# Corrections and additions for the book "Perturbation Analysis of Optimization Problems" by J.F. Bonnans and A.Shapiro

*Version of June 06, 2006*

*Some typos in the book that we noticed are of trivial nature and do not need an explanation. There are, however, more subtle corrections that need to be made. There are also simple extensions and additions to the material presented in the book which are worthwhile to mention.*

## 0.1 Corrections

**Section 2.1.4**
In Section 2.1.4, $X$ and $X^*$ are assumed to be paired locally convex topological vector spaces equipped with respective compatible topologies. The closure operation in that section is taken with respect to these topologies. In particular, if $X$ is a nonreflexive Banach space and $X^*$ is its dual, then the standard pair of compatible topologies is the strong (norm) topology in $X$ and weak$^*$ topology in $X^*$ (see the discussion after Definition 2.26 on page 25). In that case the closure in the right hand sides of formulas (2.32) and (2.33), while in the space $X^*$, should be taken in the weak$^*$ topology. This may lead to a confusion since in the "Basic Notation" $\mathrm{cl}(\cdot)$ is defined with respect to the norm (strong) topology. If $X$ is a reflexive Banach space, then one can use strong topologies in $X$ and $X^*$.

**Page 52**
In formula (2.115) the inequality $h(\omega) \geq 0$ should be replaced by $h(\omega) \leq 0$. Also, in the following line $x(\omega) > 0$ should be replaced by $x(\omega) < 0$.

**Page 66, Proposition 2.90.**
For the given definition of the set $\mathcal{T}$ formula (2.171) is in error. Correct definition of the set $\mathcal{T}$ should be

$$\mathcal{T} := \{h \in X : G(x_0) + DG(x_0)h \in K\}. \tag{0.1.1}$$

In order to see why (2.171) is wrong let us consider the convex set $K \subset Y$. Let $y_0 \in K$ and $T_K(y_0)$ be the corresponding tangent cone. Then $K \in y_0 + T_K(y_0)$ and hence $\mathrm{dist}(d, T_K(y_0)) = 0$ for any $d \in K - y_0$. Formula (2.172) then follows, under Robinson's constraint qualification, by Stability Theorem 2.87. However, if $Y$ is infinite dimensional, it is not true in general that

$$\mathrm{dist}(y, K) = o(\|y - y_0\|), \quad y \in y_0 + T_K(y_0). \tag{0.1.2}$$

Consider, for example, space $Y := \ell_1$ and the set

$$K := \left\{ (y_n) \in \ell_1 : y_1 \geq n y_n^2,\ n = 2, ... \right\}. \tag{0.1.3}$$

This is a nonempty closed convex set and

$$T_K(0) = \{ (y_n) \in \ell_1 : y_1 \geq 0 \}. \tag{0.1.4}$$

We have here that for every $n$-th coordinate vector $e_n$ and $t > 0$, $te_n \in T_K(0)$, and for $n \geq (2t)^{-1}$,

$$\operatorname{dist}(te_n, K) = \inf_{y \in [0,t]} \{ ny^2 + t - y \} = t + (4n)^{-1} \geq t = \|te_n\|. \tag{0.1.5}$$

Consequently, (0.1.2) does not hold at $y_0 = 0$.

The assertion of Corollary 2.91 is correct. The fact that, under Robinson's constraint qualification, $T^i_\Phi(x_0) = T_\Phi(x_0)$ and equation (2.173) holds can be easily derived from Stability Theorem 2.87 and the relation that for any *fixed* $d \in T_K(y_0)$ and $y(t) = y_0 + td + o(t)$ it follows that $\operatorname{dist}(y(t), K) = o(t)$, $t > 0$.

**Page 82**. Formula (2.230) of Proposition 2.118 is correct. However, its proof is imprecise. First line of the proof of Proposition 2.118 should be replaced by the following.

It follows from (2.229), applied to $f^{**}$, that $x^* \in \partial f^{**}(x)$ iff

$$f^{**}(x) = \langle x^*, x \rangle - f^{***}(x^*). \tag{0.1.6}$$

Since, by the Fenchel-Moreau-Rockafellar Theorem 2.113, $f^{***} = f^*$ the above equation is equivalent to (2.231). Starting from equation (2.231) the proof can be completed as in the book.

**Page 215**. In the right hand side of formula (3.191), "$FG(x_0)w$" should be replaced by "$DF(x_0)w$".

**Page 242**. Line before equation (3.268), "for any $z \in Z$" should be replaced by "for any $z \in \mathcal{C}$".

**Page 270**. *Errors kindly indicated to us by Alexey F. Izmailov.*
In the first line of (4.21), read $D_x g_i(x_0, u_0)$ (and not $Dg_i(x_0, u_0)$).
In the second line of (4.22), read $\varepsilon d$ instead of $d$.

**Page 275**. Formula (4.39) should be

$$\operatorname{cl}\operatorname{conv}\left( \cup_{x \in \bar{\mathcal{S}}(u_0)} D_u f(x, u_0) \right). \tag{0.1.7}$$

Under the assumptions of theorem 4.13, the set $\cup_{x \in \bar{\mathcal{S}}(u_0)} D_u f(x, u_0)$ indeed is compact. Therefore, in case the space $U$ is *finite* dimensional, the convex hull of that set is also compact and hence is closed, and hence the topological closure in the above formula can be omitted.

**Page 302**. Line before equation (4.131), "$h_n := x_n - x_0$" should be replaced by "$h_n := \kappa_n^{-1}(x_n - x_0)$".

**Pages 404.** In equation (5.11): read $F(x_0, u_0)$ instead of $D_x F(x_0, u_0)$.

**Pages 460-461.** In the proof of theorem 5.60, some terms are missing in the optimality system at the end of page 460 and for the quadratic program at the top of page 461. The missing terms are $D^2 g_i(x_0, u_0)((h_1, u_1), (h_1, u_1))$ in the expansion of constraints, and $(D^2_{(x,u)x} G(x_0, u_0)(h_1, u_1))^* \lambda_1$ in the expansion of $D_x L$, as well as the corresponding $\lambda_1 \cdot D^2_{(x,u)x} G(x_0, u_0)((h_1, u_1), h_2)$ in the cost function of the quadratic program. *One should read, starting at the display at the bottom of page 460:*

$$
\begin{aligned}
& D^2_{(x,u)x} L(x_0, \lambda_0, u_0)(h_2, u_2) + D^3_{(x,u)(x,u)x} L(x_0, \lambda_0, u_0)((h_1, u_1), (h_1, u_1)) \\
& \qquad + (D^2_{(x,u)x} G(x_0, u_0)(h_1, u_1))^* \lambda_1 + DG(x_0, u_0)^* \lambda_2 = 0, \\
& D^2 g_i(x_0, u_0)((h_1, u_1), (h_1, u_1)) + Dg_i(x_0, u_0)(h_2, u_2) = 0, \quad i \in \{1, \dots, q\} \cup I^1_+, \\
& D^2 g_i(x_0, u_0)((h_1, u_1), (h_1, u_1)) + Dg_i(x_0, u_0)(h_2, u_2) \le 0, \quad i \in I_{u_1}(x_0, u_0, h_1) \setminus I^1_+, \\
& \lambda_{2i} Dg_i(x_0, u_0)(h_2, u_2) = 0, \quad i \in I_{u_1}(x_0, u_0, h_1) \setminus I^1_+.
\end{aligned}
$$

The above system has a unique solution, since it is the optimality system of the quadratic problem

$$
\begin{aligned}
\operatorname{Min}_{h_2} \quad & D^3_{(x,u)(x,u)(x,u)} L(x_0, \lambda_0, u_0)((h_1, u_1), (h_1, u_1), (h_2, u_2)) \\
& + D^2_{(x,u)(x,u)} L(x_0, \lambda, u_0)((h_2, u_2), (h_2, u_2)) \\
& + \lambda_1 \cdot D^2_{(x,u)x} G(x_0, u_0)((h_1, u_1), h_2) \\
\text{s.t.} \quad & Dg_i(x_0, u_0)(h_2, u_2) = 0, \quad i \in \{1, \dots, q\} \cup I^1_+, \\
& Dg_i(x_0, u_0)(h_2, u_2) \le 0, \quad i \in I_{u_1}(x_0, u_0, h_1) \setminus I^1_+,
\end{aligned}
$$

whose objective, etc.

**Page 476**. Second line after equation (5.171), "$C^\infty$-smooth and $\mathcal{G} \pitchfork_x \mathcal{W}_r$" should be replaced by "$C^\infty$-smooth and $\mathcal{G} \pitchfork \mathcal{W}_r$".

**Page 516**. First line.
The sentence: "If $x_0$ is a stationary point of $(P)$, then the first order growth condition holds at $x_0$ iff $C(x_0) = \{0\}$", should be replaced by: "If $x_0$ is a stationary point of $(P)$ and the extended MF constraint qualification holds, then the first

order growth condition holds at $x_0$ iff $C(x_0) = \{0\}$".

## 0.2 Additional material

**Page 301.** The result of Proposition 4.52 can be extended as follows.

**Proposition 0.1** *Suppose that the assumptions of proposition* 4.52 *hold and let* $(\hat{x}(u), \hat{\lambda}(u))$ *be a stationary point of* $(P_u)$ *such that* $\hat{x}(u) \to x_0$ *as* $u \to u_0$. *Then*

$$\|\hat{x}(u) - x_0\| = O\left(\|u - u_0\|^{1/2}\right). \tag{0.2.8}$$

**Proof.** It suffices to show that for any sequence $u_n$ converging to $u_0$ and $x_n := \hat{x}(u_n)$, it follows that $\kappa_n = O\left(\tau_n^{1/2}\right)$, where $\kappa_n := \|x_n - x_0\|$ and $\tau_n := \|u - u_0\|$. Denote also $\lambda_n := \hat{\lambda}(u_n)$. By passing to a subsequence if necessary we can assume that $\lambda_n$ converges to some $\lambda_0 \in \Lambda(x_0)$.

We have that $f(x_n, u_n) = L(x_n, \lambda_n, u_n)$ and $f(x_0, u_0) = L(x_0, \lambda_0, u_0)$. Because of Robinson's constraint qualification, there exist points $x_n'$ which are feasible for the unperturbed problem $(P_{u_0})$ and such that $\|x_n - x_n'\| = O(\tau_n)$. Consequently by the second order growth condition we obtain

$$\begin{aligned} & f(x_n, u_n) - f(x_0, u_0) \\ & \quad = f(x_n', u_0) - f(x_0, u_0) + f(x_n, u_0) - f(x_n', u_0) + f(x_n, u_n) - f(x_n, u_0) \\ & \quad \geq c\|x_n' - x_0\|^2 - c_0\tau_n \geq c\kappa_n^2 - c_1\tau_n - c_1\tau_n\kappa_n, \end{aligned}$$

where $c$, $c_0$ and $c_1$ are some constants with the constant $c$ being positive. Since

$$|L(x_0, \lambda_0, u_n) - L(x_0, \lambda_0, u_0)| = O(\tau_n),$$

it follows that

$$L(x_n, \lambda_n, u_n) - L(x_0, \lambda_0, u_n) \geq c\kappa_n^2 - c_2\tau_n(1 + \kappa_n) \tag{0.2.9}$$

for some constant $c_2$. We also have that

$$L(x_n, \lambda_n, u_n) - L(x_n, \lambda_0, u_n) = \langle \lambda_n - \lambda_0, G(x_n, u_n) \rangle$$

and $\langle \lambda_n, G(x_n, u_n) \rangle = 0$. Moreover, since $K$ is generalized polyhedral we can assume by passing to a subsequence if necessary that $\langle \lambda_0, G(x_n, u_n) \rangle = 0$. Together with (0.2.9) this implies

$$L(x_n, \lambda_0, u_n) - L(x_0, \lambda_0, u_n) \geq c\kappa_n^2 - c_2\tau_n(1 + \kappa_n). \tag{0.2.10}$$

Consider the mapping $F(z,u) := (D_x L(x,\lambda,u), -G(x,u))$, where $z := (x,\lambda)$, and let $z_n := (x_n,\lambda_n)$ and $z_0 := (x_0,\lambda_0)$. Since the multifunction $\Gamma(z)$ is monotone we obtain by the generalized equations (4.115) that

$$\langle z_n - z_0, F(z_n,u_n) - F(z_0,u_0)\rangle \le 0. \tag{0.2.11}$$

On the other hand

$$\begin{aligned}
&\langle z_n - z_0, F(z_n,u_n) - F(z_0,u_0)\rangle \\
&\quad = \langle x_n - x_0, D_x L(x_n,\lambda_n,u_n) - D_x L(x_0,\lambda_0,u_0)\rangle - \langle \lambda_n - \lambda_0, G(x_n,u_n) - G(x_0,u_0)\rangle \\
&\quad = \langle x_n - x_0, D_x L(x_n,\lambda_0,u_n) - D_x L(x_0,\lambda_0,u_0)\rangle \\
&\qquad -\langle \lambda_n - \lambda_0, G(x_n,u_n) - G(x_0,u_0) - D_x G(x_n,u_n)(x_n - x_0)\rangle \\
&\quad = \kappa_n^2 D_{xx}^2 L(x_0,\lambda_0,u_n)(h_n,h_n) + c_3\kappa_n\tau_n + o(\kappa_n^2) \\
&\qquad -\tfrac{1}{2}\langle \lambda_n - \lambda_0, D_{xx}^2 G(x_0,u_0)(h_n,h_n) + O(\tau_n)\rangle \\
&\quad = \kappa_n^2 D_{xx}^2 L(x_0,\lambda_0,u_n)(h_n,h_n) + c_3\kappa_n\tau_n + o(\kappa_n^2) + o(\tau_n)
\end{aligned}$$

for $h_n := (x_n - x_0)/\kappa_n$ and some constant $c_3$. Together with (0.2.9) and (0.2.11) this implies that

$$0 \ge \bar{c}\kappa_n^2 + c_4\kappa_n\tau_n + c_5\tau_n$$

for some constants $\bar{c}$, $c_4$ and $c_5$ with the constant $\bar{c}$ being positive. It follows then that $\kappa_n = O\left(\tau_n^{1/2}\right)$, which completes the proof. ∎