LIBERTAS MATHEMATICA, VOL. XIV (1994)

HIGHER ORDER NECESSARY CONDITIONS FOR OPTIMIZATION

N.H. Pavel, Y.K. Huang and J.K. Kim

1. Introduction. One introduces the n^{th} order tangential cones $T^n D(x)$ to a subset D of a normed (or general topological vector) space X (see (2.7) and (2.8)) following some ideas from Pavel [5] and Pavel-Ursescu [7]. One gives (in addition to classical necessary conditions (3.5)) new necessary conditions for extremum in terms of second order tangential cones. Examples and comments are also given (Example 3.1) to illustrate some possible applications. Related ideas can be found in Ledzewicz-Schaettler [2].

2. Higher order tangential cones. Let X be a normed space, of norm $\|\cdot\|$ and let D be a nonempty subset of X. Recall that a vector $v_1 \in X$ is said to be "tangent" to D at $x \in D$, if

$$\lim_{t \downarrow 0} \frac{1}{t} \operatorname{dist}(x + tv_1; D) = 0, \tag{2.1}$$

where dist $(z; D)$ stands for the distance from $z \in X$ to D.

It is well-known (and easy to check) that (2.1) is equivalent to (cf. [6], Ch. 3)

$$x + t(v_1 + \epsilon_1(t)) \in D, \quad \text{for some } \epsilon_1(t) \to 0 \text{ as } t \downarrow 0. \tag{2.2}$$

Denote by $TD(x)$ the set of all $v_1 \in X$ satisfying (2.1). This is a cone in the sense that $v_1 \in TD(x)$ implies $\lambda v_1 \in TD(x)$ for all $\lambda > 0$. $TD(x)$ is also known as the "contingent" cone (or tangent cone) to D at x. It goes back to Bouligand in the early 1930's [6]. Many useful versions of such tangential cones were introduced by C. Ursescu [8], F. Clarke and other researchers in optimization [6].

The extension of $TD(x)$ to Banach manifold was given by Motreanu and Pavel [3]. If D is a submanifold of class C^1, then $TD(x)$ is just the tangent space to D at $x \in D$. Such tangential cones play a crucial role in both optimization theory and in the flow-invariance of D with respect to a differential equation (details in [6]). It was shown by Pavel and Ursescu [7] that the "second order" tangential directions v_2 to D at $x \in D$ plays the key role in the flow-invariance of D with respect to a second order differential equation (so, in flight-mechanics). Precisely:

Definition 2.1. A vector $v_2 \in X$ is said to be a second order tangential direction to D at $x \in D$ *if there is* $v_1 \in X$ such that

$$\lim_{t \downarrow 0} \frac{1}{t^2} \operatorname{dist}\,(x + tv_1 + \frac{t^2}{2} v_2; D) = 0. \tag{2.3}$$

It follows that $v_1 \in TD(x)$ and (2.3) is equivalent to:

$$x + tv_1 + \frac{t^2}{2}(v_2 + \epsilon_2(t)) \in D, \quad \text{for some } \epsilon_2(t) \to 0 \text{ as } t \downarrow 0. \tag{2.4}$$

Remark 2.1. Given $v_2 \in T^2D(x)$, by definition there is v_1 (which must belong to $TD(x)$) such that (2.3) holds. Is such a v_1 unique? The answer is no as shown by Remark 2.2.

Denote by $T^2D(x)$ the set of all second order tangential directions v_2 to D at x. Clearly $T^2D(x)$ is also a cone. Indeed let $v_2 \in T^2D(x)$ and $\lambda > 0$. Replace t in (2.3) by $\sqrt{\lambda}t$. It follows

$$\lim \frac{1}{t^2} \operatorname{dist}\,(x + t(\sqrt{\lambda}v_1) + \frac{t^2}{2}(\lambda v_2); D) = 0. \tag{2.5}$$

Therefore $\lambda v_2 \in T^2D(x)$. Actually, in [7] only the formula (2.3) was studied (which goes back to Paves [5] in 1975). In this form, $T^2D(x)$ is introduced here for the first time. Equivalently,

$$T^2D(x) = \{v_2 \in X;\ \text{there exists } v_1 \in T^1D(x) \text{ such that (2.3) or (2.4) holds}\} \tag{2.6}$$

with $T^1D(x) = TD(x)$ (by convention).

The generalization (i.e., the definition of $T^nD(x)$ for $n = 2, 3, \ldots,$) is straightforward:

$$\begin{aligned} T^nD(x) = \{v_n \in X;\ &\text{there exist } v_i \in T^iD(x), \\ &i = 1, 2, \ldots, n-1, \text{ such that (2.7) holds}\}, \end{aligned}$$

where (2.7) is given below

$$\lim_{t\downarrow 0} \frac{1}{t^n} \operatorname{dist}\left(x + tv_1 + \frac{t^2}{2!}v_2 + \cdots + \frac{t^n}{n!}v_n; D\right) = 0,\ n = 1, 2, \ldots. \tag{2.7}$$

Equivalently, (2.7) means

$$x + tv_1 + \cdots + \frac{t^{n-1}}{(n-1)!}v_{n-1} + \frac{t^n}{n!}(v_n + \epsilon_n(t)) \in D \text{ for some } \epsilon_n(t) \to 0 \text{ as } t \downarrow 0. \tag{2.8}$$

$T^nD(x)$ is a cone as, if $v_n \in T^nD(x)$, there are $v_1, \ldots, v_{n-1}$ satisfying (2.7). Then $\lambda v_n, \lambda^{\frac{1}{n}}v_1, \lambda^{\frac{2}{n}}v_2, \ldots, \lambda^{\frac{n-1}{n}}v_{n-1}$ satisfy (2.7) (with $t\lambda^{\frac{1}{n}}$ in place of t), so $\lambda v_n \in T^nD(x)$.

Our $T^2D(x)$ here is somehow related to the second order tangent cone $TC^{(2)}(D; x, h) \subset X \times \mathbb{R}_+$ introduced by Ledzewicz and Schaettler [2]. Namely, $TC^{(2)}(D; x, h)$ is the set of all pairs $(v, \gamma) \in X \times \mathbb{R}_+$ (the second order tangent directions to D at x in the direction of h) such that

$$x + t\sqrt{\gamma}h + t^2(v + \epsilon(t)) \in D \tag{2.9}$$

for some $\epsilon(t) \to 0$ (as $t \downarrow 0$). Clearly, such an element $v \in T^2D(x)$. Indeed, (2.9) can be rewritten as:

$$x + t(\sqrt{\gamma}h) + \frac{t^2}{2}(2v + 2\epsilon(t)) \in D \tag{2.9$'$}$$

which is just (2.4) with $v_1 = h\sqrt{\gamma}$ and $v_2 = 2v$. This means that $2v \in T^2D(x)$ and therefore $v \in T^2D(x)$ (as $T^2D(x)$ is a "positive" cone).

As an example, let us find $T^3D(x)$ in the case of $D = S_r = \{x \in X, |x| = r\} \equiv S$ and $X = H$—a Hilbert space of inner product $\langle\ \rangle$. First of all, it is known [6, p. 118] that

$$TS(x) = \{y \in X, \langle x, y\rangle = 0\}, \quad |x| = r, \tag{2.10}$$

i.e.,

$$\lim_{t\downarrow 0} \frac{1}{t}\operatorname{dist}(x + ty; S) = 0, \quad \text{if and only if } \langle x, y\rangle = 0 \tag{2.11}$$

with $|x| = r$. Moreover

$$\lim_{t\downarrow 0} \frac{1}{t^2}\operatorname{dist}(x + tv_1 + \frac{t^2}{2}v_2; S) = 0 \tag{2.12}$$

if and only if

$$\langle x, v_1\rangle = 0, \qquad |v_1|^2 + \langle x, v_2\rangle = 0. \tag{2.13}$$

The proof of the equivalence (2.12) $\Longleftrightarrow$ (2.13) is similar to the proof of Lemma 2.1 below (which is new).

Lemma 2.1. *Let $x \in S$ (i.e., $|x| = r$) and $v_1, v_2, v_3 \in H$. Then the following two conditions are equivalent:*

$$\lim_{t\downarrow} \frac{1}{t^3}\operatorname{dist}(x + tv_1 + \frac{t^2}{2}v_2 + \frac{t^3}{3!}v_3; S) = 0 \tag{2.14}$$

$$\langle x, v_1\rangle = 0, \quad |v_1|^2 + \langle x, v_2\rangle = 0, \quad \frac{1}{3}\langle x, v_3\rangle + \langle v_1, v_2\rangle = 0. \tag{2.15}$$

Proof. We know that (2.14) is equivalent to:

$$|x + tv_1 + \frac{t^2}{2}v_2 + \frac{t^3}{3!}(v_3 + \epsilon_3(t))|^2 = r^2 \tag{2.16}$$

for some $\epsilon_3(t) \to 0$ as $t \downarrow 0$ (by (2.8) with $n = 3$). Inasmuch as $|x|^2 = r^2$, it is easy to check that (2.16) implies (2.15). Vice versa, we prove that (2.15) implies (2.16), i.e., if (2.15) holds, then we can find $\epsilon_3(t)$ satisfying (2.16). Indeed, choose $\epsilon_3(t)$ given by:

$$x + tv_1 + \frac{t^2}{2}v_2 + \frac{t^3}{3!}v_3 + \frac{t^3}{3!}\epsilon_3(t) = \frac{a(t)}{|a(t)|}r \tag{2.17}$$

with $a(t) = x + tv_1 + \frac{t^2}{2}v_2 + \frac{t^3}{3!}v_3 \in X$. Then (2.16) holds. It remains to check that $\epsilon_3(t) \to 0$ as $t \downarrow 0$. Indeed:

$$\frac{t^3}{3!}|\epsilon_3(t)| = \frac{|a(t)(r - |a(t)|)|}{|a(t)|} = |r - |a(t)|\,| = \frac{|r^2 - |a(t)|^2|}{r + |a(t)|} = t^3 b(t), \tag{2.18}$$

where (2.15) was used. Here $b(t) \to 0$ as $t \downarrow 0$. This completes the proof. Therefore:

$$T^2S(x) = \{v_2 \in X; \exists v_1 \in X \text{ such that } \langle v_1, x\rangle = 0, |v_1|^2 + \langle x, v_2\rangle = 0\}. \tag{2.19}$$

Remark 2.2. Clearly $v_2 = -x \in T^2S(x)$. Indeed, take an $x^\perp$ with $\langle x^\perp, x\rangle = 0$ and $|x^\perp| = |x|$. Then with $v_1 = \pm x^\perp$ (2.13) holds. Moreover, it follows that for a given v_2 the element v_1 satisfying (2.13) is not uniquely determined.

Finally:

$$T^3S(x) = \{v_3 \in X; \exists v_1, v_2 \in X \text{ such that (2.15) holds}\}. \tag{2.20}$$

3. Optimum principles in terms of second order tangential directions.

First we recall the following general results of Pavel-Ursescu [7].

Let $g\colon X \to Y$ (Y—a finite dimensional space of inner product $\langle\ \rangle$) be a continuous function (in a neighborhood of $x_0 \in X$). Set

$$D_g = g^{-1}(0) = \{x \in X, g(x) = 0\}. \tag{3.1}$$

Lemma 3.1. [7] (1) *Suppose g is continuous in a neighborhood of x_0 and Frechet differentiable at x_0 with $\dot{g}(x_0)$ (the Frechet derivative at x_0) onto. Then*

$$TD_g(x_0) = \ker \dot{g}(x_0) = \{y \in X, \langle \dot{g}(x_0), y \rangle = 0\}. \tag{3.2}$$

(2) *If in addition to the above hypotheses, $\ddot{g}(x_0)$ (the second Frechet derivative) exists, then (2.3) is equivalent to:*

$$g(x_0) = 0, \quad \langle \dot{g}(x_0), v_1 \rangle = 0, \quad \langle \ddot{g}(x_0)v_1, v_1 \rangle + \langle \dot{g}(x_0), v_2 \rangle = 0, \tag{3.3}$$

so $T^2 D_g(x_0) = \{v_2 \in X; \exists v_1$ such that (3.3) holds$\}$.

The main result of this section is Theorem 3.1 below (optimum principle in terms of the first and second order tangential directions).

Theorem 3.1. *Let G and g be two functionals from X into $\mathbb{R}$ and let x_0 be a minimum point of G on $g^{-1}(0)$, i.e.,*

$$\inf\{G(x); g(x) = 0\} = G(x_0). \tag{3.4}$$

Suppose G is of class C^2 in a neighborhood of x_0 (G—not constant) and g satisfies the conditions in Lemma 3.1 above. Then (necessarily)

$$\langle \dot{G}(x_0), v_1 \rangle = 0, \quad \forall v_1 \in X \text{ with } \langle \dot{g}(x_0), v_1 \rangle = 0, \tag{3.5}$$

(i.e., $\dot{G}(x_0) = \lambda \dot{g}(x_0)$ for some $\lambda \neq 0$),

$$\langle \dot{G}(x_0), v_2 \rangle + \langle \ddot{G}(x_0)v_1, v_1 \rangle \geq 0, \tag{3.6}$$

for all (v_1, v_2) satisfying

$$\langle \dot{g}(x_0), v_1 \rangle = 0, \quad \langle \dot{g}(x_0), v_2 \rangle + \langle \ddot{g}(x_0)v_1, v_1 \rangle = 0. \tag*{(3.6)$'$}$$

Proof. Let v_1, v_2 satisfy (3.3). Then there are $r_i(t) \to 0$ as $t \downarrow 0, i = 1, 2$ such that

$$g(x_0 + t(v_1 + r_1(t))) = 0, \quad g(x_0 + tv_1 + \frac{t^2}{2}(v_2 + r_2(t))) = 0 \tag{3.7}$$

for all $t > 0$. Therefore

$$G(x_0 + t(v_1 + r_1(t))) - G(x_0) \geq 0 \quad \text{for all } t \in \mathbb{R} \tag{3.8}$$

which implies (3.5), and

$$G(x_0 + tv_1 + \frac{t^2}{2}(v_2 + r_2(t))) - G(x_0) \geq 0, \quad t \geq 0. \tag{3.9}$$

In view of Taylor's formula and of $\langle \dot{G}(x_0), v_1 \rangle = 0$, (3.9) yields:

$$\frac{t^2}{2}\langle \dot{G}(x_0), v_2 \rangle + \frac{t^2}{2}\langle \ddot{G}(x_0)v_1, v_1 \rangle + t^2 b(t) \geq 0 \tag{3.10}$$

with $b(t) \to 0$ as $t \to 0$ (which implies (3.6)). This completes the proof.

Remark 3.1. In [2] the condition that $\dot{g}(x_0)$ is onto is not required, but g is supposed to be three times differentiable. Condition (3.5) is the Lagrange multipliers rule, while (3.6) is new. The points x_0 satisfying (3.5) are "candidates" for infimum points in (3.4). Condition (3.6) will help to eliminate some of these "candidates". Precisely, a point x_0 satisfying (3.5) but not satisfying (3.6) is not a point where the infimum in (3.4) is assumed. Such a candidate x_0 is eliminated by (3.6). However, the necessary conditions (3.5)+(3.6) are not sufficient (for x_0 to be an infimum point as in (3.4)). The following simple example will illustrate the above situations.

Example 3.1. Take $X = \mathbb{R}^2$, $G(x) = (x_1^3 + x_2^3)/6$ and $g(x) = \frac{1}{2}(x_1^2 + x_2^2 - r^2)$ with $x = (x_1, x_2)$ and $r > 0$. Let us apply Theorem 3.1 for finding $x_0 = (x_{01}, x_{02})$ such that

$$\inf\{(x_1^3 + x_2^3)/6; \text{ subject to } x_1^2 + x_2^2 = r^2\} = \frac{1}{6}(x_{01}^3 + x_{02}^3). \tag{3.11}$$

In this case $\dot{G}(x) = \frac{1}{2}(x_1^2, x_2^2) =$ gradient of G at $x = (x_1, x_2)$ and $\ddot{G}(x) = \begin{pmatrix} x_1 & 0 \\ 0 & x_2 \end{pmatrix}$ = Hessian matrix. Set $x^\perp = (x_2, -x_1)$. Clearly $\dot{g}(x) = (x_1, x_2) = x$ and $\ddot{g}(x) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

$$\ker \dot{g}(x) = \{v_1; v_1 = \alpha x^\perp = (\alpha x_2, -\alpha x_1), \alpha \in \mathbb{R}\} = TS_r(x) \tag{3.12}$$

(the tangent space to $S_r = \{x \in \mathbb{R}^2, x_1^2 + x_2^2 = |x|^2 = r^2\}$ at $x \in S_r$, see (2.10)). Therefore, (3.5) yields (with $x_0 = (x_{01}, x_{02})$)

$$\langle \dot{G}(x_0), v_1 \rangle = 0 = x_{01}^2 x_{02} - x_{02}^2 x_{01} = x_{01} x_{02}(x_{01} - x_{02}). \tag{3.13}$$

Consequently in order for x_0 (with $|x_0| = r$) to be a solution of (3.11), we must have either

$$x_{01} x_{02} = 0 \qquad \text{or} \qquad x_{01} = x_{02} \tag{3.14}$$

which yields respectively

$$x_{01} = \epsilon r, \quad x_{02} = 0; \quad \epsilon = \pm 1 \quad (x_{01} = 0, \quad x_{02} = \epsilon r) \tag{3.15}$$

or

$$x_{01} = x_{02} = \epsilon \frac{r}{\sqrt{2}}, \quad \epsilon = \pm 1. \tag{3.16}$$

A pair (v_1, v_2) satisfying (3.6)′, i.e.,

$$\langle x_0, v_2 \rangle + |v_1|^2 = 0 \tag{3.17}$$

is given by $v_2 = -x_0$, $v_1 = x_0^\perp = (x_{02}, -x_{01})$ so (3.6) yields (with $x_{01}^2 + x_{02}^2 = r^2$)

$$-\frac{1}{2} x_{01}^3 - \frac{1}{2} x_{02}^3 + x_{01} x_{02}^2 + x_{01}^2 x_{02} \geq 0$$

which can be written as

$$(x_{01} + x_{02})(r^2 - 3 x_{01} x_{02}) \leq 0. \tag{3.18}$$

The set of all candidates $x_0 = (x_{01}, x_{02})$ for infimum points (as in (3.11)) are given by (3.15)+(3.16). Clearly (3.18) requires $\epsilon = -1$ in (3.15) and $\epsilon = 1$ in (3.16) (so those x_0 corresponding to $\epsilon = 1$ in (3.15) and $\epsilon = -1$ in (3.16) are eliminated by (3.18)). Therefore these remain

$$x_0 = (-r, 0) \quad \text{and} \quad x_0 = (0, -r) \quad \text{and} \quad x_0 = (\frac{r}{\sqrt{2}}, \frac{r}{\sqrt{2}}). \tag{3.19}$$

Finally, it is now obvious that (3.20) has two solutions

$$x_0 = (-r, 0) \quad \text{and} \quad x_0 = (0, -r) \quad \text{with}$$

$$\inf\{G(x), |x| = r\} = -\frac{r^3}{6} = G(-r, 0) = G(0, -r). \tag{3.20}$$

We are now in a position to point out that (3.5)+(3.6) are not sufficient for x_0 to be a solution of (3.11). Indeed, in this case (3.5) is just (3.13). Let us examine (3.6). A pair (v_1, v_2) satisfying (3.6)$'$ is given by

$$v_1 = \alpha x_0^\perp = (\alpha x_{02}, -\alpha x_{01}), \quad v_2 = -\alpha^2 x_0 + b x_0^\perp, \quad b \in \mathbb{R} \tag{3.21}$$

so (3.6) becomes:

$$-\alpha^2(x_{01}^3 + x_{02}^3) + b x_{01} x_{02}(x_{01} - x_{02}) + 2\alpha^2 x_{01} x_{02}(x_{01} + x_{02}) \geq 0 \tag{3.22}$$

for all $\alpha, b \in \mathbb{R}$. Or, in view of (3.22), we see that (3.23) (i.e., (3.6)) reduces to (3.18). Therefore $x_0 = (\frac{r}{\sqrt{2}}, \frac{r}{\sqrt{2}})$ satisfies (3.5) and (3.6) and however it is not a solution of (3.11).

Remark 3.2. It follows from (3.22) that

$$T^2 S_r(x) = \{v_2 = -\alpha^2 x + b x^\perp; \alpha, b \in \mathbb{R}\} \tag{3.23}$$

with $x = (x_1, x_2)$ and $x^\perp = (x_2, -x_1)$ (see also (2.12) and (2.13)).

An example of applications of $TD(x)$ (for some special D) to optimal control of PDE was recently given by Barbu and Pavel [1].

Remark 3.3. From (3.8) it also follows that (in addition to (3.5) and (3.6)) another necessary condition for x_0 to be a point of infimum (as in (3.4)) is given by

$$\langle \ddot{G}(x_0)v_1, v_1\rangle \geq 0, \quad \forall v_1 \text{ with } \langle \dot{g}(x_0), v_1\rangle = 0. \tag{3.24}$$

Or, in Example 3.1, $v_1 = (\alpha x_2, -\alpha x_1)$ with $\alpha \in \mathbb{R}$ (see (3.12)). Therefore, (3.24) becomes:

$$\alpha^2(x_1 x_2^2 + x_2 x_1^2) \geq 0 \text{ with } x_1 = x_{01}, x_2 = x_{02},$$

i.e.,

$$x_{01}x_{02}(x_{01} + x_{02}) \geq 0 \tag{3.25}$$

which will also eliminate the candidate $x_0 = (-\frac{r}{\sqrt{2}}, -\frac{r}{\sqrt{2}})$ in (3.16).

References

[1] V. Barbu and N. H. Pavel, *Optimal control of thermal conductivity,* to appear.

[2] U. Ledzewicz and H. Schaettler, *Second order conditions for extremum problems with nonregular equality constraints,* to appear in J. O. T. A.

[3] D. Motreanu and N. H. Pavel, *Quasi-tangent vectors in flow-invariance and optimization problems on Banach manifolds,* J. M. A. A., 88 (1982), 116–132.

[4] D. Motreanu and N. H. Pavel, *Flow-invariance for second order differential equations on Banach manifolds and orbital motions,* Boll. Un. Mat. Ital., (7) 1-B (1987), 943–964.

[5] N. H. Pavel, *Second order differential equations on closed subsets of a Banach space,* Boll. Un. Mat. Ital., 12 (1975), 348–353.

[6] N. H. Pavel, "Differential Equations, Flow-Invariance and Applications," Pitman Res. Notes in Math., vol. 113, 1984.

[7] N. H. Pavel and C. Ursescu, *Flow-invariant sets for autonomous second order differential equations and applications in mechanics,* Nonlinear Anal., 6 (1982), 35–74.

[8] C. Ursescu, *Tangent sets and differentiable functions,* Banach Center Publications, 1 (1974), 151–155.