Chapter 9
A Note on Quadratic Penalties for Linear Ill-Posed Problems: From Tikhonov Regularization to Mollification

Pierre Maréchal

Abstract The variational form of mollification fits in an extension of the generalized Tikhonov regularization. Using tools from variational analysis, we prove asymptotic consistency results for both this extended framework and the particular form of mollification that one obtains when building on the notion of target object.

Keywords Ill-posed problems · Regularization theory · Mollification

9.1 Introduction

Ill-posed inverse problems appear in many areas of applied mathematics, such as signal and image recovery, partial differential equations and statistics. Many of them take the form of a linear operator equation

$$Tf = g, \quad f \in F,$$

in which $T\colon F \to G$ is a bounded linear operator between the Hilbert spaces F and G and $g \in G$ is the data. Unfortunately, it frequently occurs that

$$\inf \left\{ \|Tf\| \;\middle|\; f \in (\ker T)^{\perp},\ \|f\| = 1 \right\} = 0,$$

a condition under which the pseudo-inverse $T^{\dagger}$ of T is unbounded. It results that the *natural* solution $T^{\dagger}g$ does not depend continuously on the data g and that the problem must be reformulated. Tikhonov regularization (see [15] and the references therein) initiated a vast theoretical corpus. It consists in approximating $T^{\dagger}$ by the bounded operator $R_{\alpha} = (T^*T + \alpha I)^{-1}T^*$, in which T^* denotes the adjoint of T and $\alpha > 0$ is a *regularization parameter*. The identity I may also be replaced by the more general selfadjoint operator Q^*Q, where Q is a bounded operator from F to some Hilbert

P. Maréchal (✉)
Institut de Mathématiques de Toulouse, Université Paul Sabatier, Toulouse, France
e-mail: pr.marechal@gmail.com

V. Laha et al. (eds.), *Optimization, Variational Analysis and Applications*,
Springer Proceedings in Mathematics & Statistics 355,
https://doi.org/10.1007/978-981-16-1819-2_9

space H. We then speak of *generalized Tikhonov regularization*. From the variational viewpoint, the generalized Tikhonov solution $f_\alpha = (T^*T + \alpha Q^*Q)^{-1}T^*g$ is well known to be the minimizer of the quadratic functional

$$\mathscr{F}_\alpha(f) := \| Tf - g \|^2 + \alpha \| Qf \|^2. \tag{9.1}$$

In many cases, the solution space F is a functional space such as $L^2(\Re^d)$ or a subspace of it, and the quadratic penalty term $\alpha\| Qf \|^2$ may be used to enforce smoothness of the approximate solution. For example, Q may be a second-order differential operator (see [3, Chap. 8] for a detailed exposition).

Another way to promote smoothness is via the Fourier–Plancherel transform $\hat{f}$ of f: the variational counterpart of *mollification* [1, 6, 7, 9–11] essentially consists in penalizing $(1 - \hat{\varphi}_\alpha)\hat{f}$, in which φ_α is a convolution kernel indexed by $\alpha > 0$. The function φ_α is commonly defined, for $\alpha \in (0, 1]$, as

$$\varphi_\alpha(x) = \frac{1}{\alpha^d}\varphi\left(\frac{x}{\alpha}\right), \quad x \in \Re^d, \tag{9.2}$$

in which φ is a nonnegative integrable kernel function with unit integral, and the family $(\varphi_\alpha)_{\alpha\in(0,1]}$ is referred to as an *approximate unity*. The penalty term of mollification then takes the form $\| (I - C_\alpha)f \|^2$, in which

$$C_\alpha f = \varphi_\alpha * f.$$

Mollifiers were introduced in partial differential equations by Friedrichs [4, 16]. The term *mollification* has been used in regularization theory since the eighties. Mollification was developed in several directions. In the earlier works on the subject, mollifiers served the purpose of smoothing the data prior to inversion, whenever an explicit inversion formula was available (see [5, 12] and the references therein). In [8], an alternative approach was proposed, which gave rise to the so-called *method of approximate inverses*. In this approach, the operator under consideration is not assumed to have explicit inverse, but the adjoint equation has explicit solutions. This approach opens the way to application to a large class of inverse problems and can be extended to problems in Banach spaces [14]. A third approach appeared in the same period of time. In [7], a variational formulation of the idea of mollification was proposed, in the context of Fourier synthesis and deconvolution. This formulation was further studied and extended in [1, 6, 9, 11] and is the one we consider in this paper.

Unlike Tikhonov's regularization, mollification appeals to a parameter α which is not interpreted as a weighting of the penalty term, but rather as an *objective resolution*. Therefore, strictly speaking, mollification does not belong to the generalized Tikhonov family. However, obviously, letting α go to zero makes the penalization vanish in both cases. This suggests that Tikhonov and the mollification could be put in the same framework. To phrase it differently, we could widen the contours of the generalized Tikhonov regularization to the point of admitting mollification in its realm. This is what we propose to do here.

The paper is organized as follows. In Sect. 9.2 we consider the consistency issue in the aforementioned enlarged framework. In Sect. 9.3, we build on the notion of *target object*, absent from the original Tikhonov regularization, but present in the original works on mollification [1, 2, 7].

9.2 Generalizing Tikhonov Regularization

It is sometimes convenient to consider vector-valued regularization parameters. We may call *parameter choice rule* a function

$$\begin{aligned} \alpha : \Re_+ \times G &\longrightarrow \mathscr{P} \\ (\delta, g^\delta) &\longmapsto \alpha(\delta, g^\delta) \end{aligned}$$

in which $\mathscr{P}$ is a subset of $\Re^p_+ \setminus \{0\}$, and an *a priori parameter choice rule* the particular case for which α depends on its first argument only. Following [3, Definition 3.1], we now state:

Definition 9.1 A parametrized family (R_α) of bounded operators is a *regularization of* $T^\dagger$ if for every $g \in \mathscr{D}(T^\dagger)$, there exists a parameter choice rule α such that

(1) $\sup \left\{ \|\alpha(\delta, g^\delta)\| \;\middle|\; g^\delta \in G,\ \|g^\delta - g\| \le \delta \right\} \to 0$ as $\delta \downarrow 0$;
(2) $\sup \left\{ \|R_{\alpha(\delta,g^\delta)} g^\delta - T^\dagger g\| \;\middle|\; g^\delta \in G,\ \|g^\delta - g\| \le \delta \right\} \to 0$ as $\delta \downarrow 0$.

In this case, we say that the pair (R_α, α) is a convergent regularization method for solving $Tf = g$.

Recall that the domain of the operator $T^\dagger$ is the vector subspace $\mathscr{D}(T^\dagger) = \operatorname{ran} T + (\operatorname{ran} T)^\perp$, in which $E^\perp$ denotes the orthogonal complement of E. From [3, Proposition 3.4], we straightforwardly infer that:

Proposition 9.1 *If the family of bounded operators $(R_\alpha)_{\alpha\in\mathscr{P}}$ converges pointwise to $T^\dagger$ on $\mathscr{D}(T^\dagger)$ as $\alpha \to 0$ in $\mathscr{P}$, then $(R_\alpha)_{\alpha\in\mathscr{P}}$ is a regularization of $T^\dagger$ and, for every $g \in \mathscr{D}(T^\dagger)$, there exists an a priori parameter choice rule $\alpha(\delta)$ such that (R_α, α) is a convergent regularization method for solving $Tf = g$.*

The operators $T : F \to G$ and $Q : F \to H$ are said to satisfy Morozov's *completion condition* if there exists a constant $\gamma > 0$ such that

$$\forall f \in F, \quad \|Tf\|^2 + \|Qf\|^2 \ge \gamma \|f\|^2. \tag{9.3}$$

Under the completion condition, the operator $T^*T + Q^*Q$ admits a bounded inverse, as can be easily shown. In some cases of interest, it may happen that T^*T and Q^*Q can be *diagonalized* in the same Hilbert basis. In this case, it can be shown that

$$\forall f \in F, \quad \left\| (T^*T + Q^*Q)^{-1}T^*Tf \right\|_F \leq \left\| f \right\|_F. \tag{9.4}$$

The latter assumption is in force in the rest of this paper.

Theorem 9.2.1 *Let F, G be infinite dimensional Hilbert spaces and let $T\colon F \to G$ be injective. Let $Q_\alpha\colon F \to H$ be a family of operators such that*

- *for every fixed $\alpha \in \mathscr{P}$, T and Q_α satisfy the completion condition (9.3) and Condition (9.4);*
- *for every $f \in F$, $\| Q_\alpha f \| \to 0$ as $\alpha \to 0$ in $\Re^p$.*

Then, for every $g \in \mathscr{D}(T^\dagger) = \operatorname{ran} T + \operatorname{ran} T^\perp$, $f_\alpha := (T^*T + Q_\alpha^* Q_\alpha)^{-1}T^*g$ converges strongly to $f^\dagger$, the unique least square solution of the equation $Tf = g$.

Proof We shall prove that, for every $\mathscr{P}$-valued sequence (α_n) which converges to zero, the corresponding sequence (f_{α_n}) strongly converges to $f^\dagger$. By assumption, $g = Tf^\dagger + g^\perp$, in which $f^\dagger \in F$ and $g^\perp \in (\operatorname{ran} T)^\perp = \ker T^*$. We have

$$\begin{aligned} \left\| f_\alpha \right\|_F &= \left\| (T^*T + Q_\alpha^* Q_\alpha)^{-1}T^*(Tf^\dagger + g^\perp) \right\|_F \\ &= \left\| (T^*T + Q_\alpha^* Q_\alpha)^{-1}T^*Tf^\dagger \right\|_F \\ &\leq \left\| f^\dagger \right\|_F. \end{aligned}$$

In particular, the family f_α is bounded. Now, let (α_n) be a sequence in $\mathscr{P}$ which converges to 0. In order to simplify the notation, let $f_n := f_{\alpha_n}$ and $Q_n := Q_{\alpha_n}$. Since the sequence (f_n) is bounded, we can extract a weakly convergent subsequence (f_{n_k}). Let then $\tilde{f}$ be the weak limit of this subsequence. On the one hand,

$$T^*Tf_{n_k} \rightharpoonup T^*T\tilde{f} \quad \text{as} \quad k \to \infty \tag{9.5}$$

since T^*T is bounded. On the other hand,

$$Q_{n_k}^* Q_{n_k} f_{n_k} \rightharpoonup 0 \quad \text{as} \quad k \to \infty$$

since f_{n_k} is bounded and $Q_{n_k}^* Q_{n_k}$ converges pointwise to the null operator, so that

$$\begin{aligned} T^*Tf_{n_k} &= (T^*T + Q_{n_k}^* Q_{n_k})f_{n_k} - Q_{n_k}^* Q_{n_k} f_{n_k} \\ &= T^*g - Q_{n_k}^* Q_{n_k} f_{n_k} \\ &= T^*Tf^\dagger - Q_{n_k}^* Q_{n_k} f_{n_k} \\ &\rightharpoonup T^*Tf^\dagger \end{aligned}$$

□

as $k \to \infty$. Together with (9.5), this shows that $T^*T\tilde{f} = T^*Tf^\dagger$, that is, by the injectivity of T, that $\tilde{f} = f^\dagger$. It follows that the whole sequence (f_n) converges weakly to $f^\dagger$. Finally, by the weak lower semicontinuity of the norm,

$$\| f^\dagger \| \leq \liminf_{n\to\infty} \| f_n \| \leq \limsup_{n\to\infty} \| f_n \| \leq \| f^\dagger \|,$$

which establishes that $f_n \to f^\dagger$ as $n \to \infty$. ■

In the familiar case where $F = L^2(\mathfrak{R}^d)$ or subspaces of it and $H = L^2(\mathfrak{R}^d)$, the choice $Q_\alpha = I - C_\alpha$ corresponds to the mollification method described in the introduction. The previous theorem applies in this case since, as is well known, if (C_α) is as in (9.2), then

$$C_\alpha f \to f \quad \text{as} \quad \alpha \downarrow 0.$$

Notice that Morozov's completion condition is automatically satisfied in the important case where $F = L^2(V)$, the space of square integrable functions with essential support in V, whenever V is a compact domain. As a matter of fact, in this case, it follows from [1, Lemma 12 and Proposition 5] that there exists a positive constant ν_α such that

$$\forall f \in L^2(V), \quad \|(I - C_\alpha) f\|^2 \geq \nu_\alpha \| f \|^2.$$

9.3 Target Objects

In a number of cases, the operator T gives rise to an explicit *intertwining relationship*. By this, we mean the existence of a bounded operator $\Phi_\alpha : G \to G$ such that

$$T C_\alpha = \Phi_\alpha T. \tag{9.6}$$

Note that Eq. (9.6) constrains Φ_α only on the range of T. In order to extend its definition to the whole space G, we first use the unique bounded extension of Φ_α to the closure of ran T, and then extend it further by zero on $(\operatorname{ran} T)^\perp$. With this definition of Φ_α, it is easy to see that

$$\Phi_\alpha = \operatorname{cl}(T C_\alpha T^\dagger), \tag{9.7}$$

in which cl $(\cdot)$ denotes the extension by closure. More generally, it has been shown in [2] that whenever the operator $T C_\alpha T^\dagger$ is bounded, its closure to G minimizes $\Phi \mapsto \| \Phi T - T C_\alpha \|$ over all the bounded operators on G which vanish on $(\operatorname{ran} T)^\perp$.

At all events, we may consider the following variational form of mollification:

$$f_\alpha := \operatorname{argmin} \left\| T f - \Phi_\alpha g \right\|^2 + \left\| (I - C_\alpha) f \right\|^2. \tag{9.8}$$

This form can be justified by the following heuristics. Since our *target object* is $C_\alpha f^\dagger$, the tautology $f^\dagger = C_\alpha f^\dagger + (I - C_\alpha) f^\dagger$ indicates that in addition to penalizing $(I - C_\alpha) f$ one should also aim at fitting the data corresponding to the mollified object. If $g \simeq T f^\dagger$, then

$$\Phi_\alpha g \simeq T C_\alpha f^\dagger$$

by Eq. (9.6), whence the adequacy term in (9.8). The regularized solution is then given by

$$f_\alpha := \big(T^*T + (I - C_\alpha)^*(I - C_\alpha)\big)^{-1} T^* \Phi_\alpha g.$$

Important applications allow for the introduction of the *intertwining operator* Φ_β corresponding to approximate unities such as the above-defined families (C_α). We now review a few examples.

Example 9.1 In [7], the authors studied the problem of *spectral extrapolation*, which underlies *aperture synthesis* in astronomy and space imaging. This problem corresponds to the case where

$$T = frm[o]--e_W U$$

with W a bounded domain containing an open set. Here, U denotes the Fourier–Plancherel operator. We refer to T_W as the Fourier truncation operator. Since $C_\alpha = U^{-1}\big[\hat\varphi_\alpha\big]U$, we see that

$$T C_\alpha = \mathbb{1}_W U U^{-1}\big[\hat\varphi_\alpha\big]U = \big[\hat\varphi_\alpha\big]\mathbb{1}_W U = \big[\hat\varphi_\alpha\big]T,$$

from which we infer that $\Phi_\alpha = \big[\hat\varphi_\alpha\big]$. ∎

Example 9.2 In the problem of deconvolution, as considered, e.g., in [6, 9], the situation is even simpler: since convolution operators commute, we readily see that $\Phi_\beta = C_\alpha$. ∎

Example 9.3 Finally, in computerized tomography [13], the underlying operator is the Radon transformation

$$(Tf)(\boldsymbol\theta, s) = \int f(\mathbf{x})\delta(s - \langle\boldsymbol\theta, \mathbf{x}\rangle)\,\mathrm{d}\mathbf{x}, \quad \boldsymbol\theta \in \mathscr{S}^1, \quad s \in \Re.$$

A consequence of the so-called *Fourier slice theorem* is that, for any two functions f_1, f_2,

$$T(f_1 * f_2) = Tf_1 \circledast Tf_2,$$

in which $\circledast$ denotes the convolution with respect to the variable s. It follows that, in this case,

$$\Phi_\alpha = (g \mapsto T\varphi_\alpha \circledast g),$$

a relationship which was in force in [11]. ∎

We now establish a consistency theorem for the form of mollification given in (9.8).

Theorem 9.3.1 *Let $F = L^2(\Re^d)$ and let $T\colon F \to G$ be a bounded injective operator from F to the infinite dimensional Hilbert space G. Let $C_\alpha\colon F \to F$ be an approximate unity as in (9.2). Assume that, for every fixed $\alpha \in (0, 1]$, T and $I - C_\alpha$ satisfy the completion condition (9.3). Assume at last that, for every fixed $\alpha \in (0, 1]$, the intertwining operator Φ_α exists. Then, for every $g \in \mathscr{D}(T^\dagger) = \operatorname{ran} T + \operatorname{ran} T^\perp$, $f_\alpha := (T^*T + (I - C_\alpha)^*(I - C_\alpha))^{-1}T^*\Phi_\alpha g$ converges strongly to $f^\dagger$.*

Proof We shall prove that, for every positive sequence (α_n) converging to zero, $f_{\alpha_n} \to f^\dagger$ as $n \to \infty$. Let $g = Tf^\dagger + g^\perp$, with $f^\dagger \in F$ and $g^\perp \in (\operatorname{ran} T)^\perp$. Since $\Phi_\alpha g = TC_\alpha f^\dagger$, we have:

$$\begin{aligned}\left\| f_\alpha \right\|_F &= \left\| (T^*T + (I - C_\alpha)^*(I - C_\alpha))^{-1}T^*TC_\alpha f^\dagger \right\|_F \\ &\le \left\| C_\alpha f^\dagger \right\|_F \\ &\le \left\| \varphi_\alpha \right\|_1 \cdot \left\| f^\dagger \right\|_F \;=\; \left\| f^\dagger \right\|_F .\end{aligned}$$

The last equality stems from the fact that, in Eq. (9.2), φ is assumed to be positive and to have unit integral. Therefore, the family (f_α) is bounded. Let (α_n) be a sequence in $(0, 1]$ which converges to 0, and let $f_n := f_{\alpha_n}$, $C_n := C_{\alpha_n}$ and $\Phi_n := \Phi_{\alpha_n}$. Since the sequence (f_n) is bounded, we can extract a weakly convergent subsequence (f_{n_k}). Let then $\tilde{f}$ be the weak limit of this subsequence. On the one hand,

$$T^*Tf_{n_k} \rightharpoonup T^*T\tilde{f} \quad \text{as} \quad k \to \infty \tag{9.9}$$

since T^*T is bounded. On the other hand,

$$(I - C_{n_k})^*(I - C_{n_k})f_{n_k} \rightharpoonup 0 \quad \text{as} \quad k \to \infty$$

since f_{n_k} is bounded and $(I - C_{n_k})^*(I - C_{n_k})$ converges pointwise to the null operator, so that

$$\begin{aligned}T^*Tf_{n_k} &= (T^*T + (I - C_{n_k})^*(I - C_{n_k}))f_{n_k} - (I - C_{n_k})^*(I - C_{n_k})f_{n_k} \\ &= T^*\Phi_{n_k}g - (I - C_{n_k})^*(I - C_{n_k})f_{n_k} \\ &\rightharpoonup T^*Tf^\dagger\end{aligned}$$

□

as $k \to \infty$, since $T^*\Phi_{n_k}g = T^*TC_{n_k}f^\dagger$ goes to $T^*Tf^\dagger$. Together with (9.9), this shows that $T^*T\tilde{f} = T^*Tf^\dagger$, that is, by the injectivity of T, that $\tilde{f} = f^\dagger$. Therefore, the whole sequence (f_n) converges weakly to $f^\dagger$. Finally, by the weak lower semicontinuity of the norm,

$$\| f^\dagger \| \le \liminf_{n\to\infty} \| f_n \| \le \limsup_{n\to\infty} \| f_n \| \le \| f^\dagger \|,$$

which establishes that $f_n \to f^\dagger$ as $n \to \infty$. ∎

9.4 Conclusion

We have shown that the variational form of mollification fits in an extension of the generalized Tikhonov regularization setting. Using tools from variational analysis, we have obtained asymptotic consistency results for both this extended framework and the particular form of mollification that one obtains when developing the notion of target object.

Acknowledgements The author wishes to thank Nathaël Alibaud for fruitful discussions on the subject, which led to significant improvements of this paper.

References

1. Alibaud, N., Maréchal, P., Saesor, Y.: A variational approach to the inversion of truncated Fourier operators. Inverse Probl. **25**(4), 045002 (2009)
2. Bonnefond, X., Maréchal, P.: A variational approach to the inversion of some compact operators. Pac. J. Optim. **5**(1), 97–110 (2009)
3. Engl, H.W., Hanke, M., Neubauer, A.: Regularization of Inverse Problems, vol. 375. Springer Science & Business Media (1996)
4. Friedrichs, K.O.: The identity of weak and strong extensions of differential operators. Trans. Am. Math. Soc. **55**(1), 132–151 (1944)
5. Hào, D.N.: A mollification method for ill-posed problems. Numer. Math. **68**, 469–506 (1994)
6. Hohage, T., Maréchal, P., Vanhems, A.: A mollifier approach to the deconvolution of probability densities. Part 2: Convergence rates (2020). Submitted
7. Lannes, A., Roques, S., Casanove, M.J.: Stabilized reconstruction in signal and image processing: I. partial deconvolution and spectral extrapolation with limited field. J. Mod. Opt. **34**(2), 161–226 (1987)
8. Louis, A.K., Maass, P.: A mollifier method for linear operator equations of the first kind. Inverse Probl. **6**(3), 427 (1990)
9. Maréchal, P., Simar, L., Vanhems, A.: A mollifier approach to the deconvolution of probability densities. Part 1: The methodology and its comparison to classical methods (2020). Submitted
10. Maréchal, P., Simo Tao Lee, W.C., Vanhems, A.: A mollifier approach to the nonparametric instrumental regression problem (2020). Submitted
11. Maréchal, P., Togane, D., Celler, A.: A new reconstruction methodology for computerized tomography: FRECT (fourier regularized computed tomography). IEEE Trans. Nucl. Sci. **47**(4), 1595–1601 (2000)
12. Murio, D.A.: The Mollification Method and the Numerical Solution of Ill-Posed Problems. Wiley (2011)
13. Natterer, F.: The Mathematics of Computerized Tomography. SIAM (2001)
14. Schuster, T.: The Method of Approximate Inverse: Theory and Applications, vol. 1906. Springer (2007)
15. Tikhonov, A.N., Arsenin, V.Y.: Methods for Solving Ill-Posed Problems. Wiley (1977)
16. Wikipedia contributors: Mollifier—Wikipedia, the free encyclopedia. https://en.wikipedia.org/w/index.php?title=Mollifier&oldid=950509587 (2020). Accessed 17 April 2020