NUMERICAL FUNCTIONAL ANALYSIS AND OPTIMIZATION
2016, VOL. 37, NO. 5, 521–540
http://dx.doi.org/10.1080/01630563.2016.1144070

Taylor & Francis
Taylor & Francis Group

OPEN ACCESS

Optimal Convergence Rates Results for Linear Inverse Problems in Hilbert Spaces

V. Albani[a], P. Elbau[a], M. V. de Hoop[b], and O. Scherzer[a,c]

[a]Computational Science Center, University of Vienna, Vienna, Austria; [b]Department of Computational and Applied Mathematics and Department of Earth Science, Rice University, Houston, Texas, USA; [c]Johann Radon Institute for Computational and Applied Mathematics (RICAM), Linz, Austria

ABSTRACT
In this article, we prove optimal convergence rates results for regularization methods for solving linear ill-posed operator equations in Hilbert spaces. The results generalizes existing convergence rates results on optimality to general source conditions, such as logarithmic source conditions. Moreover, we also provide optimality results under variational source conditions and show the connection to approximative source conditions.

ARTICLE HISTORY
Received 9 November 2015
Revised 15 January 2016
Accepted 15 January 2016

KEYWORDS
Approximative source conditions; convergence rates; linear inverse problems; regularization; variational source conditions

MATHEMATICS SUBJECT CLASSIFICATION
47A52; 49N45; 65J22

1. Introduction

Let $L : X \to Y$ be a bounded linear operator between two Hilbert spaces X and Y. We are interested in finding the minimum-norm solution $x^\dagger \in X$ of the equation

$$Lx = y$$

for some $y \in \mathcal{R}(L)$, that is the element $x^\dagger \in \{x \in X \mid Lx = y\}$ with the property $\|x^\dagger\| = \inf\{\|x\| \mid Lx = y\}$. It is well-known that this minimal-norm solution exists and is unique, see for example [3, Theorem 2.5].

Since y is typically not exactly known and only an approximation $\tilde{y} \in Y$ with $\|y - \tilde{y}\| \le \delta$ is given, we are looking for a family $(x_\alpha(\tilde{y}))_{\alpha \ge 0}$ of approximative solutions so that for every sequence $(\tilde{y}_k)_{k\in\mathbb{N}}$ converging to y, we find a sequence $(\alpha_k)_{k\in\mathbb{N}}$ of regularization parameters such that $(x_{\alpha_k}(\tilde{y}_k))_{k\in\mathbb{N}}$ tends to the minimum-norm solution $x^\dagger$.

A standard way to construct this family is by using Tikhonov regularization:

$$x_\alpha^{\mathrm{Tik}}(\tilde{y}) = \arg\min_{x\in X}(\|Lx - \tilde{y}\|^2 + \alpha\|x\|^2),$$

CONTACT P. Elbau peter.elbau@univie.ac.at Computational Science Center, University of Vienna, Oskar-Morgenstern-Platz 1, A-1090 Vienna, Austria.

where the minimiser can be explicitly calculated from the optimality condition and reads as follows:

$$x_\alpha^{\mathrm{Tik}}(\tilde{y}) = (\alpha + L^*L)^{-1}L^*\tilde{y}. \tag{1}$$

More generally, we want to analyze regularized solutions of the form

$$x_\alpha(\tilde{y}) = r_\alpha(L^*L)L^*\tilde{y} \tag{2}$$

with some appropriately chosen function r_α, see for example [9].

The aim of this article is to characterize for a given regularization method, generated by a family $(r_\alpha)_{\alpha>0}$, the optimal convergence rate with which $x_\alpha(\tilde{y})$ tends to the minimum-norm solution $x^\dagger$. This convergence rate depends on the solution $x^\dagger$, and we will give an explicit relation between the spectral projections of $x^\dagger$ with respect to the operator L^*L and the convergence rate; first in Section 2 for the convergence of $x_\alpha(y)$ with the exact data y, and then in Section 3 for $x_\alpha(\tilde{y})$ with noisy data $\tilde{y}$. This generalizes existing convergence rates results of [10] to general source conditions, such as logarithmic source conditions.

Afterwards, we show in Section 4 that these convergence rates can also be obtained from variational inequalities and establish the optimality of these general variational source conditions, extending the results of [1]. It is interesting to note that variational source conditions are equivalent to convergence rates of the regularized solutions, while the classical results in [5] are not.

Finally, we consider in Section 5 approximate source conditions that relate the convergence rates of the regularized solutions to the decay rate of a distance function, measuring how far away the minimum-norm solution is from the classical range condition, see [4, 9]. We can show that these approximate source conditions are indeed equivalent to the convergence rates.

2. Convergence rates for exact data

In the following, we analyze the convergence rate of the sequence $(x_\alpha(y))_{\alpha>0}$ with the exact data $y \in \mathcal{R}(L)$ to the minimum-norm solution $x^\dagger$ of $Lx = y$.

We investigate regularization methods of the form (2), which are generated by functions satisfying the following properties.

Definition 2.1. We call a family $(r_\alpha)_{\alpha>0}$ of continuous functions $r_\alpha : [0,\infty) \to [0,\infty)$ the generator of a regularization method if

(i) there exists a constant $\rho \in (0,1)$ such that

$$r_\alpha(\lambda) \le \min\left\{\frac{1}{\lambda}, \frac{\rho}{\sqrt{\alpha\lambda}}\right\} \quad \text{for every } \lambda > 0,\ \alpha > 0,$$

(ii) the error function $\tilde{r}_\alpha : [0,\infty) \to [0,\infty)$, defined by

$$\tilde{r}_\alpha(\lambda) = (1 - \lambda r_\alpha(\lambda))^2, \quad \lambda \ge 0, \tag{3}$$

is decreasing.

(iii) For fixed $\lambda \geq 0$ the map $\alpha \mapsto \tilde{r}_\alpha(\lambda)$ is continuous and increasing, and
(iv) there exists a constant $\tilde{\rho} \in (0,1)$ such that

$$\tilde{r}_\alpha(\alpha) < \tilde{\rho} \quad \text{for all } \alpha > 0.$$

Remark. These conditions do not yet enforce that $x_\alpha(y) \to x^\dagger$. To ensure this, we could additionally impose that $\tilde{r}_\alpha(\lambda) \to 0$ for every $\lambda > 0$ as $\alpha \to 0$.

Let us now fix the notation for the rest of the article.

Notation 2.2. Let $L : X \to Y$ be a bounded linear operator between two real Hilbert spaces X and Y, $y \in \mathcal{R}(L)$, and $x^\dagger \in X$ be the minimum-norm solution of $Lx = y$.

We choose a generator $(r_\alpha)_{\alpha>0}$ of a regularization method, introduce the family $(\tilde{r}_\alpha)_{\alpha>0}$ of its error functions, and the corresponding family of regularized solutions shall be given by (2).

We denote by $A \mapsto E_A$ and $A \mapsto F_A$ the spectral measures of the operators L^*L and LL^*, respectively, on all Borel sets $A \subseteq [0,\infty)$.

Next, we define the right-continuous and increasing function

$$\begin{aligned} e : [0,\infty) &\to \mathbb{R}, \\ \lambda &\mapsto \|E_{[0,\lambda]}x^\dagger\|^2. \end{aligned} \tag{4}$$

Moreover, if $f : (0,\infty) \to \mathbb{R}$ is a right-continuous, increasing, and bounded function, we write

$$\int_a^b g(\lambda)\,\mathrm{d}f(\lambda) = \int_{(a,b]} g(\lambda)\,\mathrm{d}\mu_f(\lambda)$$

for the Lebesgue–Stieltjes integral of f, where μ_f denotes the unique non-negative Borel measure defined by $\mu_f((\lambda_1,\lambda_2]) = f(\lambda_2) - f(\lambda_1)$ and $g \in L^1(\mu)$.

Remark. In this setting, we can write the error

$$x_\alpha(y) - x^\dagger = r_\alpha(L^*L)L^*y - x^\dagger = (r_\alpha(L^*L)L^*L - I)x^\dagger \tag{5}$$

according to spectral theory in the form

$$\|x_\alpha(y) - x^\dagger\|^2 = \int_0^{\|L\|^2} \tilde{r}_\alpha(\lambda)\,\mathrm{d}e(\lambda). \tag{6}$$

We want to point out here that it directly follows from the definition that the minimum-norm solution $x^\dagger$ is in the orthogonal complement $\mathcal{N}(L)^\perp$ of the nullspace of L, and we therefore do not have to consider the point $\lambda = 0$ in the integrals in equation (6).

We first want to establish a relation between the convergence rate of the regularized solution $x_\alpha(y)$ for exact data y to the minimum-norm solution $x^\dagger$ and the behaviour of the spectral function (4).

Proposition 2.3. *We use Notation 2.2 and assume that there exist an increasing function $\varphi : (0,\infty) \to (0,\infty)$ and constants $\mu \in (0,1)$ and $A > 0$ such that we have for every $\alpha > 0$ the inequality*

$$\varphi(\lambda)\tilde{r}_\alpha^\mu(\lambda) \le A\varphi(\alpha) \quad \text{for all } \lambda > 0. \tag{7}$$

Then, the following two statements are equivalent:

(i) *There exists a constant $C > 0$ with*

$$\|x_\alpha(y) - x^\dagger\|^2 \le C\varphi(\alpha) \quad \text{for all } \alpha > 0. \tag{8}$$

(ii) *There exists a constant $\tilde{C} > 0$ with*

$$e(\lambda) \le \tilde{C}\varphi(\lambda) \quad \text{for all } \lambda > 0. \tag{9}$$

Proof. According to Definition 2.1 (ii) the error function $\tilde{r}_\alpha$ is decreasing, and thus it follows together with (6) that for all $\alpha > 0$

$$\tilde{r}_\alpha(\alpha)e(\alpha) = \tilde{r}_\alpha(\alpha)\int_0^\alpha \mathrm{d}e(\lambda) \le \int_0^\alpha \tilde{r}_\alpha(\lambda)\,\mathrm{d}e(\lambda) \le \|x_\alpha(y) - x^\dagger\|^2. \tag{10}$$

Let first (8) hold. Then, it follows from (10) that for all $\alpha > 0$

$$\tilde{r}_\alpha(\alpha)e(\alpha) \le C\varphi(\alpha). \tag{11}$$

Now, we use Definition 2.1 (i), which gives that

$$\tilde{r}_\alpha(\alpha) = (1 - \alpha r_\alpha(\alpha))^2 \ge (1-\rho)^2 > 0.$$

Using this estimate in (11) yields (9) with $\tilde{C} = \frac{C}{(1-\rho)^2} > 0$.

Conversely, let (9) hold. Since $\|x_\alpha(y) - x^\dagger\|^2 \le \|x^\dagger\|^2$ (which follows from (6) with $\tilde{r}_\alpha \le 1$), it is enough to check the condition (8) for all $\alpha \in (0, \|L\|^2]$.

We use (6) and integrate the right hand side by parts, see for example [2, Theorem 6.2.2] regarding the integration by parts for Lebesgue–Stieltjes integrals, and obtain that

$$\|x_\alpha(y) - x^\dagger\|^2 = \tilde{r}_\alpha(\|L\|^2)e(\|L\|^2) + \int_0^{\|L\|^2} e(\lambda)\,\mathrm{d}(-\tilde{r}_\alpha)(\lambda). \tag{12}$$

We split up the integral on the right hand side into two terms:

$$\int_0^{\|L\|^2} e(\lambda)\,\mathrm{d}(-\tilde{r}_\alpha)(\lambda) = \int_0^\alpha e(\lambda)\,\mathrm{d}(-\tilde{r}_\alpha)(\lambda) + \int_\alpha^{\|L\|^2} e(\lambda)\,\mathrm{d}(-\tilde{r}_\alpha)(\lambda). \tag{13}$$

The first term is estimated by using that the function e is increasing and by utilising the assumption (9):

$$\int_0^\alpha e(\lambda)\,\mathrm{d}(-\tilde{r}_\alpha)(\lambda) \le e(\alpha)\int_0^\alpha \mathrm{d}(-\tilde{r}_\alpha)(\lambda) = e(\alpha)(1 - \tilde{r}_\alpha(\alpha)) \le \tilde{C}\varphi(\alpha).$$

The second integral term in (13) is estimated by using the inequalities (9) and (7):

$$\begin{aligned}\int_\alpha^{\|L\|^2} e(\lambda)\,\mathrm{d}(-\tilde r_\alpha)(\lambda) &\le \tilde C \int_\alpha^{\|L\|^2} \varphi(\lambda)\,\mathrm{d}(-\tilde r_\alpha)(\lambda)\\ &= \tilde C \int_\alpha^{\|L\|^2} \varphi(\lambda)\tilde r_\alpha^{\mu}(\lambda)\frac{1}{\tilde r_\alpha^{\mu}(\lambda)}\,\mathrm{d}(-\tilde r_\alpha)(\lambda)\\ &\le A\tilde C\varphi(\alpha)\int_\alpha^{\|L\|^2} \frac{1}{\tilde r_\alpha^{\mu}(\lambda)}\,\mathrm{d}(-\tilde r_\alpha)(\lambda)\\ &= \frac{A\tilde C}{1-\mu}\varphi(\alpha)(\tilde r_\alpha^{1-\mu}(\alpha)-\tilde r_\alpha^{1-\mu}(\|L\|^2))\\ &\le \frac{A\tilde C\tilde\rho^{1-\mu}}{1-\mu}\varphi(\alpha),\end{aligned}$$

where we used Definition 2.1 (iv) in the last step. Inserting the two estimates in (13) and in (12), we find with $e(\|L\|^2) = \|x^\dagger\|^2$ that

$$\|x_\alpha(y)-x^\dagger\|^2 \le \tilde r_\alpha(\|L\|^2)\|x^\dagger\|^2 + \tilde C\varphi(\alpha) + \frac{A\tilde C\tilde\rho^{1-\mu}}{1-\mu}\varphi(\alpha). \tag{14}$$

From (7), we deduce further that

$$\tilde r_\alpha(\|L\|^2) \le \frac{A^{\frac{1}{\mu}}}{\varphi^{\frac{1}{\mu}}(\|L\|^2)}\varphi^{\frac{1}{\mu}}(\alpha) \le \frac{A^{\frac{1}{\mu}}\varphi^{\frac{1}{\mu}-1}(\alpha)}{\varphi^{\frac{1}{\mu}}(\|L\|^2)}\varphi(\alpha) \le c\varphi(\alpha)$$
$$\text{with } c = \frac{A^{\frac{1}{\mu}}}{\varphi(\|L\|^2)},$$

since φ is increasing and $\mu < 1$.

Thus, we get from (14) that

$$\|x_\alpha(y)-x^\dagger\|^2 \le C\varphi(\alpha)$$

with

$$C = c\|x^\dagger\|^2 + \tilde C + \frac{A\tilde C\tilde\rho^{1-\mu}}{1-\mu}.$$

□

Remark. The condition (7) with the choice $\mu = \frac{1}{2}$ was already used in [4], and such a function φ was called a qualification of the regularization method.

Example 2.4. In the case of Tikhonov regularization, given by (1), we have $r_\alpha(\lambda) = \frac{1}{\alpha+\lambda}$ and therefore we get for the error function $\tilde r_\alpha$, defined by (3), the expression $\tilde r_\alpha(\lambda) = \frac{\alpha^2}{(\alpha+\lambda)^2}$. So, clearly, $\tilde r_\alpha(\alpha) = \frac{1}{4}$ and all the conditions of Definition 2.1 are fulfilled.

(i) To recover the classical equivalence results, see [10, Theorem 2.1], we set $\varphi(\alpha) = \alpha^{2\nu}$ for some $\nu \in (0,1)$ and find that the condition (7) with $A = 1$ is for every $\mu \geq \nu$ fulfilled, since we have

$$\varphi(\lambda)\tilde{r}_\alpha^\mu(\lambda) = \frac{\alpha^{2\mu}\lambda^{2\nu}}{(\alpha+\lambda)^{2\mu}} \leq \frac{\alpha^{2\mu-2\nu}}{(\alpha+\lambda)^{2\mu-2\nu}}\frac{\lambda^{2\nu}}{(\alpha+\lambda)^{2\nu}}\alpha^{2\nu} \leq \alpha^{2\nu} = \varphi(\alpha)$$

for arbitrary $\alpha > 0$ and $\lambda > 0$.

Thus, Proposition 2.3 yields for every $\nu \in (0,1)$ the equivalence of $\|x_\alpha(y) - x^\dagger\|^2 = \mathcal{O}(\alpha^{2\nu})$ and $e(\lambda) = \mathcal{O}(\lambda^{2\nu})$.

(ii) Similarly, we also get the equivalence in the case of logarithmic convergence rates. Let $0 < \nu < \mu < 1$ and define for $\alpha \in (0, \mathrm{e}^{-\frac{\nu}{\mu}}]$ the function $\varphi(\alpha) = \left|\log\alpha\right|^{-\nu}$ (for bigger values of α, we may simply set $\varphi(\alpha) = \varphi(\mathrm{e}^{-\frac{\nu}{\mu}})$). Then, we have

$$\begin{aligned}(\varphi\tilde{r}_\alpha^\mu)'(\lambda) &= \frac{\alpha^{2\mu}}{(\alpha+\lambda)^{2\mu+1}\left|\log\lambda\right|^{\nu+1}}\left(\nu\frac{\alpha+\lambda}{\lambda} - 2\mu\left|\log\lambda\right|\right)\\ &\leq -\frac{2(\mu\left|\log\lambda\right| - \nu)\alpha^{2\mu}}{(\alpha+\lambda)^{2\mu+1}\left|\log\lambda\right|^{\nu+1}} \leq 0\end{aligned}$$

for all $\lambda \in [\alpha, \mathrm{e}^{-\frac{\nu}{\mu}})$. Thus, $\varphi\tilde{r}_\alpha^\mu$ is decreasing on $[\alpha, \mathrm{e}^{-\frac{\nu}{\mu}})$, which implies (7) with $A = 1$.

So, Proposition 2.3 tells us that $\|x_\alpha(y) - x^\dagger\|^2 = \mathcal{O}(\left|\log\alpha\right|^{-\nu})$ if and only if $e(\lambda) = \mathcal{O}(\left|\log\lambda\right|^{-\nu})$.

3. Convergence rates for noisy data

We now want to estimate the distance of the regularized solution $x_\alpha(\tilde{y})$ to the minimum-norm solution $x^\dagger$ if we do not have the exact data y, but only some approximation $\tilde{y}$ of it.

In this case, we consider the regularization parameter α as a function of the noisy data $\tilde{y}$ such that the distance between $x_\alpha(\tilde{y})$ and $x^\dagger$ is minimal. Thus, we are interested in the convergence rate of the expression $\inf_{\alpha>0}\|x_\alpha(\tilde{y}) - x^\dagger\|$ to zero as the distance between $\tilde{y}$ and y tends to zero. We therefore want to find an upper bound for the expression $\sup_{\tilde{y}\in\bar{B}_\delta(y)}\inf_{\alpha>0}\|x_\alpha(\tilde{y}) - x^\dagger\|$, where $\bar{B}_\delta(y) = \{\tilde{y} \in Y \mid \|\tilde{y} - y\| \leq \delta\}$ denotes the closed ball with radius $\delta > 0$ around the data y.

Let us first consider the trivial case where $\|x_\alpha(y) - x^\dagger\| = 0$ for all α in a vicinity of 0.

Lemma 3.1. *We use Notation 2.2 and assume that there exists an $\varepsilon > 0$ such that*

$$\|x_\alpha(y) - x^\dagger\| = 0 \quad \textit{for all } \alpha \in (0,\varepsilon].$$

Then, we have

$$\sup_{\tilde{y}\in\bar{B}_\delta(y)} \inf_{\alpha>0} \|x_\alpha(\tilde{y}) - x^\dagger\|^2 \le \frac{\rho^2}{\varepsilon}\delta^2, \tag{15}$$

where $\rho > 0$ is chosen as in Definition 2.1 (i).

Proof. Let $\tilde{y} \in \bar{B}_\delta(y)$ be fixed. Then, using that $Lr_\alpha(L^*L) = r_\alpha(LL^*)L$, it follows from Definition 2.1 (i) that

$$\|x_\alpha(\tilde{y}) - x_\alpha(y)\|^2 = \left\langle \tilde{y} - y, r_\alpha^2(LL^*)LL^*(\tilde{y} - y)\right\rangle \le \delta^2 \max_{\lambda>0} \lambda r_\alpha^2(\lambda) \le \rho^2\frac{\delta^2}{\alpha}. \tag{16}$$

The right hand side is uniform for all $\tilde{y} \in \bar{B}_\delta(y)$. Thus, picking $\alpha = \varepsilon$, we get

$$\sup_{\tilde{y}\in\bar{B}_\delta(y)} \inf_{\alpha>0} \|x_\alpha(\tilde{y}) - x^\dagger\|^2 \le \inf_{\alpha>0}\left(\|x_\alpha(y) - x^\dagger\| + \rho\frac{\delta}{\sqrt{\alpha}}\right)^2 \le \frac{\rho^2}{\varepsilon}\delta^2,$$

which is (15). □

In the general case, we estimate the optimal regularization parameter α to be in the vicinity of the value α_δ, which is chosen as the solution of the implicit equation (17) and is therefore only depending on the distance δ between the correct data y and the noisy data $\tilde{y}$.

Lemma 3.2. *We use again Notation 2.2 and consider the case where $\|x_\alpha(y) - x^\dagger\| > 0$ for all $\alpha > 0$.*

If we choose for every $\delta > 0$ the parameter $\alpha_\delta > 0$ such that

$$\alpha_\delta\|x_{\alpha_\delta}(y) - x^\dagger\|^2 = \delta^2, \tag{17}$$

then there exists a constant $C_1 > 0$ such that

$$\sup_{\tilde{y}\in\bar{B}_\delta(y)} \inf_{\alpha>0} \|x_\alpha(\tilde{y}) - x^\dagger\|^2 \le C_1\frac{\delta^2}{\alpha_\delta} \quad \text{for all } \delta > 0. \tag{18}$$

Moreover, there exists a constant $C_0 > 0$ such that

$$\sup_{\tilde{y}\in\bar{B}_\delta(y)} \inf_{\alpha>0} \|x_\alpha(\tilde{y}) - x^\dagger\|^2 \ge C_0\frac{\delta^2}{\alpha_\delta} \tag{19}$$

for all $\delta > 0$ which fulfil that $\alpha_\delta \in \sigma(LL^*)$, where $\sigma(LL^*) \subset [0,\infty)$ denotes the spectrum of the operator LL^*.

Proof. First, we remark that the function

$$A : (0,\infty) \to (0,\infty), \quad A(\alpha) = \alpha\|x_\alpha(y) - x^\dagger\|^2 = \int_0^{\|L\|^2} \alpha\tilde{r}_\alpha(\lambda)\,\mathrm{d}e(\lambda)$$

is, according to Definition 2.1 (iii) together with the assumption that $\|x_\alpha(y) - x^\dagger\| > 0$ for all $\alpha > 0$, continuous and strictly increasing and satisfies $\lim_{\alpha\to 0} A(\alpha) = 0$ and $\lim_{\alpha\to\infty} A(\alpha) = \infty$. Therefore, we find for every $\delta > 0$ a unique value $\alpha_\delta = A^{-1}(\delta^2)$.

Let $\tilde{y} \in \bar{B}_\delta(y)$. Then, as in the proof of Lemma 3.1, see (16), we find that

$$\|x_\alpha(\tilde{y}) - x_\alpha(y)\|^2 \le \rho^2 \frac{\delta^2}{\alpha}.$$

From this estimate, we obtain with the triangular inequality and with the definition (17) of α_δ that

$$\sup_{\tilde{y}\in\bar{B}_\delta(y)} \inf_{\alpha>0} \|x_\alpha(\tilde{y}) - x^\dagger\|^2 \le \inf_{\alpha>0} \left(\|x_\alpha(y) - x^\dagger\| + \rho\frac{\delta}{\sqrt{\alpha}} \right)^2 \le (1+\rho)^2 \frac{\delta^2}{\alpha_\delta},$$

which is the upper bound (18) with the constant $C_1 = (1+\rho)^2$.

For the lower bound (19), we write similarly

$$\begin{aligned} \|x_\alpha(\tilde{y}) - x^\dagger\|^2 &= \|x_\alpha(y) - x^\dagger\|^2 + \|x_\alpha(\tilde{y}) - x_\alpha(y)\|^2 \\ &\quad + 2\left\langle x_\alpha(\tilde{y}) - x_\alpha(y), x_\alpha(y) - x^\dagger \right\rangle \\ &= \|x_\alpha(y) - x^\dagger\|^2 + \left\langle \tilde{y} - y, r_\alpha^2(LL^*)LL^*(\tilde{y} - y) \right\rangle \\ &\quad + 2\left\langle r_\alpha(LL^*)(\tilde{y} - y), r_\alpha(LL^*)LL^*y - y \right\rangle. \end{aligned} \tag{20}$$

Now, from the continuity of $\tilde{r}_{\alpha_\delta}$ and Definition 2.1 (iv), we find that for every $\delta > 0$ there exists a parameter $a_\delta \in (0, \alpha_\delta)$ such that $\tilde{r}_{\alpha_\delta}(a_\delta) < \tilde{\rho}$.

Then, the assumption $\alpha_\delta \in \sigma(LL^*)$ implies that the spectral measure F of the operator LL^* fulfils $F_{[a_\delta, 2\alpha_\delta]} \ne 0$.

Suppose now that

$$z_\delta = F_{[a_\delta, 2\alpha_\delta]}(r_{\alpha_\delta}(LL^*)LL^*y - y) \ne 0. \tag{21}$$

Then, choosing $\tilde{y} = y + \delta \frac{z_\delta}{\|z_\delta\|}$, equation (20) becomes

$$\|x_\alpha(\tilde{y}) - x^\dagger\|^2 = \|x_\alpha(y) - x^\dagger\|^2 + \frac{\delta^2}{\|z_\delta\|^2} \left\langle z_\delta, r_\alpha^2(LL^*)LL^*z_\delta \right\rangle + \frac{2\delta}{\|z_\delta\|} \left\langle r_\alpha(LL^*)z_\delta, z_\delta \right\rangle.$$

Thus, we may drop the last term as it is non-negative, which gives us the lower bound

$$\sup_{\tilde{y}\in\bar{B}_\delta(y)} \inf_{\alpha>0} \|x_\alpha(\tilde{y}) - x^\dagger\|^2 \ge \inf_{\alpha>0} \left(\|x_\alpha(y) - x^\dagger\|^2 + \delta^2 \min_{\lambda\in[a_\delta, 2\alpha_\delta]} \lambda r_\alpha^2(\lambda) \right).$$

Since we get from Definition 2.1 (ii) the inequality

$$\lambda r_\alpha^2(\lambda) = \frac{\left(1 - \sqrt{\tilde{r}_\alpha(\lambda)}\right)^2}{\lambda} \ge \frac{\left(1 - \sqrt{\tilde{r}_\alpha(a_\delta)}\right)^2}{2\alpha_\delta} \quad \text{for all } \lambda \in [a_\delta, 2\alpha_\delta],$$

we can estimate further

$$\sup_{\tilde{y}\in\bar{B}_\delta(y)}\inf_{\alpha>0}\|x_\alpha(\tilde{y})-x^\dagger\|^2 \geq \inf_{\alpha>0}\left(\|x_\alpha(y)-x^\dagger\|^2+\delta^2\frac{\left(1-\sqrt{\tilde{r}_\alpha(a_\delta)}\right)^2}{2\alpha_\delta}\right).$$

Now, since $\alpha \mapsto \tilde{r}_\alpha(\lambda)$ is for every $\lambda > 0$ increasing, see Definition 2.1 (iii), the first term is increasing in α, see (6), and the second term is decreasing in α. Thus, we can estimate the expression for $\alpha < \alpha_\delta$ from below by the second term at $\alpha = \alpha_\delta$, and for $\alpha \geq \alpha_\delta$ by the first term at $\alpha = \alpha_\delta$:

$$\begin{aligned}\sup_{\tilde{y}\in\bar{B}_\delta(y)}\inf_{\alpha>0}\|x_\alpha(\tilde{y})-x^\dagger\|^2 &\geq \min\left\{\|x_{\alpha_\delta}(y)-x^\dagger\|^2, \delta^2\frac{\left(1-\sqrt{\tilde{r}_{\alpha_\delta}(a_\delta)}\right)^2}{2\alpha_\delta}\right\}\\ &\geq \frac{(1-\sqrt{\tilde{\rho}})^2}{2}\frac{\delta^2}{\alpha_\delta},\end{aligned}$$

which is (19) with $C_0 = \frac{1}{2}(1-\sqrt{\tilde{\rho}})^2$.

If z_δ, as defined by (21), happens to vanish, the same argument works with an arbitrary non-zero element $z_\delta \in \mathcal{R}(F_{[a_\delta,2\alpha_\delta]})$ since the last term in (20) is zero for $\tilde{y} = y + \delta\frac{z_\delta}{\|z_\delta\|}$. □

From Lemma 3.1 and Lemma 3.2, we now get an equivalence relation between the noisy and the noise-free convergence rates.

Proposition 3.3. *We use Notation 2.2. Let further $\varphi : [0,\infty) \to [0,\infty)$ be a strictly increasing function satisfying $\varphi(0) = 0$ and*

$$\varphi(\gamma\alpha) \leq g(\gamma)\varphi(\alpha) \quad \text{for all } \alpha > 0,\ \gamma > 0 \tag{22}$$

for some increasing function $g : (0,\infty) \to (0,\infty)$.

Moreover, we assume that there exists a constant $C > 0$ with

$$\frac{\tilde{r}_\alpha(\lambda)}{\tilde{r}_\beta(\lambda)} \leq C\frac{\varphi(\alpha)}{\varphi(\beta)} \quad \text{for all } 0 < \alpha \leq \beta \leq \lambda \tag{23}$$

and there is a constant $\tilde{C}$ such that

$$\frac{\tilde{r}_\alpha(\lambda)}{\tilde{r}_\beta(\lambda)} \geq \tilde{C}\frac{\varphi(\alpha)}{\varphi(\beta)} \quad \text{for all } 0 < \lambda \leq \alpha \leq \beta. \tag{24}$$

We define

$$\tilde{\varphi}(\alpha) = \sqrt{\alpha\varphi(\alpha)} \quad \text{and} \quad \psi(\delta) = \frac{\delta^2}{\tilde{\varphi}^{-1}(\delta)}. \tag{25}$$

Then, the following two statements are equivalent:

(i) *There exists a constant $c > 0$ such that*

$$\sup_{\tilde{y}\in\bar{B}_\delta(y)}\inf_{\alpha>0}\|x_\alpha(\tilde{y})-x^\dagger\|^2 \leq c\psi(\delta) \quad \text{for all } \delta > 0. \tag{26}$$

(ii) *There exists a constant* $\tilde{c} > 0$ *such that*

$$\|x_\alpha(y) - x^\dagger\|^2 \le \tilde{c}\varphi(\alpha) \quad \textit{for all } \alpha > 0. \tag{27}$$

Proof. We first remark that (22) implies that $\tilde{\varphi}(\gamma\alpha) \le \sqrt{\gamma g(\gamma)}\tilde{\varphi}(\alpha)$, and so by setting $\tilde{g}(\gamma) = \sqrt{\gamma g(\gamma)}$, $\delta = \tilde{\varphi}(\alpha)$ and $\tilde{\gamma} = \tilde{g}(\gamma)$, we get

$$\tilde{g}^{-1}(\tilde{\gamma})\tilde{\varphi}^{-1}(\delta) \le \tilde{\varphi}^{-1}(\tilde{\gamma}\delta).$$

Thus, we have

$$\psi(\tilde{\gamma}\delta) = \frac{\tilde{\gamma}^2\delta^2}{\tilde{\varphi}^{-1}(\tilde{\gamma}\delta)} \le \frac{\tilde{\gamma}^2\delta^2}{\tilde{g}^{-1}(\tilde{\gamma})\tilde{\varphi}^{-1}(\delta)} = h(\tilde{\gamma})\psi(\delta) \tag{28}$$

where $h(\tilde{\gamma}) = \frac{\tilde{\gamma}^2}{\tilde{g}^{-1}(\tilde{\gamma})}$.

In the case where $\|x_\alpha(y) - x^\dagger\| = 0$ for all $\alpha \in (0, \varepsilon]$ for some $\varepsilon > 0$, the inequality (27) is trivially fulfilled for some $\tilde{c} > 0$. Moreover, we know from Lemma 3.1 that then the inequality (15) holds, which implies the inequality (26) for some constant $c > 0$, since we have according to the definition of the function ψ that $\psi(\delta) \ge a\delta^2$ for all $\delta \in (0, \delta_0)$ for some constants $a > 0$ and $\delta_0 > 0$.

Thus, we may assume that $\|x_\alpha(y) - x^\dagger\| > 0$ for all $\alpha > 0$.

Let (27) hold. For arbitrary $\delta > 0$, we use the regularization parameter α_δ defined in (17). Then, the inequality (27) implies that

$$\frac{\delta^2}{\alpha_\delta} \le \tilde{c}\varphi(\alpha_\delta).$$

Consequently,

$$\tilde{\varphi}^{-1}\left(\frac{\delta}{\sqrt{\tilde{c}}}\right) \le \alpha_\delta,$$

and therefore, using the inequality (18) obtained in Lemma 3.2, we find with (28) that

$$\sup_{\tilde{y} \in \bar{B}_\delta(y)} \inf_{\alpha > 0} \|x_\alpha(\tilde{y}) - x^\dagger\|^2 \le C_1 \frac{\delta^2}{\alpha_\delta} \le C_1 \tilde{c}\psi\left(\frac{\delta}{\sqrt{\tilde{c}}}\right) \le C_1 \tilde{c} h(\tfrac{1}{\sqrt{\tilde{c}}})\psi(\delta),$$

which is the estimate (26) with $c = C_1\tilde{c}h(\frac{1}{\sqrt{\tilde{c}}})$.

Conversely, if (26) holds, we choose an arbitrary $\delta > 0$ such that α_δ defined by (17) is in the spectrum $\sigma(LL^*)$. Then, we can use the inequality (19) of Lemma 3.2 to obtain from the condition (26) that

$$C_0 \frac{\delta^2}{\alpha_\delta} \le c\psi(\delta).$$

Thus, by the definition of ψ, we have

$$\tilde{\varphi}^{-1}(\delta) \le \frac{c}{C_0}\alpha_\delta.$$

So, finally, we get with (22) that

$$\|x_{\alpha_\delta}(y) - x^\dagger\|^2 = \frac{\delta^2}{\alpha_\delta} \leq \frac{c}{C_0}\varphi\left(\frac{c}{C_0}\alpha_\delta\right) \leq \frac{c}{C_0} g(\tfrac{c}{C_0})\varphi(\alpha_\delta),$$

and since this holds for every δ such that $\alpha_\delta \in \sigma(LL^*)$, we have with $\hat{c} = \frac{c}{C_0} g(\frac{c}{C_0})$ that

$$\|x_\alpha(y) - x^\dagger\|^2 \leq \hat{c}\varphi(\alpha) \quad \text{for all } \alpha \in \sigma(LL^*). \tag{29}$$

Finally, we consider some $\alpha \notin \sigma(LL^*)$, $\alpha < \|L\|^2$, and set

$$\alpha_- = \sup\{\tilde{\alpha} \in \sigma(LL^*) \cup \{0\} \mid \tilde{\alpha} < \alpha\} \quad \text{and}$$
$$\alpha_+ = \inf\{\tilde{\alpha} \in \sigma(LL^*) \mid \tilde{\alpha} > \alpha\}.$$

Then, recalling that $\sigma(L^*L)\setminus\{0\} = \sigma(LL^*)\setminus\{0\}$, see for example [6, Problem 61], we find for $\alpha_- > 0$ (for $\alpha_- = 0$, the first term in the following calculation simply vanishes) that

$$\begin{aligned}\|x_\alpha(y) - x^\dagger\|^2 &= \int_0^{\alpha_-} \tilde{r}_\alpha(\lambda)\, de(\lambda) + \int_{\alpha_+}^{\|L\|^2} \tilde{r}_\alpha(\lambda)\, de(\lambda) \\ &\leq \|x_{\alpha_-}(y) - x^\dagger\|^2 \sup_{\lambda\in[0,\alpha_-]} \frac{\tilde{r}_\alpha(\lambda)}{\tilde{r}_{\alpha_-}(\lambda)} \\ &\quad + \|x_{\alpha_+}(y) - x^\dagger\|^2 \sup_{\lambda\in[\alpha_+,\|L\|^2]} \frac{\tilde{r}_\alpha(\lambda)}{\tilde{r}_{\alpha_+}(\lambda)}.\end{aligned}$$

Using the conditions (23) and (24), we have with (29) that

$$\|x_\alpha(y) - x^\dagger\|^2 \leq \frac{\hat{c}}{\tilde{C}}\varphi(\alpha_-)\frac{\varphi(\alpha)}{\varphi(\alpha_-)} + C\hat{c}\varphi(\alpha_+)\frac{\varphi(\alpha)}{\varphi(\alpha_+)} = (C + \tfrac{1}{\tilde{C}})\hat{c}\varphi(\alpha),$$

which is (27) with $\tilde{c} = (C + \frac{1}{\tilde{C}})\hat{c}$. □

Remark. If we consider Tikhonov regularization, then we can ignore the conditions (23) and (24) in Proposition 3.3 if we have a quadratic upper bound on the function g in (22).

Indeed, let $\varphi : [0,\infty) \to [0,\infty)$ be an arbitrary increasing function fulfilling (22) for some increasing function $g : (0,\infty) \to (0,\infty)$ which is bounded by

$$g(\gamma) \leq \frac{C}{4}(1 + \gamma^2) \quad \text{for all } \gamma > 0 \tag{30}$$

for some constant $C > 0$. Then, conditions (23) and (24) are fulfilled for the error function $\tilde{r}_\alpha$ of Tikhonov regularization, given by $\tilde{r}_\alpha(\lambda) = \frac{\alpha^2}{(\alpha+\lambda)^2}$.

To see this, we remark that for $0 < \alpha \leq \beta$, the ratio

$$\frac{\tilde{r}_\alpha(\lambda)}{\tilde{r}_\beta(\lambda)} = \left(\frac{\alpha}{\beta}\frac{\beta+\lambda}{\alpha+\lambda}\right)^2$$

is decreasing in λ. Therefore, for $\lambda \geq \beta$ we get that

$$\frac{\tilde{r}_\alpha(\lambda)}{\tilde{r}_\beta(\lambda)} \leq \frac{4}{(1+\frac{\beta}{\alpha})^2} \leq \frac{4}{1+(\frac{\beta}{\alpha})^2} \leq \frac{C}{g(\frac{\beta}{\alpha})} \leq C\frac{\varphi(\alpha)}{\varphi(\beta)},$$

which is (23).

We similarly find for $\lambda \leq \alpha$ that

$$\frac{\tilde{r}_\alpha(\lambda)}{\tilde{r}_\beta(\lambda)} \geq \frac{1}{4}\left(1+\frac{\alpha}{\beta}\right)^2 \geq \frac{1}{4} \geq \frac{1}{4}\frac{\varphi(\alpha)}{\varphi(\beta)},$$

which is (24) with $\tilde{C} = \frac{1}{4}$.

We want to apply this theorem now to the two special cases discussed previously in Example 2.4.

Example 3.4. (i) In the case of Example 2.4 (i), where we considered Tikhonov regularization with a convergence rate given by $\varphi(\alpha) = \alpha^{2\nu}$ for some $\nu \in (0,1)$, the condition (22) in Proposition 3.3 is clearly fulfilled with $g(\gamma) = \gamma^{2\nu}$. In particular, g satisfies that $g(\gamma) \leq 1+\gamma^2$, which is (30) with $C = 4$, and thus the conditions (23) and (24) in Proposition 3.3 follow as in the remark above.

So, we can apply Proposition 3.3 and it only remains to calculate

$$\tilde{\varphi}^{-1}(\delta) = \delta^{\frac{2}{2\nu+1}} \quad \text{and} \quad \psi(\delta) = \delta^{2-\frac{2}{2\nu+1}} = \delta^{\frac{4\nu}{2\nu+1}}.$$

Thus, we recover the classical result, see [10, Theorem 2.6], that the convergence rate $\|x_\alpha(y) - x^\dagger\|^2 = \mathcal{O}(\alpha^{2\nu})$ for the correct data y is equivalent to the convergence rate $\sup_{\tilde{y}\in\bar{B}_\delta(y)} \inf_{\alpha>0} \|x_\alpha(\tilde{y}) - x^\dagger\|^2 = \mathcal{O}(\delta^{\frac{4\nu}{2\nu+1}})$ for noisy data.

(ii) Next, we look at Tikhonov regularization with the logarithmic convergence rate

$$\varphi(\alpha) = \begin{cases} \left|\log\alpha\right|^{-\nu} & \text{if } 0 < \alpha < \mathrm{e}^{-(1+\nu)}, \\ (1+\nu)^{-\nu} & \text{if } \alpha \geq \mathrm{e}^{-(1+\nu)}, \end{cases}$$

see Example 2.4 (ii). First, we remark that φ is concave. This is because φ is increasing, constant for $\alpha > \mathrm{e}^{-(1+\nu)}$, and for $0 < \alpha < \mathrm{e}^{-(1+\nu)}$ we have

$$\varphi''(\alpha) = \frac{\nu}{\alpha^2}\left|\log\alpha\right|^{-(\nu+2)}\left(1+\nu-\left|\log\alpha\right|\right) < 0,$$

and because $\varphi(0) = 0$, we have

$$\varphi(\gamma\alpha) \leq \gamma\varphi(\alpha) \quad \text{for all } \gamma \geq 1,\ \alpha > 0.$$

Thus, using that φ is increasing, the requirement (22) in Proposition 3.3 is fulfilled with

$$g(\gamma) = \begin{cases} 1 & \text{if } \gamma < 1, \\ \gamma & \text{if } \gamma \geq 1. \end{cases}$$

In particular, this function g satisfies the inequality (30) with $C = 4$ and therefore, also the conditions (23) and (24) in Proposition 3.3 are fulfilled according to the previous remark.

To get the corresponding function ψ, as defined in (25), we have to solve the implicit equation $\delta = \tilde{\varphi}(\frac{\delta^2}{\psi(\delta)})$, where $\tilde{\varphi}$ is defined in (25) and with the specific choice of $\varphi(\alpha) = \left|\log\alpha\right|^{-\nu}$ for $\alpha < \mathrm{e}^{-(1+\nu)}$ satisfies $\tilde{\varphi}^2(\alpha) = \alpha\left|\log\alpha\right|^{-\nu}$. This equation then reads as follows:

$$\psi(\delta) = \left|\log\frac{\delta^2}{\psi(\delta)}\right|^{-\nu}. \tag{31}$$

By solving this equation for δ, we get

$$\delta = \sqrt{\psi(\delta)}\exp\left(-\frac{1}{2\psi^{\frac{1}{\nu}}(\delta)}\right),$$

which, in particular, shows that the function ψ is increasing and furthermore, because of $\lim_{\delta\downarrow 0}\psi(\delta) = 0$, $\psi(\delta) < 1$ for sufficiently small $\delta > 0$. Therefore, we find for small $\delta > 0$ that

$$\delta \le \exp\left(-\frac{1}{2\psi^{\frac{1}{\nu}}(\delta)}\right), \quad \text{that is} \quad \psi(\delta) \ge |2\log\delta|^{-\nu}. \tag{32}$$

Moreover, if we write ψ as

$$\psi(\delta) = \left|\log\delta\right|^{-\nu} f(\delta)$$

for some function f, the implicit equation (31) becomes

$$f(\delta) = \left|\frac{\log\delta}{\log(f(\delta)) - 2\log\delta - \log(|\log\delta|^{\nu})}\right|^{\nu}.$$

Since $\lim_{\delta\downarrow 0}\frac{\log(\left|\log\delta\right|^{\nu})}{\log\delta} = 0$, we find parameters $\varepsilon \in (0,1)$ and $\delta_0 \in (0,1)$ such that we have for all $\delta < \delta_0$ the inequality $0 \le \log(\left|\log\delta\right|^{\nu}) \le \varepsilon\left|\log\delta\right|$. Assuming that $f(\delta) \ge 1$ gives

$$f(\delta) \le \left(\frac{\left|\log\delta\right|}{\log(f(\delta)) + (2-\varepsilon)\left|\log\delta\right|}\right)^{\nu} \le \frac{1}{(2-\varepsilon)^{\nu}} < 1,$$

which is a contradiction to the assumption. Thus, $f(\delta) < 1$.

Since we know already from (32) that $f(\delta) \ge 2^{-\nu}$, it therefore follows from Proposition 3.3 that the convergence rate $\|x_\alpha(y) - x^\dagger\|^2 = \mathcal{O}(\left|\log\alpha\right|^{-\nu})$ is equivalent to $\sup_{\tilde{y}\in\bar{B}_\delta(y)}\inf_{\alpha>0}\|x_\alpha(\tilde{y}) - x^\dagger\|^2 = \mathcal{O}(\left|\log\delta\right|^{-\nu})$.

4. Relation to variational inequalities

Instead of characterizing the convergence rate of the regularized solution via the behavior of the spectral decomposition of the minimum-norm solution $x^\dagger$, we may also check variational inequalities for the element $x^\dagger$, see [7, 8, 11, 12]. In [1], it was shown that for Tikhonov regularization and convergence rates of the order $\mathcal{O}(\alpha^{2\nu})$, $\nu \in (0,1)$, such variational inequalities are equivalent to specific convergence rates.

In this section, we generalize this result to cover general regularization methods and convergence rates.

Proposition 4.1. *We consider again the setting of Notation 2.2. Moreover, let $\varphi : [0,\infty) \to [0,\infty)$ be an increasing, continuous function and $\nu \in (0,1)$.*

Then, the following two statements are equivalent:

(i) *There exists a constant $C > 0$ with*

$$e(\lambda) \le C\varphi^{2\nu}(\lambda) \quad \text{for all } \lambda > 0. \tag{33}$$

(ii) *There exists a constant $\tilde{C} > 0$ such that*

$$\langle x^\dagger, x\rangle \le \tilde{C}\|\varphi(L^*L)x\|^\nu \|x\|^{1-\nu} \quad \text{for all } x \in X. \tag{34}$$

Proof. Assume first that (34) holds. Then, we have for all $\lambda > 0$

$$\begin{aligned}\|E_{[0,\lambda]}x^\dagger\|^2 &= \langle x^\dagger, E_{[0,\lambda]}x^\dagger\rangle \\ &\le \tilde{C}\|\varphi(L^*L)E_{[0,\lambda]}x^\dagger\|^\nu \|E_{[0,\lambda]}x^\dagger\|^{1-\nu} \\ &\le \tilde{C}\varphi^\nu(\lambda)\|E_{[0,\lambda]}x^\dagger\|,\end{aligned}$$

which implies (33) with $C = \tilde{C}^2$.

On the other hand, if (33) is fulfilled then we can estimate for arbitrary $\Lambda > 0$ and every $x \in X$

$$\left|\langle E_{[0,\Lambda]}x^\dagger, x\rangle\right| \le \|E_{[0,\Lambda]}x^\dagger\|\|x\| \le \sqrt{C}\,\varphi^\nu(\Lambda)\|x\|. \tag{35}$$

Furthermore, we get with the bounded, invertible operator $T = \varphi(L^*L)|_{\mathcal{R}(E_{[\Lambda,\infty)})}$ that

$$\begin{aligned}\left|\langle E_{[\Lambda,\infty)}x^\dagger, x\rangle\right| &= \left|\langle T^{-1}E_{[\Lambda,\infty)}x^\dagger, TE_{[\Lambda,\infty)}x\rangle\right| \\ &\le \|TE_{[\Lambda,\infty)}x\|\sqrt{\lim_{\varepsilon\downarrow 0}\int_{\Lambda-\varepsilon}^{\|L\|^2}\frac{1}{\varphi^2(\lambda)}\,\mathrm{d}e(\lambda)}.\end{aligned} \tag{36}$$

Integrating by parts, we can rewrite the integral in the form

$$\int_{\Lambda-\varepsilon}^{\|L\|^2}\frac{1}{\varphi^2(\lambda)}\,\mathrm{d}e(\lambda) = \frac{e(\|L\|^2)}{\varphi^2(\|L\|^2)} - \frac{e(\Lambda-\varepsilon)}{\varphi^2(\Lambda-\varepsilon)} + 2\int_{\Lambda-\varepsilon}^{\|L\|^2}\frac{e(\lambda)}{\varphi^3(\lambda)}\,\mathrm{d}\varphi(\lambda).$$

Using now (33) and dropping all negative terms, we arrive at

$$\lim_{\varepsilon\downarrow 0}\int_{\Lambda-\varepsilon}^{\|L\|^2}\frac{1}{\varphi^2(\lambda)}\,\mathrm{d}e(\lambda)\leq\frac{C}{\varphi^{2-2\nu}(\|L\|^2)}+\frac{C}{1-\nu}\frac{1}{\varphi^{2-2\nu}(\Lambda)}\leq\frac{c^2}{\varphi^{2-2\nu}(\Lambda)}$$

with the constant $c>0$ given by $c^2=C(1+\frac{1}{1-\nu})$. Plugging this into (36), we find that

$$\left|\left\langle E_{[\Lambda,\infty)}x^\dagger,x\right\rangle\right|\leq\frac{c}{\varphi^{1-\nu}(\Lambda)}\|\varphi(L^*L)x\|. \tag{37}$$

We now pick

$$\Lambda=\inf\{\lambda>0\mid\left|\left\langle E_{[0,\lambda]}x^\dagger,x\right\rangle\right|\geq\tfrac{1}{2}\left|\left\langle x^\dagger,x\right\rangle\right|\}$$

and assume that $\Lambda>0$; otherwise $\left\langle x^\dagger,x\right\rangle=0$ and (34) is trivially fulfilled. Then, the right continuity of $\lambda\mapsto\left\langle E_{[0,\lambda]}x^\dagger,x\right\rangle$ implies that

$$\left|\left\langle E_{[0,\Lambda]}x^\dagger,x\right\rangle\right|\geq\frac{1}{2}\left|\left\langle x^\dagger,x\right\rangle\right|.$$

Moreover, we have that

$$\left|\left\langle E_{[\lambda,\infty)}x^\dagger,x\right\rangle\right|\geq\left|\left\langle x^\dagger,x\right\rangle\right|-\left|\left\langle E_{[0,\lambda]}x^\dagger,x\right\rangle\right|>\frac{1}{2}\left|\left\langle x^\dagger,x\right\rangle\right|$$

for every $\lambda\in(0,\Lambda)$. Therefore, the left continuity of $\lambda\mapsto\left\langle E_{[\lambda,\infty)}x^\dagger,x\right\rangle$ implies that

$$\left|\left\langle E_{[\Lambda,\infty)}x^\dagger,x\right\rangle\right|\geq\frac{1}{2}\left|\left\langle x^\dagger,x\right\rangle\right|.$$

Thus, we get with the estimates (35) and (37) that

$$\begin{aligned}\left\langle x^\dagger,x\right\rangle&\leq 2\left|\left\langle E_{[0,\Lambda]}x^\dagger,x\right\rangle\right|^{1-\nu}\left|\left\langle E_{[\Lambda,\infty)}x^\dagger,x\right\rangle\right|^{\nu}\\&\leq 2C^{\frac{1-\nu}{2}}c^\nu\|\varphi(L^*L)x\|^\nu\|x\|^{1-\nu}.\end{aligned}$$

□

We remark that the first part of this proof also works in the limit case $\nu=1$, which shows that (34) implies (33) for $\nu=1$ as well.

Corollary 4.2. *We use again Notation 2.2. Let further $\varphi:[0,\infty)\to[0,\infty)$ be an increasing, continuous function and $\nu\in(0,1]$.*

Then, the standard source condition

$$x^\dagger\in\mathcal{R}(\varphi^\nu(L^*L)) \tag{38}$$

implies the variational inequality

$$\left\langle x^\dagger,x\right\rangle\leq C\|\varphi(L^*L)x\|^\nu\|x\|^{1-\nu}\quad\textit{for all } x\in X \tag{39}$$

for some constant $C>0$.

Conversely, the variational inequality (39) *implies that*

$$x^\dagger \in \mathcal{R}(\psi(L^*L))$$

for every continuous function $\psi : [0,\infty) \to [0,\infty)$ *with* $\psi \geq c\varphi^\mu$ *for some constant* $c > 0$ *and some* $\mu \in (0,\nu)$.

Proof. If $x^\dagger$ fulfils (38), then there exists an element $\omega \in X$ with

$$\langle x^\dagger, x\rangle = \langle \omega, \varphi^\nu(L^*L)x\rangle \leq \|\omega\| \|\varphi^\nu(L^*L)x\|. \tag{40}$$

Using the interpolation inequality, see for example [3, Chapter 2.3], we find

$$\langle x^\dagger, x\rangle \leq \|\omega\| \|\varphi(L^*L)x\|^\nu \|x\|^{1-\nu},$$

which is (39) with $C = \|\omega\|$.

If, on the other hand, (39) holds, then, according to Proposition 4.1 there exists a constant $\tilde{C} > 0$ such that $e(\lambda) \leq \tilde{C}\varphi^{2\nu}(\lambda)$. Now, similarly to the proof of Proposition 4.1 we get with $T = \psi(L^*L)|_{\mathcal{R}(E_{(\Lambda,\infty)})}$ that

$$\langle E_{(\Lambda,\infty)}x^\dagger, x\rangle \leq \langle T^{-1}E_{(\Lambda,\infty)}x^\dagger, TE_{(\Lambda,\infty)}x\rangle \leq \|TE_{(\Lambda,\infty)}x\| \sqrt{\int_\Lambda^{\|L\|^2} \frac{1}{\psi^2(\lambda)}\,\mathrm{d}e(\lambda)},$$

and, using the lower bound on ψ, that

$$\int_\Lambda^{\|L\|^2} \frac{1}{\psi^2(\lambda)}\,\mathrm{d}e(\lambda) \leq \frac{1}{c^2}\int_\Lambda^{\|L\|^2} \frac{1}{\varphi^{2\mu}(\lambda)}\,\mathrm{d}e(\lambda) \leq \tilde{c}^2\varphi^{2(\nu-\mu)}(\|L\|^2),$$

for some constant $\tilde{c} > 0$. So,

$$\langle x^\dagger, x\rangle = \lim_{\Lambda\to 0}\langle E_{(\Lambda,\infty)}x^\dagger, x\rangle \leq \tilde{c}\varphi^{\nu-\mu}(\|L\|^2)\|\psi(L^*L)x\|,$$

which implies that $x^\dagger \in \mathcal{R}(\psi(L^*L))$, see for example [11, Lemma 8.21]. □

Remark. In general, the inequality (39) does not imply the standard source condition (38). Let us for example consider the case where we have an increasing, continuous function $\varphi : [0,\infty) \to [0,\infty)$ with $\varphi(0) = 0$, $\varphi(\lambda) > 0$ for all $\lambda > 0$, and

$$c\varphi^{2\nu}(\lambda) \leq e(\lambda) \leq C\varphi^{2\nu}(\lambda) \quad \text{for all } \lambda > 0$$

for some constants $0 < c \leq C$.

Now, the standard source condition (38) would imply that we can find a $\xi \in \mathcal{N}(L)^\perp$ with $x^\dagger = \varphi^\nu(L^*L)\xi$. Thus, we would get with $T = \varphi^\nu(L^*L)|_{\mathcal{R}(E_{(\Lambda,\infty)})}$ that

$$\|\xi\|^2 = \lim_{\Lambda\to 0}\|E_{(\Lambda,\infty)}\xi\|^2 = \lim_{\Lambda\to 0}\|T^{-1}E_{(\Lambda,\infty)}x^\dagger\|^2 = \lim_{\Lambda\to 0}\int_\Lambda^{\|L\|^2}\frac{1}{\varphi^{2\nu}(\lambda)}\,\mathrm{d}e(\lambda).$$

However, in the limit $\Lambda \to 0$, we have that

$$\int_{\Lambda}^{\|L\|^2} \frac{1}{\varphi^{2\nu}(\lambda)}\,\mathrm{d}e(\lambda) = \frac{e(\|L\|^2)}{\varphi^{2\nu}(\|L\|^2)} - \frac{e(\Lambda)}{\varphi^{2\nu}(\Lambda)} + 2\nu \int_{\Lambda}^{\|L\|^2} \frac{e(\lambda)}{\varphi^{2\nu+1}(\lambda)}\,\mathrm{d}\varphi(\lambda)$$
$$\geq c - C + 2\nu c \log\left(\frac{\varphi(\|L\|^2)}{\varphi(\Lambda)}\right) \to \infty,$$

which is a contradiction to the existence of such a point ξ.

5. Connection to approximate source conditions

Another approach to weakening the standard source condition (38) in order to obtain a condition which is equivalent to the convergence rate was introduced in [9], see also [4]. The idea was that for the argument (40), which shows that the standard source condition (38) implies the variational inequality (39), it would have been enough to be able to approximate the minimum-norm solution $x^\dagger$ by a bounded sequence in $\mathcal{R}(\varphi^\nu(L^*L))$. And, the smaller the bound on the sequence, the smaller the constant C in the variational inequality (39) will be. Therefore, the distance between $x^\dagger$ and $\mathcal{R}(\varphi^\nu(L^*L)) \cap \bar{B}_R(0)$ as a function of the radius R of the closed ball $\bar{B}_R(0) = \{x \in X \mid \|x\| \leq R\}$ should be directly related to the convergence rate.

Definition 5.1. In the setting of Notation 2.2, we define the distance function d_φ of a continuous function $\varphi : [0,\infty) \to [0,\infty)$ by

$$d_\varphi(R) = \inf_{\xi \in \bar{B}_R(0)} \|x^\dagger - \varphi(L^*L)\xi\|. \tag{41}$$

Indeed, this distance function gives us directly an upper bound on the error between the regularized solution $x_\alpha(y)$ and the minimum-norm solution $x^\dagger$, see [9, Theorem 5.5] or [4, Proposition 2]. For convenience, we repeat the argument here.

Lemma 5.2. *We use Notation 2.2 and assume that $\varphi : [0,\infty) \to [0,\infty)$ is an increasing, continuous function with $\varphi(0) = 0$ so that there exists a constant $A > 0$ such that the inequality*

$$\sqrt{\tilde{r}_\alpha(\lambda)}\varphi(\lambda) \leq A\varphi(\alpha) \quad \text{for all } \lambda > 0 \tag{42}$$

holds for every $\alpha > 0$.

Then, we have for every $\xi \in X$ that

$$\|x_\alpha(y) - x^\dagger\| \leq \|x^\dagger - \varphi(L^*L)\xi\| + A\varphi(\alpha)\|\xi\| \quad \text{for all } \alpha > 0. \tag{43}$$

Proof. For every vector $\xi \in X$, we find from (5) with the definition (3) of the error function $\tilde{r}_\alpha$ that

$$\|x_\alpha(y) - x^\dagger\| = \|\tilde{r}_\alpha^{\frac{1}{2}}(L^*L)x^\dagger\| \leq \|\tilde{r}_\alpha^{\frac{1}{2}}(L^*L)(x^\dagger - \varphi(L^*L)\xi)\| + \|\tilde{r}_\alpha^{\frac{1}{2}}(L^*L)\varphi(L^*L)\xi\|.$$

Now, since $\tilde{r}_\alpha(\lambda) \le 1$, we have that $\|\tilde{r}_\alpha(L^*L)\| \le 1$. Moreover, with $e_\xi(\lambda) = \|E_{(0,\lambda]}\xi\|^2$, we get from the inequality (42) that

$$\|\tilde{r}_\alpha^{\frac{1}{2}}(L^*L)\varphi(L^*L)\xi\|^2 = \int_0^{\|L\|^2} \tilde{r}_\alpha(\lambda)\varphi^2(\lambda)\,\mathrm{d}e_\xi(\lambda) \le A^2\varphi^2(\alpha)\|\xi\|^2.$$

So, putting the two inequalities together, we obtain (43). □

Thus, taking the infimum over all $\xi \in \bar{B}_R(0)$ in (43), the error $\|x_\alpha(y)-x^\dagger\|$ can be bound by a combination of $d_\varphi(R)$ and $\varphi(\alpha)R$. By balancing these terms, we obtain from a given distance function d_φ the corresponding convergence rate.

Conversely, we can also show that an upper bound on the spectral projections of the minimum-norm solution gives us an upper bound on the distance function, which then yields another equivalent characterisation for the convergence rate of the regularization method.

Proposition 5.3. *We use Notation 2.2 and assume that $\varphi : [0,\infty) \to [0,\infty)$ is an increasing, continuous function with $\varphi(0) = 0$ so that there exists a constant $A > 0$ with*

$$\sqrt{\tilde{r}_\alpha(\lambda)}\varphi(\lambda) \le A\varphi(\alpha) \quad \text{for all } \lambda > 0,\ \alpha > 0. \tag{44}$$

Moreover, let d_φ be the distance function of φ, and let $\nu \in (0,1)$ be arbitrary.

Then, the following statements are equivalent:

(i) *There exists a constant $C > 0$ so that*

$$e(\lambda) \le C\varphi^{2\nu}(\lambda) \quad \text{for all } \lambda > 0. \tag{45}$$

(ii) *There exists a constant $\tilde{C} > 0$ so that*

$$d_\varphi(R) \le \tilde{C}R^{-\frac{\nu}{1-\nu}} \quad \text{for all } R > 0. \tag{46}$$

Proof. Assume first that (46) holds. Then, from Lemma 5.2, we get by taking the infimum of (43) over all $\xi \in \bar{B}_R(0)$ for an arbitrary $R > 0$ that

$$\|x_\alpha(y) - x^\dagger\| \le d_\varphi(R) + A\varphi(\alpha)R \le \tilde{C}R^{-\frac{\nu}{1-\nu}} + A\varphi(\alpha)R.$$

Since the first term is decreasing and the second term is increasing in R, we pick for R the value $R(\alpha)$ given by

$$R^{-\frac{\nu}{1-\nu}}(\alpha) = \varphi(\alpha)R(\alpha), \quad \text{that is} \quad R(\alpha) = \varphi^{-(1-\nu)}(\alpha).$$

Thus, we end up with

$$\|x_\alpha(y) - x^\dagger\| \le (\tilde{C} + A)\varphi^\nu(\alpha).$$

Applying Proposition 2.3 with the function φ, therein replaced by $\varphi^{2\nu}$ (we remark that the condition (7) is then fulfilled with $\mu = \nu$, since (44) implies $\varphi^{2\nu}(\lambda)\tilde{r}_\alpha^\nu(\lambda) \le A^{2\nu}\varphi^{2\nu}(\alpha)$) we find that there exists a constant $C > 0$ so that (45) holds.

Conversely, if we have the relation (45), then we define for arbitrary $\alpha > 0$ with the operator $T = \varphi(L^*L)|_{\mathcal{R}(E_{(\alpha,\infty)})}$ the element

$$\xi_\alpha = T^{-1}E_{(\alpha,\infty)}x^\dagger.$$

Now, the distance of $\varphi(L^*L)\xi_\alpha$ to the minimum-norm solution $x^\dagger$ can be estimated according to (45) by

$$\|x^\dagger - \varphi(L^*L)\xi_\alpha\|^2 = \|E_{[0,\alpha]}x^\dagger\|^2 \leq C\varphi^{2\nu}(\alpha). \tag{47}$$

Moreover, we can get an upper bound on the norm of ξ_α by

$$\|\xi_\alpha\|^2 = \int_\alpha^{\|L\|^2} \frac{1}{\varphi^2(\lambda)}\,\mathrm{d}e(\lambda) = \frac{e(\|L\|^2)}{\varphi^2(\|L\|^2)} - \frac{e(\alpha)}{\varphi^2(\alpha)} + 2\int_\alpha^{\|L\|^2} \frac{e(\lambda)}{\varphi^3(\lambda)}\,\mathrm{d}\varphi(\lambda).$$

Using assumption (45), evaluating the integral, and dropping the resulting two negative terms, we find that

$$\|\xi_\alpha\|^2 \leq \frac{C}{\varphi^{2-2\nu}(\|L\|^2)} + \frac{C}{1-\nu}\frac{1}{\varphi^{2-2\nu}(\alpha)} \leq \frac{c^2}{\varphi^{2-2\nu}(\alpha)} \tag{48}$$

with $c^2 = C(1 + \frac{1}{1-\nu})$.

So, combining (47) and (48), we have by definition (41) of the distance function d_φ with $R = c\varphi^{-(1-\nu)}(\alpha)$ that

$$d_\varphi(c\varphi^{-(1-\nu)}(\alpha)) \leq \sqrt{C}\varphi^\nu(\alpha),$$

and thus it follows by switching to the variable R that

$$d_\varphi(R) \leq \tilde{C}R^{-\frac{\nu}{1-\nu}},$$

where $\tilde{C} = \sqrt{C}\,c^{\frac{\nu}{1-\nu}}$. □

6. Conclusion

In this article, we have proven optimal convergence rates results for regularization methods for solving linear ill-posed operator equations in Hilbert spaces. The result generalizes existing convergence rates results on optimality of [10] to general source conditions, such as logarithmic source conditions. The results state that convergence rates results of regularised solution require a certain decay of the solution in terms of the spectral decomposition. Moreover, we also provide optimality results under variational source conditions, extending the results of [1]. It is interesting to note that variational source conditions are equivalent to convergence rates of the regularized solutions, while the classical results are not. Moreover, we also show that rates of the distance function developed in [4, 9] are equivalent to convergence rates of the regularized solutions.

Funding

Author Albani is supported by CNPq grant 201644/2014-2. The work of O. Scherzer has been supported by the Austrian Science Fund (FWF) within the national research network Geometry and Simulation, project S11704 (Variational Methods for Imaging on Manifolds) and Project P26687-N25 (Interdisciplinary Coupled Physics Imaging).

References

1. R. Andreev, P. Elbau, M. V. de Hoop, L. Qiu, and O. Scherzer (2015). Generalized convergence rates results for linear inverse problems in Hilbert spaces. *Numer. Funct. Anal. Optim.* 36(5):549–566.
2. M. Carter and B. van Brunt (2000). The Lebesgue–Stieltjes integral. *A Practical Introduction. Undergraduate Mathematics Series*. Springer, New York, NY.
3. H. W. Engl, M. Hanke, and A. Neubauer (1996). Regularization of inverse problems. *Mathematics and its Applications*, Vol. 375. Kluwer Academic Publishers Group, Dordrecht, Netherlands.
4. J. Flemming, B. Hofmann, and P. Mathé (2011). Sharp converse results for the regularization error using distance functions. *Inverse Probl.* 27, page 025006. doi: 10.1088/0266-5611/27/2/025006
5. C. W. Groetsch (1984). *The Theory of Tikhonov Regularization for Fredholm Equations of the First Kind*. Pitman Advanced Publishing Program, Boston, MA.
6. P. R. Halmos (1974). *A Hilbert Space Problem Book*. Graduate Texts in Mathematics, Vol. 19. Springer, New York, NY.
7. T. Hein and B. Hofmann (2009). Approximate source conditions for nonlinear ill-posed problems: Chances and limitations. *Inverse Probl.* 25:035003, 16 pages.
8. B. Hofmann, B. Kaltenbacher, C. Pöschl, and O. Scherzer (2007). A convergence rates result for Tikhonov regularization in Banach spaces with non-smooth operators. *Inverse Probl.* 23(3):987–1010.
9. B. Hofmann and P. Mathé (2007). Analysis of profile functions for general linear regularization methods. *SIAM J. Numer. Anal.* 45(3):1122–1141.
10. A. Neubauer (1997). On converse and saturation results for Tikhonov regularization of linear ill-posed problems. *SIAM J. Numer. Anal.* 34:517–527.
11. O. Scherzer, M. Grasmair, H. Grossauer, M. Haltmeier, and F. Lenzen (2009). Variational methods in imaging. *Applied Mathematical Sciences* vol. 167. Springer, New York, NY.
12. T. Schuster, B. Kaltenbacher, B. Hofmann, and K. S. Kazimierski (2012). Regularization methods in Banach spaces. *Radon Series on Computational and Applied Mathematics*, vol. 10. Walter de Gruyter, Berlin, Germany.