Analele ştiinţifice ale Universităţii „Al. I. Cuza" din Iaşi
Tomul XXXI, s. I a, Matematică, 1985, f. 3

REGULARIZATION OF SADDLE FUNCTIONS AND THE YOSIDA APPROXIMATION OF MONOTONE OPERATORS

BY

E. KRAUSS and D. TIBA

1. Introduction. Let E and F be two real reflexive Banach spaces, with dual spaces E^*, respectively F^*. We can assume all these spaces to be locally uniformly convex (cf. T r o y a n s k i [18]). By a *saddle function* $K : E\times F\to\bar{R}$ we shall always mean a concave-convex bivariate function with values in the extended real line $\bar{R}=R\cup\{\pm\infty\}$.

In [17] the second author introduced the *regularization* K^λ_λ of a closed proper saddle function by

$$(0)\qquad K^\lambda_\lambda : E\times F\to R,\ K^\lambda_\lambda(x, y) := \sup_{U\in E}\ \inf_{V\in F}\left\{\frac{\| y-v \|^2}{2\lambda}-\frac{\|x-u\|^2}{2\lambda}+K(u, v)\right\}.$$

Let us mention that in the case $K(x, y)=p(y)$, K^λ_λ is just the Moreau-Yosida approximation p_λ of the convex function p. It can be shown that many results for this Moreau-Yosida approximation remain true in the case of saddle functions. Moreover, the subdifferential of K^λ_λ can be expressed in terms of the Yosida approximation of the maximal monotone operator R_K which was associated to a saddle function by R o c k a f e l l a r [15], [16].

In [1], [2] A t t o u c h and W e t s showed that the construction (0) is efficient even if K is not concave-convex. Namely, many of the mentioned regularity results for K^λ_λ can be extended to this case. In this context we note that in [1], [2] the regularization K^μ_λ depends on two positive parameters λ and μ, a fact which gives a deeper insight into the structure of the family of regularizations. We think, however, that the dependence of only one parameter is better adapted to the special nature of our considerations.

The aim of this paper is twofold. In Section 4 we sharpen some of the regularity results for K^λ_λ. In Section 5 we rely on the monotone operator

$$(1)\qquad T_K\subseteq E\times E^*\ ;\ f\in T_Kx := [-f, f]\in\partial K(x, x),$$

which can be introduced for each proper saddle function $K : E\times E\to\bar{R}$. In [9], [10] it was shown that, in this manner, each maximal monotone operator $A\subseteq E\times E^*$ can be obtained from a saddle function (compare Section 3 of the present paper). In Section 5 we shall demonstrate that the monotone operator $T_{K^\lambda_\lambda}$ associated to K^λ_λ is just the Yosida approximation of T_K.

2. **Notations.** We suppose familiarity with the definitions and basic results of the theory of closed saddle function as developed by Rockafellar. The fundamentals of this theory can be found in [4], [8], [12], [14]. Here we recall mainly those notions which are not uniformly defined in the literature.

The *closure* of a convex function $p : F \to \overline{R}$, cl p, is the greatest lowe semicontinuous convex function which is majorized by p and does not assume the value $-\infty$. If there is no such function, we set $\mathrm{cl}\, p \equiv -\infty$. It holds $\mathrm{cl}\, p(x) = \sup\limits_{f \in F^*} \inf\limits_{v \in F} \{f(x-v)+p(v)\}$. The *closure* of a *concave* function $q : E \to \overline{R}$ is defined by $\mathrm{cl}\, q(x) := \inf\limits_{f \in E^*} \sup\limits_{v \in E} \{f(x-v)+q(v)\}$.

The *convex closure* $\mathrm{cl}_2 K$ of a saddle function K is obtained by closing $K(x, \cdot)$ as a convex function (for arbitrary $x \in E$), and similarly the *concave closure* $\mathrm{cl}_1 K$ of K. The saddle function $K : E \times F \to \overline{R}$ is called *proper* if dom $K = \mathrm{dom}_1 K \times \mathrm{dom}_2 K \neq \emptyset$, with

$$\mathrm{dom}_1\, K = \{x \in E\ ;\ \mathrm{cl}_2 K(x, w) > -\infty \text{ for all } w \in F\} \text{ and}$$

$$\mathrm{dom}_2\, K = \{y \in F\ ;\ \mathrm{cl}_1 K(v, y) < +\infty \text{ for all } v \in E\}\,.$$

The *subdifferential* of K is the mapping $\partial K : E \times F \to E^* \times F^*$,

$$[f, g] \in \partial K(x, y) := \begin{cases} K(x, y) + f(v-x) \geq K(v, y) & \text{for all } v \in E, \\ K(x, y) + g(w-y) \leq K(x, w) & \text{for all } w \in F. \end{cases}$$

3. **Saddle Functions and Monotone Operators.** In [15], [16] Rockafellar introduced, for a given closed proper saddle function $K : E \times F \to \overline{R}$, the operator

$$R : E \times F \to E^* \times F^*\ ;\ [f, g] \in R(x, y) := [-f, g] \in \partial K(x, y) \tag{2}$$

and proved its maximal monotonicity.

To a closed proper saddle function $K : E \times E \to \overline{R}$, the first author (cf. [9], [10], [11]) associated another monotone operator $T_K \subseteq E \times E^*$, which is defined via the relation (1). The importance of this construction is underlined by the following two theorems.

Theorem 1. *$T_K \subseteq E \times E^*$ is a monotone operator. It is maximal monotone if and only if K satisfies*

$$\mathrm{cl}_2 K\left(v, \frac{v+w}{2}\right) \leq \mathrm{cl}_1 K\left(\frac{v+w}{2}, w\right) \text{ for any } v, w \in E. \tag{3}$$

Note that (3) is satisfied for each saddle function with

$$\mathrm{cl}_2 K(x, x) \leq 0 \leq \mathrm{cl}_1 K(x, x) \text{ for all } x \in E. \tag{4}$$

The condition (4) holds, among others, for saddle functions $K : E \times E \to \overline{R}$ which are *skew-symmetric* in the sense

$$\mathrm{cl}_2 K(x, y) = -\mathrm{cl}_1 K(y, x) \text{ for all } x, y \in E. \tag{5}$$

Example. Let $p : E \to]-\infty, +\infty]$ be a lower semicontinuous, proper, convex function. Then

$$L(x, y) = \begin{cases} p(y) - p(x) & \text{for } x, y \in \mathrm{dom}\, p, \\ +\infty & \text{for } x \in \mathrm{dom}\, p,\ y \notin \mathrm{dom}\, p, \\ -\infty & \text{for } x \notin \mathrm{dom}\, p, \end{cases}$$

is a closed, skew-symmetric saddle function and $T_L = \partial p$, the subdifferential of p.

We have the following representation of maximal monotone operators:

Theorem 2. *Let $A \subseteq E \times E^*$ be maximal monotone. Then*

(i) *there exists a proper closed skew-symmetric saddle function $K : E \times E \to \overline{R}$ with $A = T_K$,*

(ii) *it holds $A = T_L$, where the saddle function $L(x, y) = \overline{\mathrm{co}}(Ay, y-x)$ is closed and proper and satisfies the inequality* (4).

More precisely, L is obtained by taking the Γ-regularization of the function $y \to (Ay, y-x) = \inf\limits_{f \in Ay}(f, y-x)$, with $\inf \emptyset = +\infty$, for each $x \in E$. In other words, we have

$$L(x, y) = \sup_{g \in E^*} \inf_{[v, f] \in A} \{(g, y-v) + (f, v-x)\}.$$

Note, that L is concave-convex. Let us emphazise that the skew-symmetric saddle function in Theorem 2(*i*) is obtained in a nonconstructive manner, but one has the estimation

$$-\inf_{f \in Ax}(f, x-y) \leq K(x, y) \leq \inf_{g \in Ay}(g, y-x) \text{ for all } x, y \in E. \tag{6}$$

4. **Regularization of Saddle Functions.** In this section we sharpen some results of A t t o u c h-W e t s [1], [2] and T i b a [17] concerning the properties of the saddle function K_λ^λ. We start by recalling the following.

Theorem 3. *Let $K : E \times F \to \overline{R}$ be closed and proper. Then, for each $\lambda > 0$, $K_\lambda^\lambda : E \times F \to R$ is a finite saddle function with*

$$\begin{aligned} K_\lambda^\lambda(x, y) &= \min_{v \in F} \max_{u \in E} \{\| y-v \|^2/2\lambda - \| x-u \|^2/2\lambda + K(u, v)\} \\ &= \max_{u \in E} \min_{v \in F} \{\| y-v \|^2/2\lambda - \| x-u \|^2/2\lambda + K(u, v)\}, \end{aligned} \tag{7}$$

$$K_\lambda^\lambda = (\mathrm{cl}_1 K)_\lambda^\lambda = (\mathrm{cl}_2 K)_\lambda^\lambda = (K^\lambda)_\lambda = (K_\lambda)^\lambda, \tag{8}$$

$$\mathrm{cl}_2 K(x, y) \leq \varliminf_{\lambda \to 0} K_\lambda^\lambda(x, y) \leq \varlimsup_{\lambda \to 0} K_\lambda^\lambda(x, y) \leq \mathrm{cl}_1 K(x, y) \tag{9}$$

for all $x \in E$, $y \in F$.

Above, we have denoted

$$K^\lambda(x, y) = \sup_{u \in E} \{-\| x-u \|^2/2\lambda + K(u, y)\}, \tag{10}$$

$$K_\lambda(x, y) = \inf_{v \in F} \{\| y-v \|^2/2\lambda + K(x, v)\}. \tag{11}$$

Relation (9) is a corrected and improved version of the similar results from [17], and we give an outline of the argument for the first inequality. It is enough to consider $x \in \mathrm{dom}_1 K$. Then

$$K_\lambda^\lambda(x, y) \geq K_\lambda(x, y) = \inf_{v \in F} \{\| y - v \|^2/2\lambda + cl_2 K(x, v)\}.$$

Since $\mathrm{cl}_2 K(x,.)$ is a lower semicontinuous, proper, convex function, we have $\lim_{\lambda \to 0} K_\lambda(x, y) = \mathrm{cl}_2 K(x, y)$, and the argument is finished.

Obviously, inequality (9) has the consequence

$$\lim_{\lambda \to 0} K_\lambda^\lambda(x, y) = K(x, y) \tag{12}$$

for all $x \in E$, $y \in F$ such that $\mathrm{cl}_2 K(x, y) = \mathrm{cl}_1 K(x, y)$. This is true whenever $x \in \mathrm{rad\ dom}_1 K$ or $y \in \mathrm{rad\ dom}_2 K$ (cf. [4], Theorem 3.1, [11], Lemma 2, or [16], Theorem 5). Here rad C denotes the set of all $u \in C$ such that every ray emanating from $u \in C$ contains points of C besides u.

For any saddle function $K(x, y) = p(y)$, with $p : F \to R \cup \{+\infty\}$ convex, closed and proper, it holds $\mathrm{cl}_2 K = \mathrm{cl}_1 K$. In this case (12) yields the well-known convergence result $\lim_{\lambda \to 0} p_\lambda(x) = p(x)$, $x \in E$.

In the next theorem we investigate the Fréchet differential of the regularization K_λ^λ.

Theorem 4. *Let $K : E \times F \to \overline{R}$ be closed and proper. Then $K_\lambda^\lambda : E \times F \to R$, $\lambda > 0$, is Fréchet differentiable on $E \times F$, the Fréchet differential being just the subdifferential $\partial K_\lambda^\lambda$. The mapping $[x, y] \to \partial K_\lambda^\lambda(x, y)$ is continuous and bounded on bounded sets. It is Lipschitz continuous if E and F are Hilbert spaces. Moreover, we have*

$$[f, g] = \partial K_\lambda^\lambda(x, y) \Leftrightarrow [-f, g] = (R_K)_\lambda(x, y), \tag{13}$$

where $(R_K)_\lambda$ is the Yosida approximation of the operator R_k

Remark. Attouch and Wets [1], [2] showed that one can achieve local Lipschitz continuity of K_λ^μ, also when dealing with bivariate functions which are not concave-convex.

Proof. As it has been proved in [17], K_λ^λ has a Gâteaux differential ∇K_λ^λ on $E \times F$; hence it is also subdifferentiable and $\partial K_\lambda^\lambda = \nabla K_\lambda^\lambda$ ([4], Proposition 2.4 or [7], Proposition 5.3). Moreover, formula (13) is true.

The operator $(R_K)_\lambda$ is demicontinuous and bounded as the Yosida approximation of the maximal monotone operator R_K (see [3], [5], [6]). As it is easy to infer that $\|(R_K)_\lambda\|$ is continuous and the spaces are locally uniformly convex, by a well-known criterion we obtain the continuity of $(R_K)_\lambda$. If E and F are Hilbert spaces, $(R_K)_\lambda$ is Lipschitz continuous. Due to (13), the same properties hold for ∇K_λ^λ. Especially, K_λ^λ is Fréchet differentiable as a continuously Gâteaux differentiable function. Q.e.d.

In the next proposition we show how different important notions pass over from K to K_λ^λ.

Proposition 5. *Let K, L be closed, proper saddle functions on $E \times E$. Then*

1) $K \leq L$ *implies* $K_\lambda^\lambda \leq L_\lambda^\lambda$,

2) *if K satisfies one of the conditions*

a) $\mathrm{cl}_2 K(x, x) \leq 0$ *for all* $x \in E$,

b) $\mathrm{cl}_1 K(x, x) \geq 0$ *for all* $x \in E$,

c) K *is skew-symmetric,*

d) $\mathrm{cl}_1 K(x, y) \geq -\mathrm{cl}_2 K(y, x)$ *for all* $x, y \in E$,

then K_λ^λ satisfies the same condition.

Proof. The points 1, 2a, b) are clear by the definition (0). The points 2c), d) can be derived from $L(x, y) = -K(y, x) \Rightarrow L_\lambda^\lambda(x, y) = -K_\lambda^\lambda(y, x)$ and from the relations 1), (8). Q.e.d.

Remark. Many of the results of this section can be extended to

$$K_\lambda^\mu(x,y) = \sup_{u \in E} \inf_{v \in F} \{-\| x - u \|^2/2\lambda + \| y - v \|^2/2\mu + K(u, v)\}.$$

For instance, if by $[x_\lambda^\mu, y_\lambda^\mu]$ we denote the saddle point which defines $K_\lambda^\mu(x, y)$ then formula (13) becomes

$$\partial K_\lambda^\mu(x, y) = [\partial_1 K^\lambda(x, y_\lambda^\mu), \partial_2 K_\mu(x_\lambda^\mu, y)] = [(\partial_1 K)_\lambda(x, y_\lambda^\mu), (\partial_2 K)_\mu(x_\lambda^\mu, y)].$$

The reasoning follows the same line as for (13). See also Attouch-Wets [1], [2], where other properties of K_λ^μ are discussed in a general setting.

5. The Yosida Approximation. Theorem 6. *Let $K : E \times E \to \overline{R}$ be a closed proper saddle function such that T_K is maximal monotone. Then $(T_K)_\lambda = T_{K_\lambda^\lambda}$, i.e. the Yosida approximation of T_K is just the monotone operator associated to the regularization K_λ^λ.*

Let us mention that Theorem 6, applied to the saddle function $K(x, y) = p(y)$, yields the formula $\partial(p_\lambda) = (\partial p)_\lambda$ for the subdifferential of the Moreau-Yosida approximation of the convex function p.

Proof. Let $f = (T_K)_\lambda(x)$, that is $f = -\lambda^{-1} J(x_\lambda - x)$, where x_λ is the solution of the equation

$$J(x_\lambda - x) + T_K(x_\lambda) \ni 0. \tag{14}$$

Above $J : E \to E^*$ is the duality mapping. By the definition of T_K we have

$$[-\lambda^{-1} J(x_\lambda - x), -\lambda^{-1} J(x_\lambda - x)] \in R_K(x_\lambda, x_\lambda). \tag{15}$$

Relation (15) gives $(R_K)_\lambda(x, x) = -\lambda^{-1} [J(x_\lambda - x), J(x_\lambda - x)]$. Therefore $[f, f] = (R_K)_\lambda(x, x)$, and by Theorem 4 it yields $[-f, f] = \partial K_\lambda^\lambda(x, x)$, i.e. $f \in T_{K_\lambda^\lambda} x$. We obtain the inclusion $(T_K)_\lambda \subseteq T_K$, and by the maximal monotonicity of $(T_K)_\lambda$ the proof is finished. Q.e.d.

Corollary 7. *Let $A \subseteq E \times E^*$ be maximal monotone and let $K : E \times E \to \overline{R}$ be a closed proper saddle function with $A = T_K$. Then*

$$A_\lambda x = \partial_2 K_\lambda^\lambda (x, x), \quad x \in E,$$

where the right hand side denotes the Fréchet differential of K_λ^λ with respect to the second argument.

Proof. We have $[-A_\lambda x, A_\lambda x] = \nabla K_\lambda^\lambda (x, x)$ since, by Theorem 4, K_λ^λ is Fréchet differentiable. Q.e.d.

In particular, in Corollary 7 we can choose as K the saddle function L of Theorem 2.

REFERENCES

1. Attouch, H., Wets, R. J.-B. — *A convergence theory for saddle functions.* Trans. A. M. S. 280(1983) 1, 1—41.
2. Attouch, H., Wets, R. J.-B. — *A convergence for bivariate functions aimed at the convergence of saddle values.* In: Lecture Notes in Mathematics 979, Springer (1983), 1—43.
3. Barbu, V. — *Nonlinear semigroups and differential equations in Banach spaces.* Ed. Academiei and Noordhoff Int. Publ. House (1976).
4. Barbu, V., Precupanu, Th. — *Convexity and optimization in Banach spaces.* Ed. Academiei and Sijthoff and Noordhoff Int. Publ. House (1978).
5. Brezis, H. — *Monotonicity methods in Hilbert spaces and some applications to nonlinear partial differential equations.* In: Contributions to Nonlinear Functional Analysis (E. Zarantonello, ed.), Academic Press (1971), 101—156.
6. Brezis, H., Crandall, M., Pazy, A. — *Perturbation of nonlinear maximal monotone sets.* Comm. Pure Appl. Math. 23 (1970), 123—144.
7. Ekeland, I., Temam, R. — *Convex analysis and variational problems.* North Holland Publ. Comp., Amsterdam (1976).
8. Gossez, J.-P. — *On the subdifferential of a saddle function.* Journal Funct. Analysis 11 (1972), 230—230.
9. Krauss, E. — *A representation of maximal monotone operators by saddle functions.* To appear in Revue Romaine Math. Pures et Appl.
10. Krauss, E. — *A representation of arbitrary maximal monotone operators via subgradients of skew-symmetric saddle functions.* To appear in Nonlinear Analysis TMA.
11. Krauss, E. — *Maximal monotone operators and saddle functions.* To appear in Zeitschrift für Anal. Anw.
12. McLinden, L. — *Dual operations on saddle functions.* Trans. A.M.S. 179 (1973), 363—381.
13. Moreau, J.-J — *Fonctionelles convexes,* Séminaire „Équations aux dérivées partielles". Collège de France, Paris (1967).
14. Rockafellar, R. T. — *Convex analysis.* Princeton Univ. Press, Princeton (N.J.) (1970).
15. Rockafellar, R. T. — *Monotone operators associated with saddle functions and minimax problems.* In: Nonlinear Functional Analysis, Part I (F. E. Browder, ed.), Proc. of Symp. in Pure Math. 18 (1970), 241—250.
16. Rockafellar, R. T. — *Saddle points and convex analysis.* In: Differential games and related topics. Am. Elsevier, New York (1971), 109—127.
17. Tiba, D. — *Regularization of saddle functions.* Bolletino U.M.I. (5), 17-A (1980), 420—427.
18. Troyanski, S. — *On local uniform convexity and differentiable norms in certain unseparable Banach spaces.* Studia Math. 37 (1971), 173—180.

Received 26.II. 1985

Akademie der Wissenschaften der DDR
Institut für Mathematik
DDR—1086 Berlin
Mohrenstr. 39, Postfach 1304

INCREST—Bucureşti
Bd. Păcii 220
Bucureşti, România

Analele ştiinţifice ale Universităţii „Al. I. Cuza" din Iaşi
Tomul XXXI, s. I a, Matematică, 1985, f. 3.

SUI SISTEMI LINEARI DI QUADRICHE RIDUCIBILI ED IRRIDUCIBILI

BY

LANDO DEGOLI

1 — Nello spazio lineare complesso S_r riferito a coordinate proiettive omogenee x_i $(i=0, 1, 2, \dots, r)$ si asumano $d+1$ quadriche linearmente indipendenti:

$$f_q = \sum_{i,k=0}^{r} a_q^{ik} x_i x_k \quad (q=0, 1, 2, \dots, d)$$

Un sistema lineare di quadriche di dimensione d è dato dalla relazione:

$$\sum_{q=0}^{d} \Lambda_q f_q = 0.$$

Supponiamo che la matrice Jacobiana a $r+1$ righe e $d+1$ colonne:

$$J = \left\| \frac{\partial f_q}{\partial x_s} \right\| \begin{pmatrix} q=0, 1, 2, \dots, d \\ s=0, 1, 2, \dots, r \end{pmatrix}$$

abbia caratteristica $m=r-k (k \geq 0)$.

È noto che quando la Jacobiana è identicamente nulla, l'intero S_r è luogo di punti coniugati rispetto a tutte le quadriche. Se la caratteristica e $r-h (h \geq 0)$, un punto generico di S_r è coniugato con un S_h. Un sistema lineare $L_{d/m}$ di dimensione d e caratteristica m possiede $d+1$ quadriche linearmente indipendenti, tra cui è possibile scegliere m quadriche funzionalmente indipendenti da cui dipendono tutte le altre. Un sistema $L_{d/m}$ $(m \leq d)$ che non possiede *sistemi subordinati essenziali* $L_{g/c}$ (vedi: [6]) con:

$$2 \leq g \leq d-1, \quad 2 \leq c \leq m-1, \quad c \leq g$$

è detto *irriducibile di prima specie.*

Abbiamo dimostrato il:

Teorema A: *Condizione necessaria e sufficiente affinchè un sistema lineare irriducibile di prima specie sia a Jacobiana di caratteristica $r-k \leq d (k \geq 0)$ è che le quadriche del sistema che passano per un punto generico di S_r abbiano in comune un S_{k+1}.*

Dati due sistemi lineari L_a ed L_b, aventi in comune un sistema lineare L_c, il loro *sistema congiungente* risulta di dimensione: $a+b-c$.

Diremo che il sistema $L_{d/m} (m \leq d)$ è *riducibile* se congiunge dei sistemi subordinati, tra i quali almeno uno L_{d_1/m_1} $(m_1 \leq d_1)$ non ha quadriche in comune con gli altri.