmented Lagrangian* $\bar{l}$ *defined in Definition* 2.3. *The following statements are true*:

(i) *If* $\bar{y} = 0$ *supports an exact penalty representation, then there exist* $\bar{r} > 0$ *and a neighborhood* W *of* $0 \in R^m$ *such that*

$$p(u) \geqslant p(0) - \bar{r}\sigma(u), \quad \forall u \in W.$$

(ii) *The converse of* (*i*) *is true if*

(a) $p(0)$ *is finite*;
(b) *there exist* $\bar{r} > 0$ *and* $m^* \in R$ *such that*

$$m^* = \inf\{\bar{f}(x,u) + \bar{r}\sigma(u)\colon (x,u) \in R^n \times R^m\}.$$

Corollary 4.1 improves the corresponding result in [3].

References

1. Auslender, A., Cominetti, R. and Haddou, M. (1997), Asymptotical analysis for penalty and barrier methods in convex and linear programming, *Mathematics of Operations Research*, 22, 43–62.
2. Bestsekas, D.P. (1982), *Constrained Optimization and Lagrangian Multiplier Method*, Academic Press, New York.
3. Huang, X.X. and Yang, X.Q. A unified augmented Lagrangian approach to duality and exact penalization, *Mathematics of Operation Research*, (accepted).
4. Ioffe, A. (1979), Necessary and sufficient conditions for a local minimum. 3: second-order conditions and augmented duality, *SIAM Journal Control and Optimization*, 17, 266–288.
5. Luo, Z.Q., Pang, J.S. and Ralph, D. (1996), *Mathematical Programs with Equilibrium Constraints*, Cambridge University Press, New York.
6. Rockafellar, R.T. (1974), Augmented Lagrange multiplier functions and duality in nonconvex programming, *SIAM Journal on Control and Optimization*, 12, 268–285.
7. Rockafellar, R.T. (1993), Lagrange multipliers and optimality, *SIAM Review*, 35, 183–238.
8. Rockafellar, R.T. and Wets, R.J.-B. (1998), *Variational Analysis*, Springer-Verlag, Berlin.
9. Rubinov, A.M., Huang, X.X. and Yang, X.Q. The zero duality gap property and lower semicontinuity of the perturbation function, *Mathematics of Operation Research*, (accepted).
10. Rubinov, A.M. and Yang, X.Q. (2003), *Lagrange-type Functions in Constrained Non-convex Optimization*, Kluwer Academic Publishers, Dordrecht.
11. Rubinov, A.M., Glover, B.M. and Yang, X.Q. (1999), Decreasing functions with applications to penalization, *SIAM Journal of Optimization*, 10(1), 289–313.
12. Yang, X.Q. and Huang, X.X. (2001), A nonlinear Lagrangian approach to constrained optimization problems, *SIAM Journal of Optimization*, 14, 1119–1144.