On a new notion of regularizer

2003 J. Phys. A: Math. Gen. 36 2191

(http://iopscience.iop.org/0305-4470/36/8/315)

INSTITUTE OF PHYSICS PUBLISHING JOURNAL OF PHYSICS A: MATHEMATICAL AND GENERAL

J. Phys. A: Math. Gen. **36** (2003) 2191–2195 PII: S0305-4470(03)54339-6

On a new notion of regularizer

A G Ramm

LMA/CNRS, 31 Chemin Joseph Aiguier, Marseille 13402, France
and
Mathematics Department, Kansas State University, Manhattan, KS 66506-2602, USA

E-mail: ramm@math.ksu.edu

Received 2 October 2002, in final form 2 January 2003
Published 12 February 2003
Online at stacks.iop.org/JPhysA/36/2191

Abstract
A new understanding of the notion of regularizer is proposed. It is argued that this new notion is more realistic than the old one and better fits the practical computational needs. An example of the regularizer in the new sense is given. A method for constructing regularizers in the new sense is proposed and justified.

PACS numbers: 02.30.Tb, 02.60.Jh
Mathematics Subject Classification: 47A52, 65F22, 65J20

1. Introduction

Let

$$A(u) = g \tag{1.1}$$

where $A : X \to Y$ is a closed, possibly nonlinear, map from a Banach space X into a Banach space Y. Problem (1.1) is called ill posed if A is not a homeomorphism of X onto Y, that is, either equation (1.1) does not have a solution, or the solution is non-unique, or the solution does not depend on g continuously. Let us assume that (1.1) has a solution u and this solution is unique, but A^{-1} is not continuous. Given noisy data g_δ, $\|g_\delta - g\| \leqslant \delta$, one wants to construct a stable approximation u_δ of the solution u, $\|u_\delta - u\| \to 0$ as $\delta \to 0$. This is often done with the help of a regularizer. Traditionally (see, e.g., [2]) one calls a family of operators R_h a regularizer for problem (1.1) if

(a) $R_h A(u) \to u$ as $h \to \infty$ for any $u \in D(A)$,
(b) $R_h g_\delta$ is defined for any $g_\delta \in Y$ and there exists $h(\delta) \to 0$ as $\delta \to 0$ such that

$$\|R_{h(\delta)} g_\delta - u\| \to 0 \qquad \text{as} \quad \delta \to 0 \tag{$*$}$$

where u solves (1.1).

In this definition, u is fixed and $(*)$ must hold for any $g_\delta \in B(g,\delta) := \{g_\delta : \|g_\delta - g\| \leqslant \delta\}$.

In practice, one does not know the solution u. The only available information is a family g_δ and some *a priori* information about the solution u. This information very often consists of the knowledge that $u \in \mathcal{K}$, where $\mathcal{K}$ is a compactum in X. Thus $u \in S_\delta := \{v : \|A(v) - g_\delta\| \leqslant \delta, v \in \mathcal{K}\}$. We assume that the operator A is known exactly and we always assume that $g_\delta \in B(g,\delta)$, where $g = A(u)$.

It is natural to call a family of operators $R(\delta)$ a regularizer (in the *new sense*) if

$$\sup_{v\in S_\delta} \|R(\delta)g_\delta - v\| \leqslant \eta(\delta) \to 0 \qquad \text{as} \quad \delta \to 0. \tag{1.2}$$

There is a crucial difference between our definition (1.2) and the standard definition $(*)$: in $(*)$, u is fixed, while in (1.2), v is any element of S_δ and the supremum, over all such v, of the norm in (1.2) must go to zero as $\delta \to 0$.

The new definition is more realistic and fits more computational needs because not only the solution u to (1.1) satisfies the inequality $\|A(u) - g_\delta\| \leqslant \delta$, but many v satisfy such an inequality $\|Av - g_\delta\| \leqslant \delta$, $v \in \mathcal{K}$, and the data g_δ may correspond, in fact, to any $v \in S_\delta$, and not only to the solution of equation (1.1). Therefore it is more natural to use definition (1.2) than $(*)$.

Our aim is to illustrate the practical difference in these two definitions by an example, and to construct regularizer in the sense (1.2) for problem (1.1) under the following assumptions:

(A1) $A : X \to Y$ is a closed, possibly nonlinear, injective map, $g \in \mathcal{R}(A)$, $\mathcal{R}(A)$ is the range of A,
and

(A2) $\phi : D(\phi) \to [0,\infty)$, $\phi(u) > 0$ if $u \neq 0$, domain $D(\phi) \subseteq D(A)$, the set $\mathcal{K} = \mathcal{K}_c := \{v : \phi(v) \leqslant c\}$ is compact in X and contains a sequence $v_n \to v$, and $\phi(v) \leqslant \liminf_{n\to\infty} \phi(v_n)$.

The last inequality holds if ϕ is lower semicontinuous. Any Hilbert space norm and norms in reflexive Banach spaces have this property.

Examples in which assumptions (A1) and (A2) are satisfied are numerous.

Example 1. A is a linear injective compact operator, $g \in \mathcal{R}(A)$, $\phi(v)$ is a norm on $X_1 \subset X$, where X_1 is densely embedded in X, the embedding $i : X_1 \to X$ is compact, and $\phi(v)$ is lower semicontinuous.

Example 2. A is a nonlinear injective continuous operator, $g \in \mathcal{R}(A)$, A^{-1} is not continuous, ϕ is as in example 1.

Example 3. A is linear, injective, densely defined, closed operator, $g \in \mathcal{R}(A)$, A^{-1} is unbounded, ϕ is as in example 1, $X_1 \subseteq D(A)$.

In section 2, it is shown that a regularizer in the sense $(*)$ may not be a regularizer in the sense (1.2). In section 3, a theoretical construction of a regularizer in the sense (1.2) is given.

2. Example: stable numerical differentiation

Here we use the results from [3–10].

Consider stable numerical differentiation of noisy data. The problem is

$$Au := \int_0^x u(s)\,\mathrm{d}s = g(x) \qquad g(0) = 0 \qquad 0 \leqslant x \leqslant 1. \tag{2.1}$$

The data are: g_δ and M_a, where $\|g_\delta - g\| \leqslant \delta$, the norm is $L^\infty(0,1)$ norm, and $\|u\|_a \leqslant M_a$, $a \geqslant 0$. The norm

$$\|u\|_a := \sup_{\substack{x,y\in[0,1]\\ x\neq y}} \frac{|u(x)-u(y)|}{|x-y|^a} + \sup_{0\leqslant x\leqslant 1} |u(x)| \qquad \text{if} \quad 0\leqslant a\leqslant 1$$

$$\|u\|_a := \sup_{0\leqslant x\leqslant 1} (|u(x)| + |u'(x)|) + \sup_{\substack{x,y\in[0,1]\\ x\neq y}} \frac{|u'(x)-u'(y)|}{|x-y|^{a-1}} \qquad 1 < a \leqslant 2.$$

If $a > 1$, then we define

$$R(\delta)g_\delta := \begin{cases} \dfrac{g_\delta(x+h(\delta)) - g_\delta(x-h(\delta))}{2h(\delta)} & h(\delta)\leqslant x\leqslant 1-h(\delta) \\ \dfrac{g_\delta(x+h(\delta)) - g_\delta(x)}{h(\delta)} & 0\leqslant x < h(\delta) \\ \dfrac{g_\delta(x) - g_\delta(x-h(\delta))}{h(\delta)} & 1-h(\delta) < x\leqslant 1 \end{cases} \tag{2.2}$$

where

$$h(\delta) = c_a \delta^{\frac{1}{a}} \tag{2.3}$$

and c_a is a constant given explicitly (cf [4]).

We prove that (2.2) is a regularizer for (2.1) in the sense (1.2), and $\mathcal{K} := \{v : \|v\|_a \leqslant M_a, a > 1\}$. In this example, we do not use lower semicontinuity of the norm $\phi(v)$ and do not define ϕ.

Let $S_{\delta,a} := \{v : \|Av - g_\delta\| \leqslant \delta, \|v\|_a \leqslant M_a\}$. To prove that (2.2)–(2.3) is a regularizer in the sense (1.2) we use the estimate

$$\sup_{v\in S_{\delta,a}} \|R(\delta)g_\delta - v\| \leqslant \sup_{v\in S_{\delta,a}} \{\|R(\delta)(g_\delta - Av)\| + \|R(\delta)Av - v\|\}$$

$$\leqslant \frac{\delta}{h(\delta)} + M_a h^{a-1}(\delta) \leqslant c_a \delta^{1-\frac{1}{a}} := \eta(\delta) \to 0 \qquad \text{as} \quad \delta\to 0. \tag{2.4}$$

Thus we have proved that (2.2)–(2.3) is a regularizer in the sense (1.2).

If $a = 1$ and $M_1 < \infty$, then one can prove that *there is no* regularizer for problem (2.1) in the sense (1.2) even if the regularizer is sought in the set of all operators, including nonlinear ones. More precisely, it is proved in [5] p 345 (see also [7]) that

$$\inf_{R(\delta)} \sup_{v\in S_{\delta,1}} \|R(\delta)g_\delta - v\| \geqslant c > 0$$

where $c > 0$ is a constant independent of δ and the infimum is taken over all operators $R(\delta)$ acting from $L^\infty(0,1)$ into $L^\infty(0,1)$, including nonlinear ones. On the other hand, if $a = 1$ and $M_1 < \infty$, then a regularizer in the sense $(*)$ does exist, but the rate of convergence in $(*)$ may be as slow as one wishes, if $u(x)$ is chosen suitably [9, 10].

3. Construction of a regularizer in the sense (1.2)

Assuming (A1) and (A2) (see section 1) throughout this section, let us construct a regularizer for (1.1) in the sense (1.2). We use the ideas from [10]. Define $F_\delta(v) := \|Av - g_\delta\| + \delta\phi(v)$ and consider the minimization problem of finding the infimum $m(\delta)$ of the functional $F_\delta(v)$ on a set S_δ defined below:

$$F_\delta(v) = \inf := m(\delta) \qquad v\in S_\delta := \{v : \|Av - g_\delta\| \leqslant \delta, \phi(v)\leqslant c\} \tag{3.1}$$

so that $\mathcal{K} = \mathcal{K}_c := \{v : \phi(v) \leqslant c\}$. The constant $c > 0$ can be chosen arbitrary large and fixed at the beginning of the argument, and then one can choose a smaller constant c_1 specified below. Since $F_\delta(u) = \delta + \delta\phi(u) := c_1\delta$, $c_1 := 1 + \phi(u)$, where u solves (1.1), one concludes that

$$m(\delta) \leqslant c_1\delta. \tag{3.2}$$

Let v_j be a minimizing sequence, such that $F_\delta(v_j) \leqslant 2m(\delta)$. Then $\phi(v_j) \leqslant 2c_1$. By assumption (A2), as $j \to \infty$, one has

$$v_j \to v_\delta \qquad \phi(v_\delta) \leqslant 2c_1. \tag{3.3}$$

Take $\delta = \delta_m \to 0$ and denote $v_{\delta_m} := w_m$. Then (3.3) and assumption (A2) imply existence of a subsequence, denoted again by w_m, such that

$$w_m \to w \qquad A(w_m) \to A(w) \qquad \|A(w) - g\| = 0. \tag{3.4}$$

Thus $A(w) = g$ and, since A is injective, $w = u$, where u is the unique solution to (1.1).

Define now $R(\delta)g_\delta$ by the formula $R(\delta)g_\delta := v_\delta$.

Theorem 3.1. *$R(\delta)$ is a regularizer for problem (1.1) in the sense (1.2).*

Proof. Assume the contrary

$$\sup_{v\in S_\delta} \|R(\delta)g_\delta - v\| = \sup_{v\in S_\delta} \|v_\delta - v\| \geqslant \gamma > 0 \tag{3.5}$$

where $\gamma > 0$ is a constant independent of δ. Since $\phi(v_\delta) \leqslant 2c_1$ by (3.3) and $\phi(v) \leqslant c$, one can choose convergent in X sequences $w_m := v_{\delta_m} \to \tilde{w}$, $\delta_m \to 0$ and $v_m \to \tilde{v}$, such that $\|w_m - v_m\| \geqslant \frac{\gamma}{2}$, $\|\tilde{w} - \tilde{v}\| \geqslant \frac{\gamma}{2}$, and $A(\tilde{w}) = g$, $A(\tilde{v}) = g$. Therefore, by the injectivity of A, $\tilde{w} = \tilde{v} = u$, and one gets a contradiction with the inequality $\|\tilde{w} - \tilde{v}\| \geqslant \frac{\gamma}{2} > 0$. This contradiction proves theorem 3.1.

Note that the conclusions $A(\tilde{w}) = g$ and $A(\tilde{v}) = g$ follow from the inequalities $\|A(v_\delta) - g_\delta\| \leqslant \delta$ and $\|A(v) - g_\delta\| \leqslant \delta$ after passing to the limit $\delta \to 0$, using assumption (A2). □

Remark 3.2. Our argument is closely related to a generalization of a classical result [1] which says that an injective and continuous map of a compactum in a Banach space has continuous inverse. The generalization of this result, given in [6] (ch 5, section 6, lemma 2), says that the same conclusion holds if A is an injective and closed map. This generalization, in other words, claims that if A is an injective and closed map of a compactum $\mathcal{K}$ in a Banach space into a Banach space, then the modulus of continuity $\omega(\delta)$ of the inverse operator A^{-1} on the set $A(\mathcal{K})$ tends to zero:

$$\sup_{\substack{\|Av-Aw\|\leqslant\delta \\ v,w\in\mathcal{K}}} \|v - w\| := \omega(\delta) \to 0 \qquad \text{as} \quad \delta \to 0.$$

Therefore, in our proof of theorem 3.1, one has $\sup_{v\in S_\delta} \|v_\delta - v\| \to 0$ as $\delta \to 0$.

Indeed, $v \in S_\delta$ and $v_\delta \in S_\delta$, so $\|A(v) - g_\delta\| \leqslant \delta$ and $\|A(v_\delta) - g_\delta\| \leqslant \delta$. Thus $\|A(v) - A(v_\delta)\| \leqslant 2\delta$. Since v_δ and v belong to a compactum $\mathcal{K}_{c_1} := \{v : \phi(v) \leqslant 2c_1\}$, theorem 3.1 follows from the property $\omega(\delta) \to 0$ as $\delta \to 0$.

Remark 3.3. We have assumed that A is injective, that is, equation (1.1) has at most one solution. This assumption can be relaxed: one may assume that u solves (1.1) and there is an $r > 0$, such that in the ball $B(u, r) := \{v : \|v - u\| \leqslant r\}$, the solution u is unique, but globally (1.1) may have many solutions. Our arguments remain valid if the compactum $\mathcal{K} := \{v : \phi(v) \leqslant c, v \in B(u, r)\}$.

One may drop the injectivity of A in assumption (A1) in section 1. In this case, (3.4) yields $w \in U$, where $U := \{w : A(w) = g\}$. Thus, $\lim_{\delta\to 0} \rho(R(\delta)g_\delta, U) = 0$, where $\rho(w, U)$ is the distance from an element w to the set U.

In applications, when a physical problem is reduced to equation (1.1), one wants to have a unique solution to (1.1). If the solution is non-unique, that is, U contains more than one element, then one wants to impose additional conditions which select a unique solution in the set U, thus making the operator A injective.

References

[1] Dunford N and Schwartz J 1958 *Linear Operators* (New York: Interscience)

[2] Engl H, Hanke M and Neubauer A 1996 *Regularization of Inverse Problems* (Dordrecht: Kluwer)

[3] Ramm A G 1968 On numerical differentiation *Mathematics Izvestija vuzov* **11** 131–5 (in Russian) (Engl. Transl. *Math. Rev.* **40** 5130)

[4] Ramm A G 1981 Stable solutions of some ill-posed problems *Math. Methods Appl. Sci.* **3** 336–63

[5] Ramm A G 1986 *Scattering by Obstacles* (Dordrecht: Reidel) pp 1–442

[6] Ramm A G 1990 *Random Fields Estimation Theory* (New York: Longman Scientific and Wiley)

[7] Ramm A G 2000 Inequalities for the derivatives *Math. Ineq. Appl.* **3** 129–32

[8] Ramm A G and Smirnova A 2001 On stable numerical differentiation *Math. Comput.* **70** 1131–53

[9] Ramm A G and Smirnova A B 2003 Stable numerical differentiation: when is it possible? submitted

[10] Ramm A G 2002 Regularization of ill-posed problems with unbounded operators *J. Math. Anal. Appl.* **271** 447–50