Capítulo 4

Reglas de elección del parámetro

En este capítulo estudiaremos varias reglas de elección del parámetro más utilizadas. Algunas de ellas, como veremos, forman parte de métodos de regularización de orden óptimo en conjuntos fuente del tipo $\mathcal{X}_\mu$ o $\mathcal{X}_{\mu,\rho}$, μ, $\rho > 0$, mientras que otras son reglas heurísticas. En la Sección 4.1 presentaremos reglas de elección del parámetro *a-priori*, las que dependen sólo del nivel de ruido δ y por ello, pueden definirse "antes" de la obtención de los datos y^δ. En la Sección 4.2 estudiaremos una de las reglas de elección del parámetro *a-posteriori* más importantes, y más frecuentemente utilizada, conocida como el "Principio de Discrepancia de Morozov". Por último, en la Sección 4.3 presentaremos algunas reglas heurísticas también llamadas reglas "libres de error", puesto que dependen sólo del dato y^δ y no del nivel de ruido δ. Estas reglas son muy útiles en problemas del mundo real en los cuales la información acerca del nivel de ruido en los datos no está disponible o bien no es confiable.

Es oportuno señalar aquí que el tema "elección del parámetro de regularización" constituye toda un área dentro de la especialidad y en muchos casos los resultados son escasos o inexistentes. Por ejemplo, no se conoce hasta el momento ningún método o algoritmo que permita escoger adecuadamente el parámetro de reconstrucción en el método de la inversa aproximada.

4.1. Reglas de elección del parámetro *a-priori*

En esta sección estudiaremos reglas de elección del parámetro *a-priori* que dependen sólo del nivel de ruido δ y no del dato perturbado y^δ. Una forma sencilla en que se pueden definir este tipo de reglas es como potencias de δ, es decir,

$$\alpha(\delta) = p\,\delta^s, \tag{4.1}$$

donde p y s son constantes positivas. Sean $\alpha(\delta)$ una regla de elección del parámetro *a-priori* y $h : \mathbb{R}^+ \to \mathbb{R}^+$. Si existen constantes positivas c_1 y c_2 tales que

$$c_1 h(\delta) \leq \alpha(\delta) \leq c_2 h(\delta),$$

entonces diremos que "α es del orden de $h(\delta)$" y lo denotaremos con "$\alpha \sim h(\delta)$".

Recordemos que una regla de elección del parámetro, a diferencia de una familia de operadores de regularización, se define para un dato y fijo. Veremos a continuación que, en algunos casos, cuando se tiene algún tipo de información *a-priori* acerca de la solución exacta $x^\dagger \doteq T^\dagger y$, por ejemplo cuando se conocen los valores de μ y ρ en la condición fuente $x^\dagger \in \mathcal{X}_{\mu,\rho}$, es posible definir reglas de elección del parámetro como en (4.1), donde las costantes p y s se determinan a partir de μ y ρ, de tal manera que si el MRE satisface determinadas condiciones, dichas reglas forman parte de un método de regularización de orden óptimo en $\mathcal{X}_{\mu,\rho}$.

Proposición 4.1. *Sean $\mathcal{X}$, $\mathcal{Y}$ espacios de Hilbert, $T \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$, $\{g_\alpha\}$ un MRE tal que G_α definida en (3.4) satisface*

$$G_\alpha = O\left(\frac{1}{\alpha}\right) \quad \textit{cuando } \alpha \to 0^+. \tag{4.2}$$

Sean μ, $\rho > 0$. Supongamos que existe una constante $c > 0$ independiente de α tal que $\omega_\mu : (0, \alpha_0) \to \mathbb{R}^+$ definida como en (3.14) satisface (3.9). Si $\alpha^*(\delta)$ es una regla de elección del parámetro a-priori que satisface

$$\alpha^* \sim \left(\frac{\delta}{\rho}\right)^{\frac{2}{2\mu+1}}, \tag{4.3}$$

entonces el método de regularización $(\{R_\alpha\}, \alpha^*(\delta))$ es de orden óptimo en $\mathcal{X}_{\mu,\rho}$, donde $\{R_\alpha\}$ está definido por (3.3).

Demostración. Sean $x^\dagger \doteq T^\dagger y \in \mathcal{X}_{\mu,\rho}$, $y^\delta \in \mathcal{Y}$ tal que $\left\|y^\delta - y\right\| \leq \delta$, $x_\alpha \doteq R_\alpha y$ y $x_\alpha^\delta \doteq R_\alpha y^\delta$. De los Teoremas 3.4 y 3.5 (recordar que por hipótesis $\omega_\mu(\alpha) = c\alpha^\mu$ satisface (3.9)) y de la condición (4.2) sobre G_α se sigue que

$$\left\|x_{\alpha^*}^\delta - x^\dagger\right\| \leq \left\|x_{\alpha^*} - x^\dagger\right\| + \left\|x_{\alpha^*} - x_{\alpha^*}^\delta\right\| \leq c\alpha^{*\mu}\rho + k\delta\sqrt{\frac{C}{\alpha^*}}, \tag{4.4}$$

donde k es una constante positiva independiente de α y C es la constante de la hipótesis (*H2*). Utilizando (4.3) es fácil probar que el lado derecho de la última desigualdad en (4.4) se puede acotar por $\hat{c}\,\delta^{\frac{2\mu}{2\mu+1}}\rho^{\frac{1}{2\mu+1}}$, donde $\hat{c}$ es una constante positiva independiente de δ. Luego (ver Definición 2.16) el método de regularización $(\{R_\alpha\}, \alpha^*(\delta))$ es de orden óptimo en $\mathcal{X}_{\mu,\rho}$, como queríamos probar. ∎

Observación 4.2. A menudo sucede que no se conoce el valor de ρ. En tal caso, si consideramos una regla de elección del parámetro $\alpha \sim \delta^{\frac{2}{2\mu+1}}$, se obtiene una cota para el error total que es al menos óptima con respecto a la potencia de δ, es decir, de orden óptimo en $\mathcal{X}_\mu$: $\left\|x_\alpha^\delta - x^\dagger\right\| = O\left(\delta^{\frac{2\mu}{2\mu+1}}\right)$. Si también μ es desconocido, entonces en tal caso no es apropiado usar una regla de elección del parámetro *a-priori* y se deberá considerar alguna regla *a-posteriori* como las que veremos en la siguiente sección.

Es importante observar que si $\{g_\alpha\}$ es un MRE que posee calificación clásica de orden μ_0, entonces la mejor tasa de convergencia que se puede obtener para el error total con una regla de elección del parámetro *a-priori* dada por (4.3) es $O\left(\delta^{\frac{2\mu_0}{2\mu_0+1}}\right)$, la cual se obtiene bajo el supuesto *a-priori* de que $x^\dagger \in \mathcal{X}_{\mu_0}$. Esto está fuertemente relacionado con el fenómeno de "saturación" que mencionamos en la Sección 3.1.3. Sin embargo, si $\{g_\alpha\}$ es un MRE para el cual $\mu_0 = +\infty$, entonces la condición (3.9) con ω_μ dado por (3.14) se satisface para todo $\mu > 0$. Luego, se deduce de la Proposición 4.1 que si $x^\dagger \in \mathcal{X}_\mu$ para μ arbitrariamente grande, entonces la tasa de convergencia a cero del error total puede ser arbitrariamente cercana a $O(\delta)$, aunque nunca se alcance realmente esta tasa. Por ejemplo, en la Sección 3.2 vimos que para el método TSVD y el método de Showalter se tiene que $\mu_0 = +\infty$. Luego, con una regla de elección del parámetro que satisfaga (4.3), estos métodos resultan de orden óptimo en $\mathcal{X}_{\mu,\rho}$ para todos μ, $\rho > 0$. Por otra parte, como la regularización de Tikhonov-Phillips posee calificación clásica de orden $\mu_0 = 1$ (ver Sección 3.2.2), con una regla de elección del parámetro que satisfaga (4.3), este método es de orden óptimo en $\mathcal{X}_{\mu,\rho}$ para todos $\mu \in (0, 1]$ y $\rho > 0$. La mejor tasa de convergencia a cero para el error total es entonces $O\left(\delta^{\frac{2}{3}}\right)$ y se obtiene con una regla de elección del parámetro $\alpha \sim \left(\frac{\delta}{\rho}\right)^{\frac{2}{3}}$ bajo el supuesto $x^\dagger \in \mathcal{X}_{1,\rho}$ para cualquier $\rho > 0$. Esto coincide con lo obtenido previamente en el Capítulo 3.

4.2. Reglas de elección del parámetro *a-posteriori*: el Principio de Discrepancia de Morozov

En esta sección presentaremos una regla de elección del parámetro *a-posteriori* muy utilizada, conocida como el *"Principio de Discrepancia"* de V. A. Morozov ([46]). Como señalamos en la Observación 4.2, este tipo de reglas de elección del parámetro es muy apropiado en los casos en que no se dispone de ninguna información *a-priori* acerca de la regularidad de la solución exacta, en particular, no se conoce ningún valor de $\mu > 0$ tal que $x^\dagger \in \mathcal{X}_\mu$.

Dado un MRE $\{g_\alpha\}$, el Principìo de Discrepancia consiste esencialmente en elegir el parámetro de regularización mediante una comparación entre la *"discrepancia"* $\left\|Tx_\alpha^\delta - y^\delta\right\| = \left\|(TR_\alpha - I)y^\delta\right\|$ (de allí el nombre) y una cota asociada con el nivel de ruido δ, donde R_α está dado por (3.3). Más precisamente, supongamos que queremos resolver el problema $Tx = y$, pero en lugar del dato exacto y sólo se conoce un dato con ruido y^δ tal que $\left\|y - y^\delta\right\| \leq \delta$. En este caso, no tiene sentido tratar de encontrar una solución aproximada $\tilde{x}$ que produzca una discrepancia $\left\|T\tilde{x} - y^\delta\right\|$ que sea *estrictamente menor* que δ; lo mejor que podríamos obtener es una discrepancia de orden δ. Por otra parte, hemos visto en la Sección 3.1 que para un nivel de ruido fijo $\delta > 0$, el error total aumenta a medida que el parámetro de regularización α tiende a cero, con lo cual la estabilidad disminuye. Este simple razonamiento nos permite concluir que lo mejor que podemos hacer es elegir el mayor parámetro de regularización posible que resulte en una discrepancia de orden δ. Es precisamente la formalización de este razonamiento lo que constituye el Principio de Discrepancia de Morozov. Sea τ una constante positiva tal que

$$\tau > \sup_{\alpha>0,\, \lambda\in[0,\|T\|^2]} |r_\alpha(\lambda)|\,, \tag{4.5}$$

donde $r_\alpha(\lambda)$ es como en (3.7). Notar que $\tau > 1$ pues $r_\alpha(0) = 1$. El parámetro de regularización definido mediante el Principio de Discrepancia está dado por

$$\alpha_d(\delta, y^\delta) \doteq \sup\{\alpha > 0 : \left\|Tx_\alpha^\delta - y^\delta\right\| \leq \tau\delta\}. \tag{4.6}$$

Es fácil probar que si para cada $\lambda > 0$, el mapeo $\alpha \to g_\alpha(\lambda)$ es continuo por izquierda entonces el funcional $\alpha \to \left\|Tx_\alpha^\delta - y^\delta\right\|$ también es continuo por izquierda. Luego, se alcanza el supremo en (4.6) y así,

$$\left\|Tx^\delta_{\alpha_d(\delta,y^\delta)} - y^\delta\right\| \leq \tau\delta. \tag{4.7}$$

Si $\left\|Tx_\alpha^\delta - y^\delta\right\| \leq \tau\delta$ para todo $\alpha > 0$ entonces $\alpha_d(\delta, y^\delta) = +\infty$ y en tal caso $x^\delta_{\alpha_d(\delta,y^\delta)}$ tiene que entenderse en el sentido del límite de x_α^δ cuando $\alpha \to +\infty$. Notar que si $\{g_\alpha\}$

es tal que se satisface (3.23) con G_α dado por (3.4), entonces

$$x_\infty^\delta \doteq \lim_{\alpha\to+\infty} x_\alpha^\delta = 0. \tag{4.8}$$

En efecto, de la definición de R_α (ver (3.3)) se sigue que

$$\left\|x_\alpha^\delta\right\| = \left\|R_\alpha y^\delta\right\| = \left\|g_\alpha(T^*T)T^*y^\delta\right\| \leq \|g_\alpha(T^*T)\| \left\|T^*y^\delta\right\|. \tag{4.9}$$

Para $x \in \mathcal{X}$, resulta de (3.4) que

$$\|g_\alpha(T^*T)x\|^2 = \int_0^{\|T\|^2+} g_\alpha^2(\lambda)\, d\, \|E_\lambda x\|^2 \leq G_\alpha^2 \|x\|^2,$$

y entonces, $\|g_\alpha(T^*T)\| \leq G_\alpha$. Puesto que G_α satisface (3.23), existe una constante positiva c tal que $G_\alpha \leq c/\alpha$ para todo $\alpha > 0$. Se sigue entonces que $\|g_\alpha(T^*T)\| \leq c/\alpha$ para todo $\alpha > 0$, lo cual junto a (4.9) implica que $\left\|x_\alpha^\delta\right\| \leq (c/\alpha)\left\|T^*y^\delta\right\|$. Haciendo tender α a infinito, se obtiene (4.8).

Es oportuno observar aquí que la discrepancia $\left\|Tx_\alpha^\delta - y^\delta\right\|$ nunca es menor que $\|y - Qy\| - \delta$, donde Q es la proyección ortogonal de $\mathcal{Y}$ sobre $\overline{Ran(T)}$. En efecto, para $y \in \mathcal{Y}$ se tiene que $\|Qy - y\| \leq \|Tx - y\|$ para todo $x \in \mathcal{X}$ puesto que $Tx \in Ran(T)$ y Q es la proyección ortogonal de $\mathcal{Y}$ sobre $\overline{Ran(T)}$. Como $x_\alpha^\delta \in \mathcal{X}$ se tiene entonces que

$$\|Qy - y\| \leq \left\|Tx_\alpha^\delta - y\right\| \leq \left\|Tx_\alpha^\delta - y^\delta\right\| + \left\|y - y^\delta\right\| \leq \left\|Tx_\alpha^\delta - y^\delta\right\| + \delta,$$

es decir, $\left\|Tx_\alpha^\delta - y^\delta\right\| \geq \|Qy - y\| - \delta$. Luego, el conjunto dado en (4.6) podría ser vacío (por ejemplo, si $\|Qy - y\| = 3\delta$ y $\tau = 3/2$). Para evitar esto, aquí supondremos siempre que y es alcanzable, es decir, que $y \in Ran(T)$, con lo cual $Qy = y$ y el conjunto en (4.6) no es vacío. Esto se deduce de observar lo siguiente:

$$\begin{aligned}
\left\|Tx_\alpha^\delta - y^\delta\right\| &= \left\|r_\alpha(TT^*)y^\delta\right\| \\
&\leq \|r_\alpha(TT^*)y\| + \left\|r_\alpha(TT^*)(y - y^\delta)\right\| \\
&\leq \|r_\alpha(TT^*)y\| + \delta \sup_{\alpha>0,\, \lambda\in[0,\|T\|^2]} |r_\alpha(\lambda)|
\end{aligned}$$

$$\doteq \; c_\alpha(y) + \delta \sup_{\alpha>0,\, \lambda\in[0,\|T\|^2]} |r_\alpha(\lambda)| \,. \tag{4.10}$$

Ahora, si $y \in Ran(T)$, en virtud de (1.7) y de que $R_\alpha \to T^\dagger$ puntualmente sobre $Dom(T^\dagger)$, se sigue que $c_\alpha(y)$ tiende a cero cuando $\alpha \to 0^+$. Luego, el lado derecho en (4.10), cuando $\alpha \to 0^+$ converge a $\delta \sup_{\alpha>0,\, \lambda\in[0,\|T\|^2]} |r_\alpha(\lambda)|$ que es estrictamente menor que $\tau\delta$ en virtud de (4.5). Así, existe $\alpha_0 > 0$ tal que $\left\|Tx^\delta_{\alpha_0} - y^\delta\right\| < \tau\delta$ y el conjunto en (4.6) es efectivamente no vacío si $y \in Ran(T)$.

A continuación probaremos algunas propiedades de convergencia y optimalidad de métodos de regularización $(\{R_\alpha\}, \alpha_d)$, donde R_α está definido como en (3.3) y α_d es la regla de elección del parámetro dada por el Principio de Discrepancia (4.6). Debido a la importancia y trascendencia de estos resultados decidimos incluir aquí las demostraciones de los mismos, las cuales contienen ideas y desarrollos propios. En el enfoque utilizado para estos desarrollos se ha puesto énfasis en la utilización de las poderosas herramientas que proporciona la teoría espectral introducida en la Sección 1.3.

Teorema 4.3. *Sean $\mathcal{X}$, $\mathcal{Y}$ espacios de Hilbert, $T \in \mathcal{L}(\mathcal{X},\mathcal{Y})$, $\{g_\alpha\}$ un MRE con calificación clásica de orden $\mu_0 > \frac{1}{2}$ y tal que G_α definida por (3.4) satisface la condición (3.23) para alguna constante positiva $\hat{c}$. Entonces el método de regularización $(\{R_\alpha\}, \alpha_d)$, donde R_α está dado por (3.3) y la regla de elección del parámetro α_d está definida mediante el Principio de Discrepancia (4.6), es convergente para todo $y \in Ran(T)$ y de orden óptimo en $\mathcal{X}_{\mu,\rho}$ para todo $\mu \in (0, \mu_0 - \frac{1}{2}]$ y para todo $\rho > 0$.*

Demostración. En primer lugar notemos que se puede suponer sin pérdida de generalidad que para todo δ suficientemente pequeño, $\alpha_d(\delta, y^\delta) < +\infty$ para todo $y^\delta \in \mathcal{Y}$ tal que $\left\|y - y^\delta\right\| \le \delta$. En efecto, supongamos que existe una sucesión $\delta_n \to 0^+$ tal que para los correspondientes y^{δ_n} se tiene que $\alpha_d(\delta_n, y^{\delta_n}) = +\infty$ para todo $n \in \mathbb{N}$. Como G_α satisface la condición (3.23) se sigue de (4.8) que $x^{\delta_n}_\infty = 0$ y en virtud de (4.6) resulta que $\left\|y^{\delta_n}\right\| \le \tau\delta_n$ para todo n. Esto, junto con el hecho que $\left\|y - y^{\delta_n}\right\| \le \delta_n$, implica que $\|y\| \leq (\tau+1)\delta_n$ para todo $n \in \mathbb{N}$ y por lo tanto $y = 0$ con lo cual se tiene el caso trivial $x^\dagger \doteq T^\dagger y = 0$.

Sean $\mu \in (0, \mu_0 - \frac{1}{2}]$, $\rho > 0$, $y \in Ran(T)$ y $x^\dagger \doteq T^\dagger y \in \mathcal{X}_{\mu,\rho}$, es decir, existe $w \in \mathcal{X}$ con $\|w\| \leq \rho$ tal que $x^\dagger = (T^*T)^\mu w$. Sea $y^\delta \in \mathcal{Y}$ tal que $\left\|y - y^\delta\right\| \leq \delta$ y $\alpha_d \doteq \alpha_d(\delta, y^\delta)$. Para probar que el método de regularización $(\{R_\alpha\}, \alpha_d)$ es de orden óptimo en $\mathcal{X}_{\mu,\rho}$ debemos demostrar que existe una constante positiva c independiente de δ y de $x^\dagger \in \mathcal{X}_{\mu,\rho}$ tal que $\left\|x^\delta_{\alpha_d} - x^\dagger\right\| \leq c\, \rho^{\frac{1}{2\mu+1}} \delta^{\frac{2\mu}{2\mu+1}}$. Para ello, en virtud de (2.20) es suficiente probar que tanto para el error de regularización $\left\|x_{\alpha_d} - x^\dagger\right\|$ como para el error asociado al ruido $\left\|x_{\alpha_d} - x^\delta_{\alpha_d}\right\|$ pueden obtenerse similares acotaciones.

Para el error de regularización, se sigue de (3.12) que

$$\left\|x_{\alpha_d} - x^\dagger\right\| = \|r_{\alpha_d}(T^*T)(T^*T)^\mu w\| = \|(T^*T)^\mu r_{\alpha_d}(T^*T)w\|.$$

De esto, de la desigualdad de interpolación (1.42) aplicada a $x \doteq r_{\alpha_d}(T^*T)w$ con $r = \mu$ y $q = \mu + \frac{1}{2}$ y recordando que $(T^*T)^k$ y $r_{\alpha_d}(T^*T)$ conmutan para todo $k \in \mathbb{R}$, se obtiene que

$$\begin{aligned} \left\|x_{\alpha_d} - x^\dagger\right\| &\leq \|r_{\alpha_d}(T^*T)w\|^{\frac{1}{2\mu+1}} \left\|(T^*T)^{\frac{1}{2}} r_{\alpha_d}(T^*T)(T^*T)^\mu w\right\|^{\frac{2\mu}{2\mu+1}} \\ &= \|r_{\alpha_d}(T^*T)w\|^{\frac{1}{2\mu+1}} \left\|(T^*T)^{\frac{1}{2}} r_{\alpha_d}(T^*T)x^\dagger\right\|^{\frac{2\mu}{2\mu+1}}. \end{aligned} \tag{4.11}$$

Ahora,

$$\begin{aligned} \left\|(T^*T)^{\frac{1}{2}} r_{\alpha_d}(T^*T)x^\dagger\right\| &= \left\|T r_{\alpha_d}(T^*T)x^\dagger\right\| && \text{(por (3.13))} \\ &= \left\|T(g_{\alpha_d}(T^*T)T^*T - I)x^\dagger\right\| && \text{(por (3.7))} \\ &= \left\|(T g_{\alpha_d}(T^*T)T^*T - T)T^\dagger y\right\| && \text{(pues } x^\dagger = T^\dagger y) \\ &= \left\|(T g_{\alpha_d}(T^*T)T^* - I)TT^\dagger y\right\| \\ &= \|TR_{\alpha_d} y - y\| && \text{(por (3.3) y por (1.7) pues } y \in Ran(T)) \\ &= \|T x_{\alpha_d} - y\|. \end{aligned} \tag{4.12}$$

Sustituyendo con (4.12) en (4.11) obtenemos entonces que

$$\left\|x_{\alpha_d} - x^\dagger\right\| \leq \|r_{\alpha_d}(T^*T)w\|^{\frac{1}{2\mu+1}} \|T x_{\alpha_d} - y\|^{\frac{2\mu}{2\mu+1}},$$

y definiendo γ como el supremo en (4.5), es decir,

$$\gamma \doteq \sup_{\alpha>0,\, \lambda\in[0,\|T\|^2]} |r_\alpha(\lambda)|\,, \tag{4.13}$$

resulta que

$$\left\|x_{\alpha_d} - x^\dagger\right\| \leq (\gamma\rho)^{\frac{1}{2\mu+1}} \left\|Tx_{\alpha_d} - y\right\|^{\frac{2\mu}{2\mu+1}}, \tag{4.14}$$

donde hemos utilizado que $\|w\| \leq \rho$.

Ahora, como $x_{\alpha_d} - x^\delta_{\alpha_d} = g_{\alpha_d}(T^*T)T^*(y-y^\delta)$, se sigue que

$$\begin{aligned}
\left\|T(x_{\alpha_d} - x^\delta_{\alpha_d}) - (y-y^\delta)\right\| &= \left\|(Tg_{\alpha_d}(T^*T)T^* - I)(y-y^\delta)\right\| \\
&= \left\|(TT^*g_{\alpha_d}(TT^*) - I)(y-y^\delta)\right\| \quad \text{(por (1.35))} \\
&= \left\|r_{\alpha_d}(TT^*)(y-y^\delta)\right\| \quad \text{(pues } 1-\lambda g_\alpha(\lambda) = r_\alpha(\lambda)) \\
&\leq \gamma\delta \qquad \text{(por (4.13) y } \left\|y-y^\delta\right\| \leq \delta).
\end{aligned} \tag{4.15}$$

Luego,

$$\begin{aligned}
\|Tx_{\alpha_d} - y\| &\leq \left\|Tx^\delta_{\alpha_d} - y^\delta\right\| + \left\|T(x_{\alpha_d} - x^\delta_{\alpha_d}) - (y-y^\delta)\right\| \\
&\leq (\tau+\gamma)\delta \quad \text{(por (4.7) y (4.15))}.
\end{aligned} \tag{4.16}$$

Sustituyendo (4.16) en (4.14) se obtiene la acotación que queríamos encontrar para el error de regularización, es decir,

$$\left\|x_{\alpha_d} - x^\dagger\right\| \leq c_1 \rho^{\frac{1}{2\mu+1}} \delta^{\frac{2\mu}{2\mu+1}},$$

donde $c_1 = \gamma^{\frac{1}{2\mu+1}}(\tau+\gamma)^{\frac{2\mu}{2\mu+1}}$.

Vayamos ahora al error asociado al ruido. Se deduce del Teorema 3.4 y del hecho que G_α satisface la condición (3.23) que

$$\left\|x_{\alpha_d} - x^\delta_{\alpha_d}\right\| \leq k\delta\sqrt{\frac{C}{\alpha_d}},$$

donde C es la constante de la hipótesis (*H2*) y k es una constante positiva independiente de α_d. Luego, para obtener la acotación que necesitamos es suficiente con probar que existe una constante positiva c_2 tal que

$$\alpha_d \geq c_2 \left(\frac{\delta}{\rho}\right)^{\frac{2}{2\mu+1}}. \tag{4.17}$$

Por definición de α_d y puesto que $\alpha_d < +\infty$ se tiene que

$$\left\|Tx^\delta_{2\alpha_d} - y^\delta\right\| > \tau\delta. \tag{4.18}$$

Usando la desigualdad triangular y (4.15) con $2\alpha_d$ en lugar de α_d resulta que

$$\left\|Tx^\delta_{2\alpha_d} - y^\delta\right\| - \|Tx_{2\alpha_d} - y\| \leq \left\|T(x_{2\alpha_d} - x^\delta_{2\alpha_d}) - (y - y^\delta)\right\| \leq \gamma\delta.$$

Esto junto a (4.18) implica que

$$\|Tx_{2\alpha_d} - y\| \geq \left\|Tx^\delta_{2\alpha_d} - y^\delta\right\| - \gamma\delta \geq (\tau - \gamma)\delta, \tag{4.19}$$

donde $\tau - \gamma > 0$ en virtud de (4.5) y (4.13) (aquí se ve claramente por qué la elección de τ dada por (4.5) es crucial). Como $\{g_\alpha\}$ tiene calificación clásica de orden μ_0 y $\mu + \frac{1}{2} \leq \mu_0$, se tiene que existe una constante positiva $\tilde{c}$ independiente de α tal que $\omega_{\mu+\frac{1}{2}}(\alpha) \doteq \tilde{c}\alpha^{\mu+\frac{1}{2}}$ satisface la condición (3.9). Luego, se sigue del hecho que $Tx^\dagger = TT^\dagger y = y$ junto con el Teorema 3.5 que

$$\|Tx_{2\alpha_d} - y\| = \left\|Tx_{2\alpha_d} - Tx^\dagger\right\| \leq \omega_{\mu+\frac{1}{2}}(2\alpha_d)\rho = \tilde{c}(2\alpha_d)^{\mu+\frac{1}{2}}\rho.$$

De esto y (4.19) se deduce (4.17) con $c_2 \doteq (\tau - \gamma)/(2^{\mu+\frac{1}{2}}\tilde{c})$, con lo cual queda probado que el método de regularización $(\{R_\alpha\}, \alpha_d)$ es de orden óptimo en $\mathcal{X}_{\mu,\rho}$.

Finalmente, para probar la convergencia de este método de regularización para todo $y \in Ran(T)$ usamos el Teorema 2.17. Notar que se satisfacen las hipótesis de este teorema con $\mu^* = \mu_0 - \frac{1}{2}$ y $\tau_0 \doteq \gamma$, donde γ está dada por (4.13). Observar que la regla de elección del parámetro α_τ definida en (2.31) corresponde precisamente a la dada por el Principio de Discrepancia (4.6) con el respectivo parámetro τ. En efecto, sea $\alpha(\delta, y^\delta) \doteq \alpha_d(\frac{\delta}{\tau}, y^\delta)$. Entonces se tiene que $\alpha_\tau(\delta, y^\delta) = \alpha(\tau\delta, y^\delta) = \alpha_d(\delta, y^\delta)$. ■

Observación 4.4. Es fácil verificar que todas las conclusiones del Teorema 4.3 siguen siendo válidas para cualquier regla de elección del parámetro $\hat{\alpha} = \hat{\alpha}(\delta, y^\delta)$ que satisfaga

$$\left\|Tx_{\hat{\alpha}}^\delta - y^\delta\right\| \le \tau\delta \le \left\|Tx_\beta^\delta - y^\delta\right\| \tag{4.20}$$

para algún β tal que $\hat{\alpha} \le \beta \le 2\hat{\alpha}$, donde τ está dado por (4.5). (Notar que si $\hat{\alpha}$ satisface (4.20) entonces en virtud de (4.6) se sigue que necesariamente $\hat{\alpha} \le \alpha_d$). Esto es importante desde el punto de vista práctico ya que proporciona un cierto margen de libertad en la determinación de la regla $\alpha_d(\delta, y^\delta)$ a partir de (4.6).

Observación 4.5. En virtud de (4.3), la desigualdad (4.17) permite afirmar que la regla de elección del parámetro *a-posteriori* dada por el Principio de Discrepancia (independiente de los parámetros μ y ρ, desconocidos en este caso) nunca es “mejor” que la regla de elección del parámetro *a-priori* que se obtiene a partir del conocimiento de μ y ρ. Esto es razonable pues precisamente el Principio de Discrepancia no utiliza ninguna información *a-priori* sobre la solución exacta.

Si $\{g_\alpha\}$ es un MRE para el cual $\mu_0 = +\infty$, entonces la condición (3.9) con ω_μ dado por (3.14) se satisface para todo $\mu > 0$. En tal caso se sigue de la demostración del Teorema 4.3 que el método $(\{R_\alpha\}, \alpha_d)$ es de orden óptimo en $\mathcal{X}_{\mu,\rho}$ para todos μ, $\rho > 0$. Esto ocurre, por ejemplo, con el método TSVD y el método de Showalter presentados en la Sección 3.2.

Es oportuno notar que mediante argumentos similares a los usados en la demostración del Teorema 4.3 se puede probar que el método de regularización $(\{R_\alpha\}, \alpha_d)$, donde R_α está dado por (3.3) y la regla de elección del parámetro α_d está definida mediante el Principio de Discrepancia (4.6), tiene la propiedad

$$\left\|x_{\alpha_d(\delta,y^\delta)}^\delta - x^\dagger\right\| = o\left(\delta^{\frac{2\mu}{2\mu+1}}\right)$$

para $x^\dagger \doteq K^\dagger y \in \mathcal{X}_\mu$ y $\mu \in [0, \mu_0 - \frac{1}{2})$, donde μ_0 es como en la Definición 3.6. En general, este método de regularización no es de orden óptimo en $\mathcal{X}_\mu$ para $\mu > \mu_0 - \frac{1}{2}$, como se puede deducir a partir del siguiente resultado para el método de Tikhonov-Phillips.

Proposición 4.6. *Sean $\mathcal{X}$, $\mathcal{Y}$ espacios de Hilbert, $K : \mathcal{X} \to \mathcal{Y}$ un operador lineal y compacto con sistema singular $\{(\sigma_n; u_n, v_n)\}$, $R_\alpha \doteq (K^*K + \alpha I)^{-1}K^*$ y $\alpha_d(\delta, y^\delta)$ la regla de elección del parámetro definida mediante el Principio de Discrepancia (4.6). Si existe $m \in \mathbb{N}$ tal que para todo $y \in \operatorname{span}\{v_m\}$ y para todo $y^\delta \in \mathcal{Y}$ con $\left\| y - y^\delta \right\| \leq \delta$ se verifica*

$$\left\| x^\delta_{\alpha_d(\delta, y^\delta)} - K^\dagger y \right\| = o\left(\sqrt{\delta}\right), \tag{4.21}$$

entonces $\dim(Ran(K)) < +\infty$ y por lo tanto el problema $Kx = y$ es bien condicionado.

Demostración. Supongamos que $\dim(Ran(K)) = +\infty$, entonces hay infinitos valores singulares σ_n y

$$\lim_{n \to +\infty} \sigma_n = 0. \tag{4.22}$$

Sean $y \doteq v_m$, $\{\delta_n\} \subset \mathbb{R}^+$ una sucesión arbitraria tal que $\delta_n \overset{n\to\infty}{\to} 0^+$, $y^{\delta_n} \doteq y + \delta_n v_n$ y $\alpha_n \doteq \alpha_d(\delta_n, y^{\delta_n})$. Entonces $\left\| y - y^{\delta_n} \right\| = \delta_n$ y se sigue de (1.18) que

$$x^\dagger \doteq K^\dagger y = K^\dagger v_m = \sum_{n=1}^{\infty} \frac{\langle v_m, v_n \rangle}{\sigma_n} u_n = \frac{u_m}{\sigma_m}.$$

Luego,

$$\begin{aligned}
x^{\delta_n}_{\alpha_n} - x^\dagger &= R_{\alpha_n} y^{\delta_n} - \frac{u_m}{\sigma_m} \\
&= R_{\alpha_n}(v_m + \delta_n v_n) - \frac{u_m}{\sigma_m} \\
&= (K^*K + \alpha_n I)^{-1} K^* v_m + \delta_n (K^*K + \alpha_n I)^{-1} K^* v_n - \frac{u_m}{\sigma_m} \qquad \text{(por def. de } R_\alpha) \\
&= \frac{\sigma_m}{\sigma_m^2 + \alpha_n} u_m + \frac{\delta_n \sigma_n}{\sigma_n^2 + \alpha_n} u_n - \frac{u_m}{\sigma_m} \qquad (\text{pues } K^* v_j = \sigma_j u_j \; \forall\, j \in \mathbb{N}).
\end{aligned}$$

Eligiendo $\delta_n \doteq \sigma_n^2$ (lo cual es posible en virtud de (4.22)), para $n \neq m$ se tiene entonces que

$$\left\| x^{\delta_n}_{\alpha_n} - x^\dagger \right\|^2 \geq \frac{\delta_n^3}{(\delta_n + \alpha_n)^2} = \left(\frac{\sqrt{\delta_n}}{1 + \frac{\alpha_n}{\delta_n}} \right)^2. \tag{4.23}$$

Puesto que $y \doteq v_m$, se satisface (4.21), lo cual implica junto a (4.23) que

$$\frac{\sqrt{\delta_n}}{1+\frac{\alpha_n}{\delta_n}} = o(\sqrt{\delta_n})$$

y por lo tanto

$$\lim_{n\to+\infty} \frac{\alpha_n}{\delta_n} = +\infty. \tag{4.24}$$

Por otra parte, se sigue de (4.7) que

$$\left\|y^{\delta_n}\right\| - \tau\delta_n \leq \left\|y^{\delta_n}\right\| - \left\|Kx_{\alpha_n}^{\delta_n} - y^{\delta_n}\right\| \leq \left\|Kx_{\alpha_n}^{\delta_n}\right\|. \tag{4.25}$$

Ahora bien, como

$$\begin{aligned} Kx_{\alpha_n}^{\delta_n} &= \frac{1}{\alpha_n}K[(\alpha_n I + K^*K) - K^*K]x_{\alpha_n}^{\delta_n} \\ &= \frac{1}{\alpha_n}K[(\alpha_n I + K^*K) - K^*K](K^*K+\alpha_n I)^{-1}K^*y^{\delta_n} \\ &= \frac{1}{\alpha_n}KK^*[y^{\delta_n} - Kx_{\alpha_n}^{\delta_n}], \end{aligned}$$

resulta que

$$\left\|Kx_{\alpha_n}^{\delta_n}\right\| \leq \frac{\|K\|^2}{\alpha_n}\left\|Kx_{\alpha_n}^{\delta_n} - y^{\delta_n}\right\| \leq \frac{\|K\|^2}{\alpha_n}\tau\delta_n, \tag{4.26}$$

donde la última desigualdad se sigue nuevamente de (4.7). De (4.26) y (4.25) se tiene entonces que

$$\left\|y^{\delta_n}\right\| - \tau\delta_n \leq \frac{\|K\|^2}{\alpha_n}\tau\delta_n,$$

de donde

$$\frac{\left\|y^{\delta_n}\right\| - \tau\delta_n}{\|K\|^2\tau} \leq \frac{\delta_n}{\alpha_n},$$

y así, haciendo tender n a $+\infty$, se deduce que

$$0 < \frac{1}{\|K\|^2\tau} \leq \limsup_{n\to+\infty}\frac{\delta_n}{\alpha_n}, \quad (\text{pues } \left\|y^{\delta_n}\right\| = 1+\delta_n)$$

lo cual contradice (4.24). Esta contradicción proviene de suponer que $\dim(Ran(K)) = +\infty$. Por lo tanto, se tiene que $\dim(Ran(K)) < +\infty$ como queríamos probar. ■

El resultado anterior nos dice que para la regularización de Tikhonov-Phillips $\{R_\alpha\}$ (en cuyo caso $\mu_0 = 1$) con la regla de elección del parámetro α_d dada por el Principio de Discrepancia, el error total $\left\|x^\delta_{\alpha_d(\delta,y^\delta)} - K^\dagger y\right\|$ no puede converger a cero estrictamente más rápido que $O(\sqrt{\delta})$, a menos que $Ran(K)$ tenga dimensión finita, es decir, que el problema $Kx = y$ sea bien condicionado. Luego, el método de regularización $(\{R_\alpha,\}, \alpha_d)$ no es de orden óptimo en $\mathcal{X}_\mu$ para ningún $\mu > \mu_0 - \frac{1}{2} = \frac{1}{2}$. En efecto, si lo fuese, entonces (4.21) se verifica para todo $x^\dagger \doteq K^\dagger y \in \mathcal{X}_\mu$. Ahora bien, se sigue de la Proposición 2.11 que si $y \in \text{span}\{v_m\}$ para cualquier $m \in \mathbb{N}$, entonces $K^\dagger y \in \mathcal{X}_\mu$ para todo $\mu > 0$. Luego, (4.21) se satisface, en particular, para todo $y \in \text{span}\{v_m\}$ para todo $m \in \mathbb{N}$ y en tal caso el Teorema 4.3 implicaría que $Ran(K)$ tiene dimensión finita y por lo tanto sería cerrado (recordar que según la Definición 2.16, los métodos de orden óptimo en $\mathcal{X}_\mu \doteq Ran((T^*T)^\mu)$ se definen sólo para operadores con rango no cerrado).

4.3. Reglas heurísticas de elección del parámetro

En esta sección describiremos dos reglas heurísticas frecuentemente utilizadas para elegir el parámetro de regularización en problemas concretos: el criterio de la curva L y el de validación cruzada generalizada (VCG). Estas reglas son "puramente *a-posteriori*", pues dependen sólo del dato y^δ y no del nivel de ruido δ. Por este motivo se las conoce también como "reglas libres de error". Como vimos en la Sección 2.1, A. B. Bakushinskii ([6]) probó que tales reglas no pueden formar parte de métodos de regularización convergentes para un problema mal condicionado (ver Teorema 2.2), lo cual no implica que puedan tener un buen comportamiento para niveles de ruido δ pequeños, como en efecto sucede frecuentemente.

4.3.1. El criterio de la curva L

En el año 1992 P. C. Hansen ([18]) propuso una regla de elección del parámetro libre de error conocida como "el criterio de la curva L". La regla de elección surge de relacionar la norma de los residuos de las aproximaciones calculadas al aplicar un método de regularización, con la norma de las aproximaciones propiamente dichas. Esto se hace graficando $\|x_\alpha^\delta\|$ versus $\|y^\delta - Tx_\alpha^\delta\|$ en una escala $\log - \log$ considerando un amplio rango de valores de α. La motivación y fundamentación heurística es la siguiente. Cuando los valores de α no son muy pequeños, se tiene que $\|y^\delta - Tx_\alpha^\delta\|$ decrece con una tasa de convergencia de orden $O(\alpha^\nu)$, donde el valor del parámetro $\nu > 0$ depende de la regularidad de la solución y del método de regularización utilizado; mientras que en tal caso $\|x_\alpha^\delta\|$ es típicamente del mismo orden de magnitud que $\|x^\dagger\|$, y varía relativamente poco cuando α decrece. Una consecuencia de esto es que la parte de la gráfica de la curva $(\|y^\delta - Tx_\alpha^\delta\|, \|x_\alpha^\delta\|)$ que corresponde a valores de α no muy pequeños, en general es casi horizontal. Por otra parte, para valores de α cercanos a cero (es decir, para valores pequeños con respecto a su *valor óptimo*) la norma de las aproximaciones tiende a infinito (por ejemplo con el orden $O(\delta/\sqrt{\alpha})$), mientras que la norma del residuo se mantiene en el orden de δ. Esto hace que para valores de α muy pequeños la gráfica de la curva $(\|y^\delta - Tx_\alpha^\delta\|, \|x_\alpha^\delta\|)$ sea empinada. Luego, la curva resultante tiene una forma similar a la letra "L" y de ahí el nombre de este criterio (ver, por ejemplo, las Figuras 8.27 y 8.28 en la Sección 8.1).

Observando las gráficas de estas curvas se puede inferir que para valores del parámetro de regularización α cercanos a la "esquina" de la curva L, tanto la norma del residuo como la norma de la solución regularizada se mantienen pequeñas y por ello, es razonable seleccionar como parámetro al valor de α correspondiente a la esquina de dicha curva. Una dificultad práctica de este método es que generalmente hay varios valores del parámetro de regularización cercanos a dicha esquina y por ello, es conveniente utilizar métodos numéricos de optimización para hallar el valor óptimo. Por ejemplo, P. C. Hansen y D. P. O'Leary propusieron un algoritmo que busca el punto del gráfico con máxima curvatura (ver [22]).

Es importante señalar aquí que Hansen probó para el método de Tikhonov-Phillips que la norma de la solución regularizada es una función monótona decreciente de la norma del residuo (ver [18], Teorema 1). Sin embargo, esto ocurre también para otros métodos de regularización espectrales para los cuales para cada λ fijo, se satisface que $g_\alpha(\lambda)$ y $r_\alpha(\lambda)$ como funciones de α son monótonas creciente y decreciente, respectivamente. Esta es una condición necesaria para aplicar el criterio de la curva L, aunque no garantiza que la curva tenga efectivamente forma de "L".

Para el método TSVD, como veremos en la Sección 8.1, el gráfico de la curva L está dado por un conjunto discreto de puntos, lo cual es consecuencia de la definición del método (ver por ejemplo la Figura 8.21).

4.3.2. Validación cruzada generalizada

Otro método muy utilizado para elegir el parámetro de regularización es el llamado criterio de "validación cruzada generalizada" (VCG), el cual fue introducido en el año 1979 por G. H. Golub, M. T. Heath y G. Wahba ([16]). Este método se aplica a problemas donde el operador K es compacto con rango de dimensión finita. Uno de tales problemas es el de *discretización de momentos* introducido en el Ejemplo 2.26, que surge al discretizar una ecuación integral de Fredholm de primera clase. El criterio de validación cruzada generalizada se origina a partir de consideraciones estadísticas y uno de los supuestos fundamentales es que la perturbación en los datos es de tipo ruido blanco discreto, lo cual permite realizar un análisis más detallado del error ([12]).

Con VCG el parámetro de regularización óptimo se elige como el valor de α que minimiza el funcional $V(\alpha)$ definido por

$$V(\alpha) \doteq \left(\frac{\left\| y^\delta - K x_\alpha^\delta \right\|}{\text{traza}\{\frac{1}{m} r_\alpha(KK^*)\}} \right)^2 .$$

Si bien este método es frecuentemente utilizado, presenta algunas dificultades. Por ejemplo, a menudo el gráfico del funcional $V(\alpha)$ es aplanado y tiene mínimos locales, con lo cual resulta dificultoso localizar numéricamente el mínimo global. Otra de las dificultades es que el método puede fallar cuando los errores son altamente correlacionados. Más detalles de este método se pueden encontrar en [65].

Capítulo 5

Divergencia con velocidad arbitraria del método de mínimos cuadrados en dimensión infinita

Sean $\mathcal{X}$, $\mathcal{Y}$ espacios de Hilbert, $T : \mathcal{X} \to \mathcal{Y}$ un operador lineal y continuo, $\{\mathcal{X}_N\}$ una sucesión creciente de subespacios de dimensión finita de $\mathcal{X}$ tales que $\overline{\bigcup_{N=1}^{\infty} \mathcal{X}_N} = \mathcal{X}$ y $x_N = T_N^{\dagger} y$, donde $T_N = TP_{\mathcal{X}_N}$ y $P_{\mathcal{X}_N}$ es la proyección ortogonal de $\mathcal{X}$ sobre $\mathcal{X}_N$. Como mencionamos en la Sección 2.3, en 1980 T. Seidman ([54]) probó que si el problema $Tx = y$ es mal condicionado, entonces, sin imponer ninguna condición de regularidad a $x^{\dagger}$, no se puede garantizar que la sucesión de soluciones de mínimos cuadrados $\{x_N\}$ converja a $x^{\dagger}$ y que es posible que $\left\| x_N - x^{\dagger} \right\|$ se incremente sin cota. En este capítulo extenderemos el resultado anterior en el siguiente sentido: probaremos que si los subespacios $\mathcal{X}_N$ no se eligen cuidadosamente, es posible que las soluciones de mínimos cuadrados x_N diverjan de la solución exacta $x^{\dagger}$ con cualquier velocidad arbitrariamente grande, prescrita *a-priori*. En la Sección 5.1 probaremos resultados de divergencia para problemas inversos con datos exactos, mientras que en la Sección 5.2

lo haremos para el caso en que sólo se dispone de datos con ruido. Algunos de los resultados originales que presentamos en este capítulo han sido ya publicados en la revista Inverse Problems (ver [57]).

5.1. Divergencia para problemas inversos con datos exactos

En esta sección, generalizaremos los resultados del Teorema 2.22 y del Corolario 2.24. Más precisamente, mostraremos que para cualquier espacio de Hilbert $\mathcal{X}$, separable y de dimensión infinita, para cualquier $b \in \mathcal{X}$, $b \neq 0$, y para cualquier función $f : \mathbb{N} \to \mathbb{R}^+$, es posible encontrar un operador inyectivo A y una sucesión de subespacios $\mathcal{X}_N \subset \mathcal{X}$ tales que $\|x_N - A^{-1}b\| \geq f(N)$ para todo $N \in \mathbb{N}$, donde x_N es la solución de mínimos cuadradros del problema $Ax = b$ en $\mathcal{X}_N$. Necesitaremos introducir previamente las siguientes definiciones.

Definición 5.1. Sean $\mathcal{X}$ un espacio de Hilbert separable de dimensión infinita y $B \doteq \{e_n\}_{n=1}^{\infty}$ una base ortonormal de $\mathcal{X}$. Decimos que un elemento $x \in \mathcal{X}$ es "*degenerado con respecto a* B" si es combinación lineal finita de elementos de B; caso contrario se dice que x es "*no degenerado con respecto a* B". En particular, si $\langle x, e_n \rangle \neq 0 \ \forall\, n \in \mathbb{N}$ se dice que x es "*fuertemente no degenerado con respecto a* B". Si $\langle x, e_n \rangle \neq 0$ para infinitos valores de n, se dice que x es "*débilmente no degenerado con respecto a* B". Claramente, todo elemento fuertemente no degenerado con respecto a una base es también débilmente no degenerado con respecto a la misma base.

A continuación, probaremos dos lemas que necesitaremos luego, para demostrar los resultados principales.

Lema 5.2. *Sean $\mathcal{X}$ un espacio de Hilbert separable de dimensión infinita y $b \in \mathcal{X}$,*

$b \neq 0$. *Entonces existe una base ortonormal B de $\mathcal{X}$ tal que b es fuertemente no degenerado con respecto a B.*

Demostración. Dado $b \in \mathcal{X}$, $b \neq 0$, sea $\{g_n\}_{n=1}^{\infty}$ una base ortonormal de $[\mathrm{span}\{b\}]^{\perp}$ y definamos $f_n \doteq b + g_n$, $n \in \mathbb{N}$. Supongamos que para una cierta sucesión de escalares $\alpha_1, \alpha_2, \ldots$, se tiene que

$$0 = \sum_n \alpha_n f_n = \left(\sum_n \alpha_n\right) b + \sum_n \alpha_n g_n.$$

Como $\{b\} \cup \{g_n\}_{n=1}^{\infty}$ es un sistema ortogonal, es linealmente independiente y por lo tanto, la ecuación anterior implica que $\alpha_n = 0$ para todo $n \in \mathbb{N}$. Luego, $\{f_n\}_{n=1}^{\infty}$ es linealmente independiente.

Usaremos ahora el proceso de Gram-Schmidt (ver [64]) para construir un sistema ortonormal a partir de f_n (notar que $\langle f_i, f_j \rangle = \|b\|^2 + \delta_{i,j}$ para i, $j \in \mathbb{N}$). Para ello, definimos

$$e_1 \doteq \frac{f_1}{\|f_1\|},$$

y para $n \geq 2$, definimos e_n recursivamente como

$$e_n \doteq \left(f_n - \sum_{j=1}^{n-1} \langle f_n, e_j \rangle e_j\right) \left\| f_n - \sum_{j=1}^{n-1} \langle f_n, e_j \rangle e_j \right\|^{-1}.$$

Entonces, resulta que

$$\mathrm{span}\{f_n\}_{n=1}^{\infty} = \mathrm{span}\{e_n\}_{n=1}^{\infty} \tag{5.1}$$

y $\{e_n\}_{n=1}^{\infty}$ es un sistema ortonormal en $\mathcal{X}$ (ver [64]). Además, este sistema es completo en $\mathcal{X}$. En efecto, supongamos que existe $y_1 \in \mathcal{X}$ tal que $\langle y_1, e_n \rangle = 0$ para todo $n \in \mathbb{N}$. Esto implica que $\langle y_1, f_n \rangle = 0$ para todo $n \in \mathbb{N}$ en virtud de (5.1) y por lo tanto

$$\langle y_1, g_n \rangle = -\langle y_1, b \rangle \qquad \text{para todo } n \in \mathbb{N}. \tag{5.2}$$

Pero, puesto que $\{g_n\}_{n=1}^{\infty}$ es un sistema ortonormal en $\mathcal{X}$, la sucesión $\{\langle y_1, g_n \rangle\}_{n=1}^{\infty}$ debe estar en ℓ^2. Así, como por (5.2) esta sucesión es constante, se tiene que $\langle y_1, g_n \rangle = 0$ para todo $n \in \mathbb{N}$, lo cual nuevamente en virtud de (5.2), implica que también $\langle y_1, b\rangle = 0$. Así, $y_1 = 0$ puesto que $\{b, g_1, g_2, \ldots\}$ es completo en $\mathcal{X}$. Luego $\{e_n\}_{n=1}^{\infty}$ es un sistema ortonormal completo en $\mathcal{X}$.

Probaremos a continuación que b es fuertemente no degenerado con respecto a la base $B \doteq \{e_n\}_{n=1}^{\infty}$. Supongamos que no lo es, es decir, que existe $n^* \in \mathbb{N}$ tal que $\langle e_{n^*}, b\rangle = 0$ y $\langle e_n, b\rangle \neq 0$ para todo $n < n^*$. Como $\langle e_1, b\rangle = \|f_1\|^{-1}\|b\|^2 \neq 0$ se tiene que $n^* \geq 2$. Sean $F_n \doteq \text{span}\{f_1, \ldots, f_n\}$ y $E_n \doteq \text{span}\{e_1, \ldots, e_n\}$. Puesto que $\{f_n\}_{n=1}^{\infty}$ es un conjunto linealmente independiente por el proceso de Gram-Schmidt, resulta que $F_n = E_n$ para todo $n \in \mathbb{N}$ (ver [64]). Sean ahora P_{F_n} y P_{E_n} las proyecciones ortogonales de $\mathcal{X}$ sobre F_n y E_n, respectivamente. Como $\langle e_{n^*}, b\rangle = 0$, se sigue que

$$P_{F_{n^*}} b = P_{E_{n^*}} b = P_{E_{n^*-1}} b \; = \; P_{F_{n^*-1}} b,$$

y por lo tanto $\langle f_{n^*}, b\rangle = 0$. Pero esto contradice el hecho que $\langle f_{n^*}, b\rangle = \langle b, b\rangle + \langle g_{n^*}, b\rangle = \|b\|^2 \neq 0$. Así, tal n^* no existe y $\langle e_n, b\rangle \neq 0$ para todo $n \in \mathbb{N}$. Luego, concluímos que b es fuertemente no degenerado con respecto a la base B de $\mathcal{X}$. ■

El siguiente lema está basado en las ideas presentadas en el Ejemplo 3.1 de [54].

Lema 5.3. *Sean $\mathcal{X}$ un espacio de Hilbert separable de dimensión infinita, $B \doteq \{e_n\}_{n=1}^{\infty}$ una base ortonormal de $\mathcal{X}$ y sean $\{\alpha_n\}_{n=1}^{\infty}$, $\{\beta_n\}_{n=1}^{\infty}$ dos sucesiones en ℓ^2 que satisfacen*

$$\alpha_n \neq 0 \quad \forall\, n \in \mathbb{N}, \qquad \beta_1 = 0, \qquad \beta_n \neq 0 \quad \forall\, n \geq 2. \tag{5.3}$$

Entonces, dado $x = \sum_{n=1}^{\infty} \xi_n e_n \in \mathcal{X}$, el operador lineal $A : \mathcal{X} \to \mathcal{X}$ definido mediante

$$Ax \doteq \sum_{n=1}^{\infty} (\alpha_n \xi_n + \beta_n \xi_1)\, e_n \tag{5.4}$$

es compacto, inyectivo y de rango denso en $\mathcal{X}$.

Demostración. Veamos primero que el operador A es compacto. Para cada $N \in \mathbb{N}$, sea $\tilde{A}_N : \mathcal{X} \to \mathcal{X}_N$ el operador definido mediante

$$\tilde{A}_N x \doteq \sum_{i=1}^{N} (\alpha_i \xi_i + \beta_i \xi_1)\, e_i,$$

donde $x = \sum_{n=1}^{\infty} \xi_n e_n \in \mathcal{X}$ y $\mathcal{X}_N \doteq \text{span}\{e_1, \ldots, e_N\}$. Para todo $N \in \mathbb{N}$, $\tilde{A}_N$ es un operador compacto puesto que claramente es un operador de rango finito. Además,

$$\begin{aligned}
\left\|Ax - \tilde{A}_N x\right\|^2 &= \left\|\sum_{i=N+1}^{\infty} (\alpha_i \xi_i + \beta_i \xi_1)\, e_i\right\|^2 \\
&= \sum_{i=N+1}^{\infty} (\alpha_i \xi_i + \beta_i \xi_1)^2 \\
&\leq \|x\|^2 \left(\sum_{i=N+1}^{\infty} \alpha_i^2 + 2 \sum_{i=N+1}^{\infty} \alpha_i \beta_i + \sum_{i=N+1}^{\infty} \beta_i^2\right) \\
&\leq \|x\|^2 \left[\sum_{i=N+1}^{\infty} \alpha_i^2 + 2 \left(\sum_{i=N+1}^{\infty} \alpha_i^2\right)^{\frac{1}{2}} \left(\sum_{i=N+1}^{\infty} \beta_i^2\right)^{\frac{1}{2}} + \sum_{i=N+1}^{\infty} \beta_i^2\right],
\end{aligned}$$

donde la penúltima desigualdad se debe a que $\xi_i^2 \leq \|x\|^2$ para todo $i \in \mathbb{N}$, y en la última se utilizó la desigualdad de Cauchy-Schwarz. Como $\{\alpha_n\}$ y $\{\beta_n\}$ son dos sucesiones en ℓ^2, se tiene que $\left\|A - \tilde{A}_N\right\| \to 0$ cuando $N \to \infty$. Por lo tanto, A resulta compacto por ser límite uniforme de operadores compactos.

Veamos ahora que el operador A es inyectivo. Sea $x = \sum_{i=1}^{\infty} \xi_i e_i \in \mathcal{X}$ y supongamos que $Ax = 0$, entonces

$$\sum_{i=1}^{\infty} \alpha_i \xi_i e_i = -\sum_{i=1}^{\infty} \beta_i \xi_1 e_i.$$

Igualando los coeficientes de e_1 resulta que $\alpha_1 \xi_1 = -\beta_1 \xi_1$. Como $\beta_1 = 0$ y $\alpha_1 \neq 0$ se sigue que $\xi_1 = 0$. Similarmente para todo $i \geq 2$, $\alpha_i \xi_i = -\beta_i \xi_1 = 0$, lo cual implica que $\xi_i = 0$ puesto que $\xi_1 = 0$ y $\alpha_i \neq 0$. Luego, $x = 0$ y el operador A es inyectivo.

Probaremos ahora que $Ran(A)$ es denso en $\mathcal{X}$. Sea $x = \sum_{i=1}^{\infty} \xi_i e_i \in \mathcal{X}$. Para cada $N \in \mathbb{N}$, existen escalares $c_1, \ldots, c_N$ tales que $z_N \doteq \sum_{i=1}^{N} c_i e_i \in \mathcal{X}_N$ satisface

$$Az_N = \sum_{i=1}^{N} (\alpha_i c_i + \beta_i c_1) e_i = \sum_{i=1}^{N} \xi_i e_i.$$

En efecto, igualando los coeficientes de e_i, $i = 1, \ldots, N$, en esta expresión y recordando que $\beta_1 = 0$ y $\alpha_i \neq 0$ para todo $i \in \mathbb{N}$, se obtiene inmediatamente que $c_1 = \frac{\xi_1}{\alpha_1}$ y

$c_i = \alpha_i^{-1}\left(\xi_i - \frac{\beta_i \xi_1}{\alpha_1}\right)$ para $2 \leq i \leq N$. Además,

$$\|x - Az_N\|^2 = \left\| \sum_{i=N+1}^{\infty} \xi_i e_i \right\|^2 = \sum_{i=N+1}^{\infty} |\xi_i|^2$$

y como $\{\xi_i\}_{i=1}^{\infty}$ es una sucesión en ℓ^2 concluímos que $\|x - Az_N\| \to 0$ cuando $N \to \infty$. Luego, $Ran(A)$ es denso en $\mathcal{X}$, lo cual completa la demostración del lema. ∎

En el siguiente resultado probamos que para un espacio de Hilbert separable arbitrario $\mathcal{X}$ de dimensión infinita y para cualquier elemento no nulo $b \in \mathcal{X}$, existe un operador lineal A tal que las soluciones de mínimos cuadrados del problema $Ax = b$ en $\mathcal{X}_N$ divergen de la solución exacta con velocidad arbitrariamente grande.

Teorema 5.4. *Sean $\mathcal{X}$ un espacio de Hilbert separable de dimensión infinita y $f : \mathbb{N} \to \mathbb{R}^+$ una función creciente arbitraria. Entonces para cada $b \in \mathcal{X}$, $b \neq 0$, existen una sucesión creciente de subespacios $\mathcal{X}_N$ cuya unión es densa en $\mathcal{X}$ y un operador lineal $A = A(b, f) : \mathcal{X} \to \mathcal{X}$, compacto, inyectivo y de rango denso en $\mathcal{X}$, tal que $b = Ax^*$ para algún $x^* \in \mathcal{X}$, y si x_N denota la solución de mínimos cuadrados de $Ax = b$ en $\mathcal{X}_N$, entonces $\|x_N - x^*\| \geqslant f(N)$ para todo $N \in \mathbb{N}$.*

Demostración. Sea $b \in \mathcal{X}$, $b \neq 0$. Del Lema 5.2 se sigue que existe una base ortonormal $B \doteq \{e_n\}_{n=1}^{\infty}$ de $\mathcal{X}$ tal que b es fuertemente no degenerado con respecto a B. Definimos $\mathcal{X}_N \doteq \text{span}\{e_1, ..., e_N\}$. Sean $\{\alpha_n\}_{n=1}^{\infty}$ y $\{\beta_n\}_{n=1}^{\infty}$ dos sucesiones en ℓ^2 que satisfacen (5.3) y sea A el operador lineal definido en (5.4). En virtud del Lema 5.3 se tiene que A es compacto, inyectivo y de rango denso en $\mathcal{X}$.

Ahora veremos que dado cualquier elemento $b \in \mathcal{X}$ fuertemente no degenerado con respecto a B, es decir, tal que $b_n \doteq \langle b, e_n \rangle \neq 0$ para todo $n \in \mathbb{N}$, es posible elegir las sucesiones $\{\alpha_n\}_{n=1}^{\infty}$, $\{\beta_n\}_{n=1}^{\infty} \in \ell^2$ de manera tal que, además de satisfacer (5.3), se tenga que $b \in Ran(A)$ y, si x_N denota la solución de mínimos cuadrados de $Ax = b$ en $\mathcal{X}_N$, entonces $\|x_N - A^{-1}b\| \geq f(N)$ para todo $N \in \mathbb{N}$.

Las sucesiones $\{\alpha_n\}_{n=1}^{\infty}$ y $\{\beta_n\}_{n=1}^{\infty}$ serán construidas recursivamente de la siguiente manera:

Paso 1: Elegimos α_1 arbitrario, $\alpha_1 \neq 0$ y $\beta_1 \doteq 0$.

Paso 2: Definimos $\alpha_2 \doteq \frac{1}{2}$,

$$\beta_2 \doteq \begin{cases} \frac{\alpha_1 b_2}{2b_1}, & \text{si } b_2 = \frac{\alpha_2}{2^{\frac{2}{2}+1}} = \frac{1}{8}, \\ \frac{\alpha_1}{b_1}\left(b_2 - \frac{\alpha_2}{2^{\frac{2}{2}+1}}\right) = \frac{\alpha_1}{b_1}\left(b_2 - \frac{1}{8}\right), & \text{en caso contrario.} \end{cases}$$

Es suficiente elegir $\beta_2 \neq 0$ tal que $\frac{b_1\beta_2}{\alpha_1} \in \left(b_2 - \frac{\alpha_2}{2^{\frac{2}{2}}}, b_2\right)$.

Paso 3: Definimos $\alpha_4 \doteq \frac{1}{4}$,

$$\beta_4 \doteq \begin{cases} \frac{\alpha_1 b_4}{2b_1}, & \text{si } b_4 = \frac{\alpha_4}{2^{\frac{4}{2}+1}} = \frac{1}{32}, \\ \frac{\alpha_1}{b_1}\left(b_4 - \frac{\alpha_4}{2^{\frac{4}{2}+1}}\right) = \frac{\alpha_1}{b_1}\left(b_4 - \frac{1}{32}\right), & \text{en caso contrario.} \end{cases}$$

También aquí, es suficiente elegir $\beta_4 \neq 0$ tal que $\frac{b_1\beta_4}{\alpha_1} \in \left(b_4 - \frac{\alpha_4}{2^{\frac{4}{2}}}, b_4\right)$.

Paso 4: Definimos $\alpha_6 \doteq \frac{1}{6}$,

$$\beta_6 \doteq \begin{cases} \frac{\alpha_1 b_6}{2b_1}, & \text{si } b_6 = \frac{\alpha_6}{2^{\frac{6}{2}+1}} = \frac{1}{96}, \\ \frac{\alpha_1}{b_1}\left(b_6 - \frac{\alpha_6}{2^{\frac{6}{2}+1}}\right) = \frac{\alpha_1}{b_1}\left(b_6 - \frac{1}{96}\right), & \text{en caso contrario.} \end{cases}$$

Como antes, es suficiente elegir $\beta_6 \neq 0$ tal que $\frac{b_1\beta_6}{\alpha_1} \in \left(b_6 - \frac{\alpha_6}{2^{\frac{6}{2}}}, b_6\right)$. Por simplicidad, definimos $K \doteq \alpha_1^2\left[1 + b_1^{-2}\left(\frac{5}{4}\|b\|^2 + \sqrt{2}\,\|b\| + 2\right)\right]$.

Paso 5: El término α_3 se elige según sea el signo de b_3, como sigue:

- si $b_3 < 0$, entonces

$$\alpha_3 \doteq \min\left\{\frac{1}{3}; \left|\frac{\alpha_1 b_3 \beta_6}{b_1}\right| \left(b_6 - \frac{\beta_6 b_1}{\alpha_1}\right) [Kf(4)]^{-1}\right\} > 0. \tag{5.5}$$

  Notar que $\alpha_3 > 0$ puesto que $K > 0$ y de acuerdo a la elección de β_6 hecha en el paso 4, se tiene que $b_6 - \frac{\beta_6 b_1}{\alpha_1} > 0$,

- si $b_3 > 0$, entonces se define

$$\alpha_3 \doteq \text{mín}\left\{\frac{1}{3}; 2^{\frac{3}{2}}\, b_3; b_3 \left[\frac{Kf(4)}{\left|\frac{\alpha_1\beta_6}{b_1}\right|\left(b_6 - \frac{\beta_6 b_1}{\alpha_1}\right)} + 2^{-\frac{3}{2}}\right]^{-1}\right\} > 0. \tag{5.6}$$

En (5.5) y (5.6) es suficiente elegir α_3 como cualquier número positivo menor o igual que el mínimo respectivo.

Paso 6: Definimos

$$\beta_3 \doteq \begin{cases} \frac{\alpha_1 b_3}{2b_1}, & \text{si } b_3 = \frac{\alpha_3}{2^{\frac{3}{2}+1}}, \\ \frac{\alpha_1}{b_1}\left(b_3 - \frac{\alpha_3}{2^{\frac{3}{2}+1}}\right), & \text{en caso contrario.} \end{cases}$$

Aquí, de nuevo es suficiente elegir $\beta_3 \neq 0$ tal que $\frac{b_1\beta_3}{\alpha_1} \in \left(b_3 - \frac{\alpha_3}{2^{\frac{3}{2}}}, b_3\right)$.

Se repiten pasos similares a 4, 5 y 6 para construir α_{2j}, β_{2j} y α_{2j-3}, β_{2j-3} para $j \geq 4$. Más precisamente:

Paso 7: Para todo $j \geq 4$, se definen

$$\alpha_{2j} \doteq \frac{1}{2j},$$

$$\beta_{2j} \doteq \begin{cases} \frac{\alpha_1 b_{2j}}{2b_1}, & \text{si } b_{2j} = \frac{\alpha_{2j}}{2^{j+1}}, \\ \frac{\alpha_1}{b_1}\left(b_{2j} - \frac{\alpha_{2j}}{2^{j+1}}\right), & \text{en caso contrario.} \end{cases}$$

Aquí, es suficiente elegir $\beta_{2j} \neq 0$ tal que $\frac{b_1\beta_{2j}}{\alpha_1} \in \left(b_{2j} - \frac{\alpha_{2j}}{2^j}, b_{2j}\right)$.

Paso 8: Para todo $j \geq 4$, el término α_{2j-3} se elige según sea el signo de b_{2j-3} como sigue:

- si $b_{2j-3} < 0$, se define

$$\alpha_{2j-3} \doteq \text{mín}\left\{\frac{1}{2j-3}; \left|\frac{\alpha_1 b_{2j-3}\beta_{2j}}{b_1}\right|\left(b_{2j} - \frac{\beta_{2j}b_1}{\alpha_1}\right)[Kf(2j-2)]^{-1}\right\} > 0. \tag{5.7}$$

Notar que $\alpha_{2j-3} > 0$ pues según la elección de β_{2j} se tiene que $b_{2j} - \frac{\beta_{2j}b_1}{\alpha_1} > 0$.

- Si $b_{2j-3} > 0$, se define

$$\alpha_{2j-3} \doteq \text{mín}\left\{ \frac{1}{2j-3}; 2^{\frac{2j-3}{2}} b_{2j-3}; b_{2j-3} \left[\frac{Kf(2j-2)}{\left|\frac{\alpha_1 \beta_{2j}}{b_1}\right| \left(b_{2j} - \frac{\beta_{2j} b_1}{\alpha_1}\right)} + 2^{-\frac{2j-3}{2}} \right]^{-1} \right\} > 0. \tag{5.8}$$

En (5.7) y (5.8) basta cualquier elección positiva de α_{2j-3} menor o igual que el mínimo.

Finalmente, se define

$$\beta_{2j-3} \doteq \begin{cases} \frac{\alpha_1 b_{2j-3}}{2b_1}, & \text{si } b_{2j-3} = \frac{\alpha_{2j-3}}{2^{\frac{2j-3}{2}+1}}, \\ \frac{\alpha_1}{b_1}\left(b_{2j-3} - \frac{\alpha_{2j-3}}{2^{\frac{2j-3}{2}+1}}\right), & \text{en caso contrario.} \end{cases} \tag{5.9}$$

Aquí también es suficiente elegir $\beta_{2j-3} \neq 0$ tal que $\frac{b_1 \beta_{2j-3}}{\alpha_1} \in \left(b_{2j-3} - \frac{\alpha_{2j-3}}{2^{\frac{2j-3}{2}}}, b_{2j-3}\right)$.

En la Figura 5.1 se representa el orden de recursión usado para la construcción de las sucesiones $\{\alpha_n\}_{n=1}^{\infty}$ y $\{\beta_n\}_{n=1}^{\infty}$. Es fácil probar que estas sucesiones satisfacen las condiciones (5.3) y que además están en ℓ^2. En efecto,

$$\sum_{n=1}^{\infty} \alpha_n^2 \leq \sum_{n=1}^{\infty} \frac{1}{n^2} < \infty,$$

pues $0 < \alpha_n \leq \frac{1}{n}$ para todo $n \in \mathbb{N}$ debido a la elección de α_n. Por otro lado, teniendo en cuenta la elección de la sucesión $\{\beta_n\}_{n=1}^{\infty}$ se sigue que

$$\begin{aligned}
\sum_{n=1}^{\infty} \beta_n^2 &\leq \sum_{n=1}^{\infty} \left(\frac{\alpha_1 b_n}{2b_1}\right)^2 + \sum_{n=1}^{\infty} \left(\frac{\alpha_1}{b_1}\right)^2 \left(b_n - \frac{\alpha_n}{2^{\frac{n}{2}+1}}\right)^2 \\
&= \frac{\alpha_1^2}{b_1^2}\left[\frac{1}{4}\sum_{n=1}^{\infty} b_n^2 + \sum_{n=1}^{\infty}\left(b_n^2 - \frac{b_n \alpha_n}{2^{\frac{n}{2}}} + \frac{\alpha_n^2}{2^{n+2}}\right)\right] \\
&\leq \frac{\alpha_1^2}{b_1^2}\left[\frac{5}{4}\|b\|^2 + \left(\sum_{n=1}^{\infty} b_n^2\right)^{\frac{1}{2}} \left(\sum_{n=1}^{\infty} \frac{\alpha_n^2}{2^n}\right)^{\frac{1}{2}} + \sum_{n=1}^{\infty} \alpha_n^2\right] \quad \begin{array}{l}\text{(por Cauchy-Schwarz y} \\ \text{porque } 2^{n+2} > 1 \,\forall\, n \in \mathbb{N})\end{array} \\
&\leq \frac{\alpha_1^2}{b_1^2}\left[\frac{5}{4}\|b\|^2 + \|b\| \left(\sum_{n=1}^{\infty} \frac{1}{n^2}\right)^{\frac{1}{2}} + \sum_{n=1}^{\infty} \frac{1}{n^2}\right] \quad \left(\text{pues } \alpha_n \leq \frac{1}{n} \text{ y } 2^n > 1 \,\forall\, n \in \mathbb{N}\right) \\
&\leq \frac{\alpha_1^2}{b_1^2}\left(\frac{5}{4}\|b\|^2 + \sqrt{2}\,\|b\| + 2\right) \qquad \left(\text{pues } \sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6} < 2\right) \\
&= K - \alpha_1^2 < \infty.
\end{aligned} \tag{5.10}$$

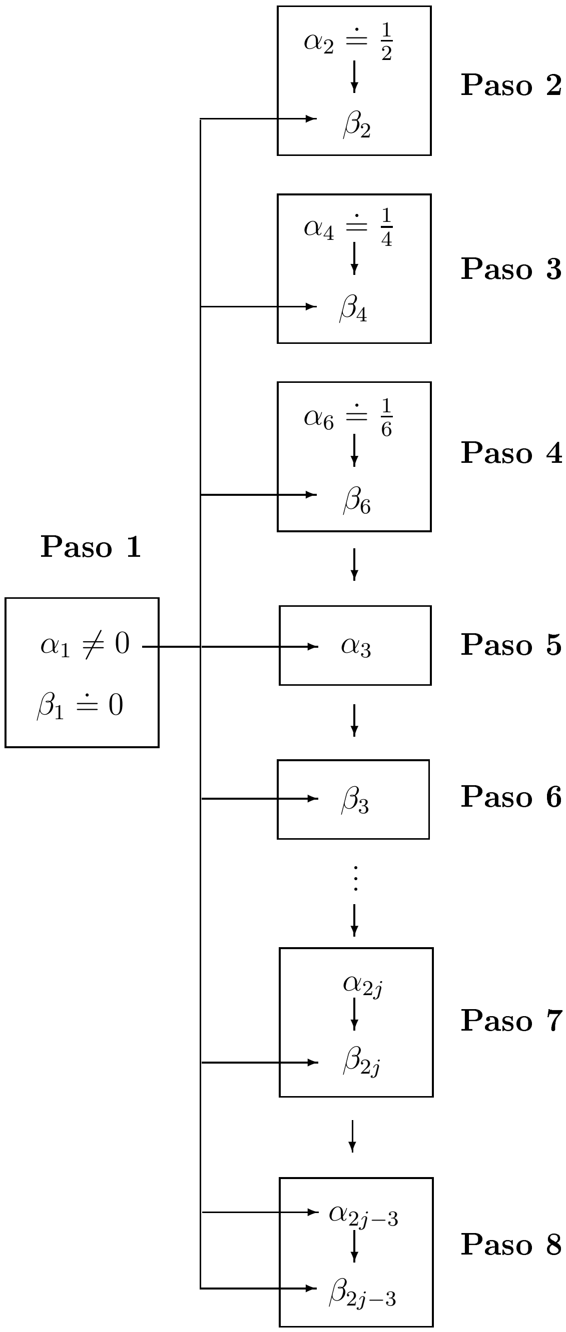

Figura 5.1: Orden de recursión usado para la construcción de las sucesiones $\{\alpha_n\}_{n=1}^{\infty}$ y $\{\beta_n\}_{n=1}^{\infty}$, en el Teorema 5.4.

Veamos ahora que efectivamente $b \in Ran(A)$, es decir, que existe $x^* = \sum_{n=1}^{\infty} \xi_n^* e_n \in \mathcal{X}$ tal que $Ax^* = b$. Para ello, probaremos que el sistema de ecuaciones que se deduce de la igualdad

$$\sum_{n=1}^{\infty}(\alpha_n \xi_n^* + \beta_n \xi_1^*) e_n = \sum_{n=1}^{\infty} b_n e_n,$$

tiene una única solución $\{\xi_n^*\} \in \ell^2$. Igualando primero los coeficientes de e_1 tenemos que $\alpha_1 \xi_1^* + \beta_1 \xi_1^* = b_1$. Como $\beta_1 = 0$, se sigue que $\xi_1^* = \frac{b_1}{\alpha_1}$. Similarmente, igualando los coeficientes de e_n para $n \geq 2$ se sigue que $\alpha_n \xi_n^* + \beta_n \xi_1^* = b_n$, lo cual implica que

$$\xi_n^* = \left(b_n - \frac{b_1 \beta_n}{\alpha_1} \right) \alpha_n^{-1}. \tag{5.11}$$

Luego,

$$\begin{aligned}
\sum_{n=1}^{\infty} \xi_n^{*2} &= \frac{b_1^2}{\alpha_1^2} + \sum_{n=2}^{\infty} \left(b_n - \frac{b_1 \beta_n}{\alpha_1} \right)^2 \alpha_n^{-2} \\
&= \frac{b_1^2}{\alpha_1^2} + \sum_{\substack{n=2 \\ b_n = \frac{\alpha_n}{2^{\frac{n}{2}+1}}}}^{\infty} \left(b_n - \frac{b_1 \beta_n}{\alpha_1} \right)^2 \alpha_n^{-2} + \sum_{\substack{n=2 \\ b_n \neq \frac{\alpha_n}{2^{\frac{n}{2}+1}}}}^{\infty} \left(b_n - \frac{b_1 \beta_n}{\alpha_1} \right)^2 \alpha_n^{-2} \\
&= \frac{b_1^2}{\alpha_1^2} + \sum_{\substack{n=2 \\ b_n = \frac{\alpha_n}{2^{\frac{n}{2}+1}}}}^{\infty} \left(b_n - \frac{b_n}{2} \right)^2 \alpha_n^{-2} + \sum_{\substack{n=2 \\ b_n \neq \frac{\alpha_n}{2^{\frac{n}{2}+1}}}}^{\infty} \left(\frac{\alpha_n}{2^{\frac{n}{2}+1}} \right)^2 \alpha_n^{-2} \quad \text{(por nuestra elección de } \beta_n\text{)} \\
&\leq \frac{b_1^2}{\alpha_1^2} + \sum_{n=2}^{\infty} \frac{1}{2^{n+4}} + \sum_{n=2}^{\infty} \frac{1}{2^{n+2}} < \infty.
\end{aligned}$$

Así, $\{\xi_n^*\} \in \ell^2$, $x^* = \sum_{n=1}^{\infty} \xi_n^* e_n$ y $b = Ax^* \in Ran(A)$.

Sea ahora x_N la solución de mínimos cuadrados de $Ax = b$ en $\mathcal{X}_N$ y escribimos $x_N = \sum_{n=1}^{N} \xi_n e_n$. Probaremos que $\|x_N - x^*\| \geq f(N)$ para todo $N \in \mathbb{N}$. Definiendo el funcional $\mathcal{J}(x_N) \doteq \frac{1}{2} \|Ax_N - Ax^*\|^2$ e igualando a cero cada una de sus derivadas parciales con respecto a ξ_i, $i = 1, \ldots, N$ (por razones de brevedad omitimos los detalles aquí) se sigue inmediatamente que

$$\xi_1 = \xi_1^* + \frac{\sum_{n=N+1}^{\infty} \alpha_n \beta_n \xi_n^*}{\alpha_1^2 + \sum_{n=N+1}^{\infty} \beta_n^2},$$

$$\xi_n = \xi_n^* - \frac{\beta_n}{\alpha_n}(\xi_1 - \xi_1^*), \qquad \text{para } 2 \leq n \leq N.$$

Por lo tanto,

$$\begin{aligned} \|x_N - x^*\|^2 &= \frac{\left(\sum_{n=N+1}^{\infty} \alpha_n \beta_n \xi_n^*\right)^2}{\left(\alpha_1^2 + \sum_{n=N+1}^{\infty} \beta_n^2\right)^2} \left(1 + \sum_{n=2}^{N} \frac{\beta_n^2}{\alpha_n^2}\right) + \sum_{n=N+1}^{\infty} \xi_n^{*2} \\ &\geq \frac{\left(\sum_{n=N+1}^{\infty} \alpha_n \beta_n \xi_n^*\right)^2}{\left(\alpha_1^2 + \sum_{n=N+1}^{\infty} \beta_n^2\right)^2} \sum_{n=2}^{N} \frac{\beta_n^2}{\alpha_n^2}. \end{aligned} \tag{5.12}$$

Luego,

(a) Para N par,

$$\begin{aligned} \|x_N - x^*\|^2 &\geq \frac{\left(\alpha_{N+2}\beta_{N+2}\xi_{N+2}^*\right)^2}{\left(\alpha_1^2 + \sum_{n=N+1}^{\infty} \beta_n^2\right)^2} \frac{\beta_{N-1}^2}{\alpha_{N-1}^2} \\ &= \frac{\beta_{N+2}^2 \left(b_{N+2} - \frac{b_1 \beta_{N+2}}{\alpha_1}\right)^2}{\left(\alpha_1^2 + \sum_{n=N+1}^{\infty} \beta_n^2\right)^2} \frac{\beta_{N-1}^2}{\alpha_{N-1}^2} \qquad \text{(reemplazando } \xi_{N+2}^* \text{ por (5.11)).} \end{aligned} \tag{5.13}$$

(a.1) Si $b_{N-1} < 0$, entonces por la elección de β_n se tiene que $\frac{b_1 \beta_{N-1}}{\alpha_1} < b_{N-1} < 0$ y por lo tanto $\beta_{N-1}^2 > \left(\frac{\alpha_1 b_{N-1}}{b_1}\right)^2$, lo cual junto a la elección de α_{N-1} (ver (5.7) en Paso 8) implica que

$$\begin{aligned} \|x_N - x^*\|^2 &\geq \frac{[f(N)]^2 K^2}{\left(\alpha_1^2 + \sum_{n=N+1}^{\infty} \beta_n^2\right)^2} \\ &\geq \frac{[f(N)]^2 \left(\alpha_1^2 + \sum_{n=1}^{\infty} \beta_n^2\right)^2}{\left(\alpha_1^2 + \sum_{n=N+1}^{\infty} \beta_n^2\right)^2} \qquad \text{(por (5.10)).} \\ &\geq [f(N)]^2, \end{aligned}$$

(a.2) Si $b_{N-1} > 0$, entonces como α_{N-1} se eligió tal que $\alpha_{N-1} \leq 2^{\frac{N-1}{2}} b_{N-1}$ (ver (5.8) en Paso 8) y recordando nuestra elección de β_{N-1} (ver (5.9)), resulta que $0 \leq b_{N-1} - \frac{\alpha_{N-1}}{2^{\frac{N-1}{2}}} < \frac{b_1 \beta_{N-1}}{\alpha_1}$. Luego, $\beta_{N-1}^2 > \left(\frac{\alpha_1}{b_1}\right)^2 \left(b_{N-1} - \frac{\alpha_{N-1}}{2^{\frac{N-1}{2}}}\right)^2$. Sustituyendo esto en (5.13) se tiene finalmente que

$$\|x_N - x^*\|^2 \geq \frac{\left[\beta_{N+2}\left(b_{N+2} - \frac{b_1 \beta_{N+2}}{\alpha_1}\right)\right]^2}{\left(\alpha_1^2 + \sum_{n=N+1}^{\infty} \beta_n^2\right)^2} \frac{\alpha_1^2}{b_1^2} \left(b_{N-1} - \frac{\alpha_{N-1}}{2^{\frac{N-1}{2}}}\right)^2 \frac{1}{\alpha_{N-1}^2}$$

$$
\begin{aligned}
&= \frac{\left[\beta_{N+2}\left(b_{N+2}-\frac{b_1\beta_{N+2}}{\alpha_1}\right)\right]^2}{\left(\alpha_1^2+\sum_{n=N+1}^{\infty}\beta_n^2\right)^2}\frac{\alpha_1^2}{b_1^2}\left(\frac{b_{N-1}}{\alpha_{N-1}}-\frac{1}{2^{\frac{N-1}{2}}}\right)^2 \\
&\geq \frac{[f(N)]^2 K^2}{\left(\alpha_1^2+\sum_{n=N+1}^{\infty}\beta_n^2\right)^2} \qquad \text{(por (5.8))} \\
&\geq \frac{[f(N)]^2}{\left(\alpha_1^2+\sum_{n=N+1}^{\infty}\beta_n^2\right)^2}\left(\alpha_1^2+\sum_{n=1}^{\infty}\beta_n^2\right)^2 \qquad \text{(por (5.10))} \\
&\geq [f(N)]^2 .
\end{aligned}
$$

(b) Para N impar, se sigue de (5.12) que

$$\|x_N - x^*\|^2 \geq \frac{\left(\alpha_{N+3}\beta_{N+3}\xi_{N+3}^*\right)^2}{\left(\alpha_1^2+\sum_{n=N+1}^{\infty}\beta_n^2\right)^2}\frac{\beta_N^2}{\alpha_N^2}.$$

Siguiendo los mismos pasos que en el caso previo (N par) resulta que

$$\|x_N - x^*\|^2 \geq [f(N+1)]^2 \geq [f(N)]^2 ,$$

pues f es una función creciente. Esto completa la demostración del teorema. ■

Luego, en el Corolario 5.7, veremos que bajo ciertos supuestos adicionales generales sobre b, el operador A del Teorema previo puede construirse de tal manera que sea además autoadjunto.

Para nuestro próximo resultado necesitaremos previamente el siguiente Lema.

Lema 5.5. *Sean $\mathcal{X}$ un espacio de Hilbert, N un subespacio cerrado de $\mathcal{X}$, $A : \mathcal{X} \to \mathcal{X}$ un operador lineal y acotado, M un subespacio cerrado de $\mathcal{X}$ tal que M y $M^\perp$ son invariantes por A y $b \in M$. Si x^* es la solución de mínimos cuadrados de $Ax = b$ en N, entonces la solución de mínimos cuadrados de $Ax = b$ en $M \cap N$ es $P_M x^*$, donde P_M es la proyección ortogonal de $\mathcal{X}$ sobre M.*

Demostración. Supongamos que $P_M x^*$ no es la solución de mínimos cuadrados de $Ax = b$ en $M \cap N$. Entonces existe $z \in M \cap N$, $z \neq P_M x^*$, tal que

$$\|Az - b\| < \|AP_M x^* - b\| . \tag{5.14}$$

Sea $\widetilde{x^*} \doteq z + P_{M^\perp}x^* \in N$. Como $z \neq P_M x^*$, resulta que $\widetilde{x^*} \neq x^*$. Entonces, puesto que $b \in M$, $z \in M$ y M, $M^\perp$ son ambos invariantes por A, se tiene que

$$\begin{aligned}
\left\|A\widetilde{x^*} - b\right\|^2 &= \|Az + AP_{M^\perp}x^* - b\|^2 \\
&= \|Az - b\|^2 + \|AP_{M^\perp}x^*\|^2 \\
&< \|AP_M x^* - b\|^2 + \|AP_{M^\perp}x^*\|^2 \quad \text{(por (5.14))} \\
&= \|Ax^* - b\|^2,
\end{aligned}$$

lo cual es absurdo pues $\widetilde{x^*} \in N$ y x^* es la solución de mínimos cuadrados de $Ax = b$ en N. Por lo tanto, $P_M x^*$ es la solución de mínimos cuadrados de $Ax = b$ en $M \cap N$. ∎

En aplicaciones concretas del método de mínimos cuadrados, por lo general la base $B = \{e_n\}_{n=1}^\infty$ de $\mathcal{X}$ está dada y $\mathcal{X}_N$ se elige como $\text{span}\{e_1, \ldots, e_N\}$. En el siguiente Corolario mostramos que en este caso, para cualquier $b \in \mathcal{X}$ que sea débilmente no degenerado con respecto a esa base y para cualquier función $f : \mathbb{N} \to \mathbb{R}^+$ creciente, también es posible construir un operador $A = A(b, f)$ tal que las soluciones de mínimos cuadrados de $Ax = b$ en $\mathcal{X}_N$ diverjan de la solución exacta con velocidad mayor o igual que f.

Corolario 5.6. *Sean $\mathcal{X}$ un espacio de Hilbert separable de dimensión infinita, $f : \mathbb{N} \to \mathbb{R}^+$ una función creciente arbitraria, $B \doteq \{e_n\}_{n=1}^\infty$ una base ortonormal de $\mathcal{X}$, $b \in \mathcal{X}$ débilmente no degenerado con respecto a B y $\mathcal{X}_N \doteq \text{span}\{e_1, ..., e_N\}$. Entonces, existe un operador lineal $A = A(b, f) : \mathcal{X} \to \mathcal{X}$, compacto, inyectivo y de rango denso en $\mathcal{X}$ tal que $b = A\hat{x}$ para algún $\hat{x} \in \mathcal{X}$ y, si x_N es la solución de mínimos cuadrados de $Ax = b$ en $\mathcal{X}_N$, entonces $\|x_N - \hat{x}\| \geqslant f(N)$ para todo $N \in \mathbb{N}$.*

Demostración. Definimos

$$\Delta \doteq \{n \in \mathbb{N} : \langle b, e_n \rangle \neq 0\}, \quad \Gamma \doteq \mathbb{N} \setminus \Delta,$$

$$B^\Delta \doteq \{e_{n_j} : n_j \in \Delta\}, \quad B^\Gamma \doteq \{e_{m_j} : m_j \in \Gamma\} \tag{5.15}$$

y sea $\mathcal{X}_b \doteq \overline{\operatorname{span} B^\Delta}$. Claramente,

$$\overline{\operatorname{span} B^\Gamma} = \mathcal{X}_b^\perp \tag{5.16}$$

y b es fuertemente no degenerado con respecto a la base B^Δ de $\mathcal{X}_b$. Como $\mathcal{X}_b$ es un espacio de Hilbert separable de dimensión infinita, el Teorema 5.4 implica que existe un operador lineal $A^\Delta : \mathcal{X}_b \to \mathcal{X}_b$ compacto, inyectivo y de rango denso en $\mathcal{X}_b$ tal que $b = A^\Delta \hat{x}$ para algún $\hat{x} \in \mathcal{X}_b$ y si x_M^Δ es la solución de mínimos cuadrados de $A^\Delta x = b$ en

$$\mathcal{X}_M^\Delta \doteq \operatorname{span}\{e_{n_j}\}_{j=1}^M \subset \mathcal{X}_b, \tag{5.17}$$

entonces

$$\left\|x_M^\Delta - \hat{x}\right\| \geq \tilde{f}(M)\ \forall\, M \in \mathbb{N}, \tag{5.18}$$

donde $\tilde{f}(j) \doteq f(n_{j+1})$.

A continuación construiremos el operador A como una extensión apropiada de A^Δ a todo el espacio $\mathcal{X}$. La forma en que se realiza esta extensión dependerá de la cardinalidad del conjunto Γ definido en (5.15).

Caso I. El conjunto Γ no es finito. En este caso $\mathcal{X}_b^\perp$ es un espacio de Hilbert separable de dimensión infinita. Sea $\{\gamma_n\}_{n=1}^\infty$ una sucesión en ℓ^2 tal que $\gamma_n \neq 0\ \forall\, n \in \mathbb{N}$ y definimos $\tilde{b} \doteq \sum_{n\in\Gamma} \gamma_n e_n$. Claramente, $\tilde{b}$ es fuertemente no degenerado con respecto a la base B^Γ de $\mathcal{X}_b^\perp$. En virtud del Teorema 5.4, existe un operador lineal $A^\Gamma : \mathcal{X}_b^\perp \to \mathcal{X}_b^\perp$ compacto, inyectivo y de rango denso en $\mathcal{X}_b^\perp$ tal que $\tilde{b} = A^\Gamma \tilde{x}$ para algún $\tilde{x} \in \mathcal{X}_b^\perp$. Entonces definimos A como sigue:

$$Ax \doteq A^\Delta x_1 + A^\Gamma x_2, \tag{5.19}$$

donde $x = x_1 + x_2$ con $x_1 \in \mathcal{X}_b$ y $x_2 \in \mathcal{X}_b^\perp$.

Caso II. El conjunto Γ es finito. En este caso, simplemente definimos A como

$$Ax \doteq A^\Delta x_1 + x_2, \tag{5.20}$$

donde nuevamente $x = x_1 + x_2$ con $x_1 \in \mathcal{X}_b$ y $x_2 \in \mathcal{X}_b^\perp$.

Es claro que en ambos casos el operador A así definido es lineal. Veamos ahora que es inyectivo, compacto y que tiene rango denso en $\mathcal{X}$.

En primer lugar consideramos el caso I (ver (5.19)). Si $Ax = 0$, entonces $A^{\Delta}x_1 = -A^{\Gamma}x_2$. Como $A^{\Delta}x_1 \in \mathcal{X}_b$ y $-A^{\Gamma}x_2 \in \mathcal{X}_b^{\perp}$ se tiene que $A^{\Delta}x_1 = -A^{\Gamma}x_2 = 0$, y por lo tanto $x_1 = x_2 = 0$ pues A^{Δ} y A^{Γ} son ambos inyectivos. Luego, $x = x_1 + x_2 = 0$ y por lo tanto A es inyectivo.

Probemos ahora que $Ran(A)$ es denso en $\mathcal{X}$. Por definición de A y debido a que $Ran(A^{\Delta}) \subset \mathcal{X}_b$ y $Ran(A^{\Gamma}) \subset \mathcal{X}_b^{\perp}$, resulta que $Ran(A^{\Delta})$ y $Ran(A^{\Gamma})$ son dos subespacios de $\mathcal{X}$ ortogonales entre sí y por lo tanto $Ran(A) = Ran(A^{\Delta}) \oplus Ran(A^{\Gamma})$. Luego,

$$\begin{aligned}
\overline{Ran(A)} &= \overline{Ran(A^{\Delta})} \oplus \overline{Ran(A^{\Gamma})} \\
&= \left(\mathcal{X}_b \cap \overline{Ran(A^{\Delta})}\right) \oplus \left(\mathcal{X}_b^{\perp} \cap \overline{Ran(A^{\Gamma})}\right) \qquad \text{(pues } Ran(A^{\Delta}) \subset \mathcal{X}_b \text{ y } Ran(A^{\Gamma}) \subset \mathcal{X}_b^{\perp}\text{)} \\
&= \overline{Ran(A^{\Delta})}^{\mathcal{X}_b} \oplus \overline{Ran(A^{\Gamma})}^{\mathcal{X}_b^{\perp}} \\
&= \mathcal{X}_b \oplus \mathcal{X}_b^{\perp} = \mathcal{X},
\end{aligned}$$

donde la penúltima igualdad se sigue del hecho que $Ran(A^{\Delta})$ y $Ran(A^{\Gamma})$ son densos en $\mathcal{X}_b$ y $\mathcal{X}_b^{\perp}$, respectivamente. Por lo tanto, $Ran(A)$ es denso en $\mathcal{X}$.

Para probar que A es compacto, notemos que A puede expresarse en la forma

$$A = A^{\Delta}P_{\mathcal{X}_b} + A^{\Gamma}P_{\mathcal{X}_b^{\perp}},$$

donde $P_{\mathcal{X}_b}$ y $P_{\mathcal{X}_b^{\perp}}$ son las proyecciones ortogonales de $\mathcal{X}$ sobre los subespacios $\mathcal{X}_b$ y $\mathcal{X}_b^{\perp}$, respectivamente. Como A^{Δ} y A^{Γ} son ambos operadores compactos y $P_{\mathcal{X}_b}$ y $P_{\mathcal{X}_b^{\perp}}$ son ambos acotados, resulta que A es compacto por ser suma de operadores compactos.

Consideremos ahora el caso II, es decir, cuando el operador A está definido como en (5.20). El operador A es inyectivo. En efecto, si $Ax = 0$ entonces $A^{\Delta}x_1 = -x_2$ y como $A^{\Delta}x_1 \in \mathcal{X}_b$ y $-x_2 \in \mathcal{X}_b^{\perp}$ se tiene que $A^{\Delta}x_1 = -x_2 = 0$. Debido al hecho que A^{Δ} es inyectivo resulta que $x_1 = 0$ y por lo tanto $x = x_1 + x_2 = 0$.

Veamos ahora que $Ran(A)$ es denso en $\mathcal{X}$. Por definición de A y puesto que $Ran(A^{\Delta}) \subset \mathcal{X}_b$, resulta que $Ran(A^{\Delta})$ y $\mathcal{X}_b^{\perp}$ son subespacios de $\mathcal{X}$ ortogonales entre sí. Entonces,

$$\begin{aligned}
\overline{Ran(A)} &= \overline{Ran(A^{\Delta})} \oplus \overline{\mathcal{X}_b^{\perp}} \\
&= \left(\mathcal{X}_b \cap \overline{Ran(A^{\Delta})}\right) \oplus \mathcal{X}_b^{\perp} \quad (\text{pues } Ran(A^{\Delta}) \subset \mathcal{X}_b \text{ y } \mathcal{X}_b^{\perp} \text{ es cerrado}) \\
&= \overline{Ran(A^{\Delta})}^{\mathcal{X}_b} \oplus \mathcal{X}_b^{\perp} \\
&= \mathcal{X}_b \oplus \mathcal{X}_b^{\perp} \quad (\text{pues } Ran(A^{\Delta}) \text{ es denso en } \mathcal{X}_b) \\
&= \mathcal{X}.
\end{aligned}$$

Veamos ahora que también en este caso A resulta compacto. Para ello escribimos $A = A^{\Delta} P_{\mathcal{X}_b} + P_{\mathcal{X}_b^{\perp}}$. Como en este caso Γ es finito, se sigue de (5.15) y (5.16) que $P_{\mathcal{X}_b^{\perp}}$ es un operador de rango finito y, por lo tanto, compacto. Luego, A resulta compacto por ser suma de dos operadores compactos.

Ahora probemos que en ambos casos (I y II) $b \in Ran(A)$. Como $\hat{x} \in \mathcal{X}_b$ y $A^{\Delta}\hat{x} = b$, se tiene de (5.19) y (5.20) que $A\hat{x} = A^{\Delta}\hat{x} + A^{\Gamma}0 = b$ en el caso I, y $A\hat{x} = A^{\Delta}\hat{x} + 0 = b$ en el caso II.

Finalmente, sea x_N la solución de mínimos cuadrados de $Ax = b$ en $\mathcal{X}_N \doteq \text{span}\{e_1, ..., e_N\}$, $x_N = \sum_{j=1}^{N} \xi_j\, e_j$. Probaremos que $\|x_N - \hat{x}\| \geq f(N)$ para todo $N \in \mathbb{N}$:

$$\begin{aligned}
\|x_N - \hat{x}\|^2 &= \left\| \sum_{j=1}^{N} \xi_j\, e_j - \hat{x} \right\|^2 \\
&= \left\| \sum_{j=1,\, j\in\Delta}^{N} \xi_j\, e_j + \sum_{j=1,\, j\in\Gamma}^{N} \xi_j\, e_j - \hat{x} \right\|^2 \\
&= \left\| \sum_{j=1,\, j\in\Delta}^{N} \xi_j\, e_j - \hat{x} \right\|^2 + \left\| \sum_{j=1,\, j\in\Gamma}^{N} \xi_j\, e_j \right\|^2 \\
&\geq \left\| \sum_{j=1,\, j\in\Delta}^{N} \xi_j\, e_j - \hat{x} \right\|^2, \qquad (5.21)
\end{aligned}$$

donde en la última igualdad se usó que $\sum_{j=1,\, j\in\Delta}^{N} \xi_j\, e_j - \hat{x} \in \mathcal{X}_b$ y $\sum_{j=1,\, j\in\Gamma}^{N} \xi_j\, e_j \in \mathcal{X}_b^{\perp}$.

Para cada $N \in \mathbb{N}$ definimos $k_N \doteq \max\{j : n_j \leq N\}$ ó, equivalentemente, $k_N \doteq \#\{e_j : j \leq N, e_j \in B^{\Delta}\} = \#\left(B_N \cap B^{\Delta}\right)$, donde $B_N \doteq \{e_1, \ldots, e_N\}$. Se sigue de (5.17) que $\mathcal{X}_{k_N}^{\Delta} = \operatorname{span}\{e_{n_j}\}_{j=1}^{k_N}$. Como $\mathcal{X}_b$ es un subespacio cerrado de $\mathcal{X}$ tal que $\mathcal{X}_b$ y $\mathcal{X}_b^{\perp}$ son invariantes por A y $b \in \mathcal{X}_b$, se sigue del Lema 5.5 que la solución de mínimos cuadrados de $Ax = b$ en $\mathcal{X}_b \cap \mathcal{X}_N = \mathcal{X}_{k_N}^{\Delta}$ es

$$P_{\mathcal{X}_b} x_N = P_{\mathcal{X}_{k_N}^{\Delta}} x_N = \sum_{j=1,\, j\in\Delta}^{N} \xi_j\, e_j,$$

donde $x_N = \sum_{j=1}^{N} \xi_j\, e_j$ es la solución de mínimos cuadrados de $Ax = b$ en $\mathcal{X}_N$.

Ahora bien, como $A|_{\mathcal{X}_b} = A^{\Delta}$ y $\mathcal{X}_{k_N}^{\Delta} \subset \mathcal{X}_b$, se tiene que $A\eta = A^{\Delta}\eta \;\forall\, \eta \in \mathcal{X}_{k_N}^{\Delta}$ y por lo tanto $\sum_{\substack{j=1\\ j\in\Delta}}^{N} \xi_j\, e_j$ es también la solución de mínimos cuadrados de $A^{\Delta}x = b$ en $\mathcal{X}_{k_N}^{\Delta}$, es decir, $x_{k_N}^{\Delta} = \sum_{\substack{j=1\\ j\in\Delta}}^{N} \xi_j\, e_j$ (ver (5.17)). Luego, sustituyendo lo anterior en (5.21) y usando (5.18) resulta que

$$\|x_N - \hat{x}\|^2 \geq \left\|x_{k_N}^{\Delta} - \hat{x}\right\|^2 \geq [\tilde{f}(k_N)]^2 = [f(n_{k_N+1})]^2. \tag{5.22}$$

Ahora, por definición de k_N, si $j_0 \in \mathbb{N}$ y $n_{j_0} \leq N$ entonces $k_N \geq j_0$. Por lo tanto, la desigualdad $n_{k_N+1} \leq N$ implicaría que $k_N \geq k_N + 1$, lo cual es una contradicción. Por lo tanto, $n_{k_N+1} > N$ ó, equivalentemente, $n_{k_N+1} \geq N + 1$.

Como f es creciente y positiva resulta que

$$[f(n_{k_N+1})]^2 \geq [f(N+1)]^2 \geq [f(N)]^2. \tag{5.23}$$

De (5.22) y (5.23) concluímos finalmente que $\|x_N - \hat{x}\| \geq f(N) \;\forall\, N \in \mathbb{N}$. ∎

A continuación probaremos que tanto en el Teorema 5.4 como en el Corolario 5.6, el operador A se puede construir de tal manera que sea además autoadjunto.

Corolario 5.7. *Sean $\mathcal{X}$ un espacio de Hilbert separable de dimensión infinita, $B \doteq \{e_n\}_{n=1}^{\infty}$ una base ortonormal de $\mathcal{X}$, $b \in \mathcal{X}$ no degenerado con respecto a B, $\mathcal{X}_N \doteq$*

span$\{e_1, \ldots, e_N\}$ *y* $f : \mathbb{N} \to \mathbb{R}^+$ *una función creciente arbitraria. Entonces, existen un operador lineal* $A = A(b, f) : \mathcal{X} \to \mathcal{X}$ *compacto, autoadjunto, inyectivo, de rango denso en* $\mathcal{X}$ *y un operador unitario* $V : \mathcal{X} \to \mathcal{X}$ *tal que* $b^* \doteq Vb \in Ran(A)$ *y, si* x_N *la solución de mínimos cuadrados de* $Ax = b^*$ *en* $\mathcal{X}_N$*, entonces* $\|x_N - A^{-1}b^*\| \geq f(N)$ *para todo* $N \in \mathbb{N}$.

Demostración. Puesto que $b \in \mathcal{X}$ es no degenerado con respecto a B, se sigue del Corolario 5.6 que existe un operador lineal $\tilde{A} = \tilde{A}(b, f)$ compacto, inyectivo, de rango denso en $\mathcal{X}$ tal que $b \in Ran(\tilde{A})$ y, si x_N denota la solución de mínimos cuadrados de $\tilde{A}x = b$ en $\mathcal{X}_N$, entonces $\left\|x_N - \tilde{A}^{-1}b\right\| \geq f(N)$ para todo $N \in \mathbb{N}$. Sea $\{(\sigma_n; u_n, v_n)\}$ el sistema singular asociado al operador compacto $\tilde{A}$ y $V : \mathcal{X} \to \mathcal{X}$ el operador unitario definido por $Vv_n = u_n$. Entonces, es fácil probar que el operador lineal $A \doteq V\tilde{A}$, además de ser compacto, inyectivo y de rango denso en $\mathcal{X}$, es autoadjunto. Esto se sigue inmediatamente del hecho que $V\tilde{A}u_n = \tilde{A}^*V^*u_n$ para todo $n \in \mathbb{N}$.

Como $b \in Ran(\tilde{A})$, se tiene que $b^* \doteq Vb \in Ran(A)$. Ahora, puesto que $\|Ax - b^*\| = \left\|\tilde{A}x - b\right\|$, resulta que x_N también es la solución de minimos cuadrados de $Ax = b^*$ en $\mathcal{X}_N$. Por lo tanto, $f(N) \leq \left\|x_N - \tilde{A}^{-1}b\right\| = \|x_N - A^{-1}b^*\|$ para todo $N \in \mathbb{N}$. ■

Puede pensarse que el resultado del Corolario 5.6 tiene poca relevancia desde el punto de vista práctico pues podría suceder que en un entorno de b, el único elemento $\eta \in Ran(A)$ para el cual se tiene que las soluciones de mínimos cuadrados de $Ax = \eta$ en $\mathcal{X}_N$ satisfacen que $\|x_N - A^{-1}\eta\| \geqslant f(N)$ para todo $N \in \mathbb{N}$ sea, precisamente, $\eta = b$. Sin embargo, en el siguiente resultado probaremos que esto no es así. En efecto, veremos que en todo entorno reducido de radio $\delta > 0$ con centro en b, existe una sucesión $\{b_k^\delta\}_{k=1}^\infty \subset Ran(A)$ tal que $b_k^\delta \to b$ cuando $k \to \infty$ y $\left\|x_N^k - A^{-1}b_k^\delta\right\| \geqslant f(N)$ para todo $N \in \mathbb{N}$, donde x_N^k es la solución de mínimos cuadrados de $Ax = b_k^\delta$ en $\mathcal{X}_N$. Es decir, para cada elemento de la sucesión $\{b_k^\delta\}_{k=1}^\infty$ (i.e. para cada k fijo), las soluciones de mínimos cuadrados aproximantes $\{x_N^k\}_{N=1}^\infty$, también se alejan de la solución exacta con velocidad arbitrariamente grande, de la misma forma con que lo hacen las soluciones de mínimos cuadrados obtenidas con dato igual a b.

Corolario 5.8. *Sean $\mathcal{X}$ un espacio de Hilbert separable de dimensión infinita, $B \doteq \{e_n\}_{n=1}^{\infty}$ una base ortonormal de $\mathcal{X}$, $b \in \mathcal{X}$ no degenerado con respecto a B, $\mathcal{X}_N \doteq \operatorname{span}\{e_1, ..., e_N\}$, $f : \mathbb{N} \to \mathbb{R}^+$ una función creciente arbitraria y $A = A(b, f)$ el operador cuya existencia se probó en el Corolario 5.6. Entonces, en todo entorno reducido de radio $\delta > 0$ y centro en b, existe una sucesión $\{b_k^\delta\}_{k=1}^{\infty} \subset Ran(A)$ tal que $b_k^\delta \to b$ cuando $k \to \infty$ y para cada $k \in \mathbb{N}$, la solución de mínimos cuadrados x_N^k de $Ax = b_k^\delta$ en $\mathcal{X}_N$ verifica que $\left\|x_N^k - A^{-1}b_k^\delta\right\| \geqslant f(N)$ para todo $N \in \mathbb{N}$.*

Demostración. Sea $\delta > 0$ dado, definimos $b_k^\delta \doteq \left(1 + \frac{\delta}{2k\|b\|}\right) b$ para cada $k \in \mathbb{N}$ ($b \neq 0$ por ser no degenerado con respecto a B). Como $0 < \left\|b_k^\delta - b\right\| = \frac{\delta}{2k} < \delta$, se tiene que $\{b_k^\delta\}_{k=1}^{\infty}$ está en el entorno reducido de centro en b y radio δ. Claramente, $b_k^\delta \to b$ cuando $k \to \infty$. Puesto que $Ran(A)$ es un subespacio de $\mathcal{X}$, resulta que $b_k^\delta \in Ran(A)\, \forall\ k \in \mathbb{N}$. Si x_N es la solución de mínimos cuadrados de $Ax = b$ en $\mathcal{X}_N$, entonces $x_N^k \doteq \left(1 + \frac{\delta}{2k\|b\|}\right) x_N$ es la solución de mínimos cuadrados de $Ax = b_k^\delta$ en $\mathcal{X}_N$. Ahora bien,

$$\begin{aligned}
\left\|x_N^k - A^{-1}b_k^\delta\right\| &= \left(1 + \frac{\delta}{2k\,\|b\|}\right)\left\|x_N - A^{-1}b\right\| \\
&\geq \left\|x_N - A^{-1}b\right\| \\
&\geq f(N)\ \forall\ N \in \mathbb{N},
\end{aligned}$$

donde la última desigualdad se sigue del Corolario 5.6. ■

Observación 5.9. En la demostración se usó el siguiente resultado que puede probarse de una manera muy simple: si x_N es la solución de mínimos cuadrados de $Ax = b$ en $\mathcal{X}_N$, entonces σx_N con $\sigma \neq 0$, es la solución de mínimos cuadrados de $Ax = \sigma b$ en $\mathcal{X}_N$.

Para el caso en que el operador $A : \mathcal{X} \to \mathcal{X}$ está dado, como sucede la mayoría de las veces en la práctica, Seidman ([54]) probó que para "casi todo" $b \in Ran(A)$

existe una sucesión de supespacios $\mathcal{X}_N$ tales que la correspondiente sucesión $\{x_N\}$ de soluciones de mínimos cuadrados de $Ax = b$ no está acotada. Aunque creemos que los resultados de esta sección también se pueden extender a este caso, es decir, que cuando el operador A está dado también se pueden obtener tasas arbitrarias de divergencia debido a la aplicación del método de mínimos cuadrados, todavía no pudimos demostrar esta conjetura aún permanece abierta.

5.2. Divergencia para problemas inversos con datos perturbados

En la sección anterior hemos considerado problemas con dato exacto, es decir, ecuaciones de la forma $Ax = b^*$, donde $b^* \doteq Ax^*$ y x^* es la solución exacta. Como es de esperar, la no convergencia señalada en el Corolario 5.6 también puede producirse con datos con ruido. T. Seidman ([54]) probó que si A es un operador compacto (además de ser lineal, inyectivo, autoadjunto, positivo y de rango denso), entonces para cada sucesión creciente de subespacios de dimensión finita $\{\mathcal{X}_N\}$ cuya unión es densa en $\mathcal{X}$ y para cada $b^* \in Ran(A)$ existe una sucesión $b_N \to b^*$ para la cual las soluciones de mínimos cuadrados x_N de $Ax = b_N$ en $\mathcal{X}_N$ satisfacen $\|x_N - x^*\| \to \infty$, donde $Ax^* = b^*$. En un cierto sentido, el siguiente corolario generaliza este resultado. Más precisamente, se muestra que también para el caso de datos con ruido la aplicación sin cuidado del método de mínimos cuadrados puede conducir a tasas de divergencia arbitrarias.

Corolario 5.10. *(tasa de divergencia arbitraria para datos con ruido). Sean $\mathcal{X}$ un espacio de Hilbert separable de dimensión infinita, $B \doteq \{e_n\}_{n=1}^{\infty}$ una base ortonormal de $\mathcal{X}$, $b \in \mathcal{X}$ no degenerado con respecto a B, $\mathcal{X}_N \doteq \mathrm{span}\{e_1, ..., e_N\}$, $f : \mathbb{N} \to \mathbb{R}^+$ una función creciente arbitraria y $A = A(b, f)$ el operador cuya existencia se probó en el Corolario 5.6. Entonces $\forall\, \alpha \in (0, 1)$ existe una sucesión $\{b_N\}_{N=1}^{\infty} \subset Ran(A)$ tal que*

$b_N \to b$ cuando $N \to \infty$ y la solución de mínimos cuadrados x_N de $Ax = b_N$ en $\mathcal{X}_N$ satisface $\|x_N - A^{-1}b\| \geqslant \alpha f(N)$ para todo $N \in \mathbb{N}$.

Demostración. Sea $\{\alpha_N\} \subset \mathbb{R}$ una sucesión tal que $\alpha_N \neq 0$ y

$$|1-\alpha_N| \leq \text{mín}\left\{\frac{1}{\|x_N^*\|}, \frac{(1-\alpha)f(N)}{\|x_N^*\|}\right\}, \tag{5.24}$$

donde x_N^* es la solución de mínimos cuadrados de $Ax = b$ en $\mathcal{X}_N$. Se sigue del Corolario 5.6 que,

$$\left\|x_N^* - A^{-1}b\right\| \geq f(N) \; \forall \; N \in \mathbb{N}. \tag{5.25}$$

Entonces $\|x_N^*\| \to \infty$, de lo cual se deduce que $|1-\alpha_N| \to 0$ y por lo tanto, $\alpha_N \to 1$ cuando $N \to \infty$. Definamos $b_N \doteq \alpha_N b$ para cada $N \in \mathbb{N}$. Luego, $b_N \to b$ y como $Ran(A)$ es un subespacio de $\mathcal{X}$, $\{b_N\}_{N=1}^{\infty} \subset Ran(A)$. De la Observación 5.9 resulta que la solución de mínimos cuadrados x_N de $Ax = b_N = \alpha_N b$ en $\mathcal{X}_N$ es $\alpha_N x_N^*$. Entonces

$$\begin{aligned}
\left\|x_N - A^{-1}b\right\| &\geq \left\|x_N^* - A^{-1}b\right\| - \|x_N^* - x_N\| \\
&= \left\|x_N^* - A^{-1}b\right\| - |1-\alpha_N| \, \|x_N^*\| \\
&\geq f(N) - (1-\alpha)f(N) \quad \text{(por (5.24) y (5.25))} \\
&= \alpha f(N).
\end{aligned}$$

■

Por simplicidad, sólo hemos considerado el caso en el cual $\mathcal{X}$ e $\mathcal{Y}$ son espacios de Hilbert reales. Sin embargo, estos resultados siguen siendo válidos para el caso de espacios de Hilbert complejos, con las obvias modificaciones. Además, se puede usar un procedimiento similar al empleado en el Corolario 5.6 para mostrar que la hipótesis de separabilidad del espacio $\mathcal{X}$ puede ser ignorada.

Capítulo 6

Saturación global de métodos de regularización arbitrarios

Como mencionamos en la Sección 3.1.3, en 1994 A. Neubauer ([49]) demostró que ciertos métodos de regularización espectrales "*saturan*", es decir son incapaces de continuar extrayendo información acerca de la solución exacta aún bajo hipótesis adicionales de regularidad sobre la misma. En ese artículo Neubauer introdujo por primera vez "la idea" del concepto de "*saturación*" para métodos de regularización. La idea del concepto se refiere al mejor orden de convergencia (del error total) que un método puede alcanzar independientemente de los supuestos de regularidad de la solución exacta y de la selección de la regla de elección del parámetro. No obstante el concepto es bastante intuitivo y puede observarse claramente aún en métodos clásicos tales como el de Tikhonov-Phillips (ver [50]). El mismo ha escapado siempre a una adecuada formalización matemática que permita describirlo con claridad y echar luz sobre las causas por las cuales algunos métodos de regularización saturan y otros no.

En 2001, P. Mathé y S. V. Pereverzev ([43]) utilizaron escalas de Hilbert para estudiar la eficiencia de soluciones aproximadas basadas en observaciones con ruido (estocástico o determinístico). En este contexto se puede cuantificar el grado de mal condicionamiento de un problema y obtener condiciones generales sobre métodos de proyección para que estos alcancen orden óptimo de convergencia. Estos conceptos fueron extendidos luego por los mismos autores ([44]) quienes estudiaron el problema de convergencia óptima en escalas de Hilbert variables. En este trabajo se mostró que existe una relación muy estrecha entre la convergencia óptima de un método y la regularidad "*a-priori*" (en términos de conjuntos fuente) para métodos espectrales que poseen calificación clásica.

En 2004, P. Mathé ([42]) propuso definiciones de los conceptos de calificación y saturación para métodos de regularización espectrales. Sin embargo, el concepto de saturación definido por Mathé no es aplicable a métodos de regularización arbitrarios y no es totalmente compatible con la idea original de saturación propuesta por Neubauer en [49]. En particular, la definición de saturación dada en [42] no implica unicidad y por lo tanto, tampoco un orden óptimo de convergencia.

En este capítulo desarrollaremos una teoría general de saturación para métodos de regularización arbitrarios (espectrales o no). La teoría que aquí desarrollamos atendiendo a la idea original de Neubauer, parte de observar que existe una marcada relación de optimalidad y dualidad entre el mejor orden global de convergencia (del error total) del método, y el conjunto fuente sobre el cual se alcanza ese mejor orden de convergencia. En particular, probaremos que para que una familia de operadores de regularización posea saturación, el error total debe ser "doblemente optimal", en el sentido que debe ser orden de convergencia óptimo sobre un cierto conjunto y a su vez este conjunto debe satisfacer una cierta condición de optimalidad con respecto al error. En la Sección 6.1 definiremos cotas de convergencia para una familia de operadores de regularización y desarrollaremos un andamiaje apropiado para la comparación de las mismas. En la Sección 6.2 introduciremos el concepto de saturación global, mostraremos su relación con el error total y con cotas de convergencia óptimas y daremos condiciones necesarias y suficientes para la existencia de saturación. En la Sección 6.3 demostraremos algunos resultados recíprocos que son utilizados luego, junto con los conceptos desarrollados en la Sección 6.2, para caracterizar la saturación de métodos espectrales con calificación clásica y de métodos espectrales con calificación máxima. Algunos de los resultados de este capítulo forman parte de un artículo que ha sido enviado para su publicación (ver [58]).

6.1. Cotas de convergencia para métodos de regularización

De ahora en más en este capítulo, a menos que se especifique lo contrario, supondremos que todos los subconjuntos del espacio de Hilbert $\mathcal{X}$ que consideraremos son no vacíos y no contienen al cero y además, sin pérdida de generalidad, que el operador T es invertible. Dado $M \subset \mathcal{X}$, denotaremos con $\mathcal{F}_M$ al conjunto de las siguientes funciones: diremos que $\psi \in \mathcal{F}_M$ si existe $a = a(\psi) > 0$ tal que ψ está definida en $M \times (0, a)$, a valores en $(0, \infty)$ y satisface las siguientes condiciones:

1. $\lim\limits_{\delta \to 0^+} \psi(x, \delta) = 0$ para todo $x \in M$, y
2. ψ es continua y creciente como función de δ en $(0, a)$ para cada $x \in M$ fijo.

Definiremos ahora distintas relaciones entre dos funciones de $\mathcal{F}_M$, donde M es un subconjunto de $\mathcal{X}$.

Definición 6.1. Sean $M \subset \mathcal{X}$ y $\psi, \tilde{\psi} \in \mathcal{F}_M$.

(i) Decimos que "*ψ precede a $\tilde{\psi}$ sobre M*", y lo denotamos $\psi \overset{M}{\preceq} \tilde{\psi}$, si existen una constante $r > 0$ y $p : M \to (0, \infty)$ tales que $\psi(x, \delta) \leq p(x)\tilde{\psi}(x, \delta)$ para todo $x \in M$ y para todo $\delta \in (0, r)$.

(ii) Decimos que "*$\psi, \tilde{\psi}$ son equivalentes sobre M*", y lo denotamos $\psi \overset{M}{\approx} \tilde{\psi}$, si $\psi \overset{M}{\preceq} \tilde{\psi}$ y $\tilde{\psi} \overset{M}{\preceq} \psi$.

(iii) Decimos que "*ψ precede estrictamente a $\tilde{\psi}$ sobre M*" y lo denotamos $\psi \overset{M}{\prec} \tilde{\psi}$ si $\psi \overset{M}{\preceq} \tilde{\psi}$ y $\limsup\limits_{\delta\to 0^+} \frac{\psi(x,\delta)}{\tilde{\psi}(x,\delta)} = 0$ para todo $x \in M$.

Las siguientes observaciones se siguen inmediatamente de estas definiciones.

- Dado que $\psi, \tilde{\psi} > 0$, en **(iii)** la condición $\limsup\limits_{\delta\to 0^+} \frac{\psi(x,\delta)}{\tilde{\psi}(x,\delta)} = 0$ es equivalente a $\lim\limits_{\delta\to 0^+} \frac{\psi(x,\delta)}{\tilde{\psi}(x,\delta)} = 0$, es decir, $\psi(x,\delta) = o(\tilde{\psi}(x,\delta))$ cuando $\delta \to 0^+$.
- La relación "$\overset{M}{\preceq}$" introduce un orden parcial en $\mathcal{F}_M$ y "$\overset{M}{\approx}$" es una relación de equivalencia en $\mathcal{F}_M$.
- Si $\psi \overset{M}{\preceq} (\overset{M}{\approx}, \overset{M}{\prec})\, \tilde{\psi}$ entonces $\psi \overset{\tilde{M}}{\preceq} (\overset{\tilde{M}}{\approx}, \overset{\tilde{M}}{\prec})\, \tilde{\psi}$ para todo $\tilde{M} \subset M$.

Con $\not\preceq$, $\not\prec$ y $\not\approx$ denotaremos la negación de las relaciones $\preceq$, $\prec$ y $\approx$, respectivamente.

Lema 6.2. *Sean $M \subset \mathcal{X}$ y $\psi, \tilde{\psi} \in \mathcal{F}_M$. Si $\psi \overset{M}{\prec} \tilde{\psi}$ entonces $\tilde{\psi} \overset{\tilde{M}}{\not\preceq} \psi$ para todo $\tilde{M} \subset M$.*

Demostración. Por el contrarrecíproco. Supongamos que existe $\tilde{M} \subset M$ tal que $\tilde{\psi} \overset{\tilde{M}}{\preceq} \psi$. Sea $x_0 \in \tilde{M}$, entonces $\tilde{\psi} \overset{\{x_0\}}{\preceq} \psi$, es decir, existen constantes $0 < p < \infty$ y $r > 0$ tales que $\sup\limits_{\delta\in(0,r)} \frac{\tilde{\psi}(x_0,\delta)}{\psi(x_0,\delta)} \le p < \infty$. Luego,

$$
\begin{aligned}
\limsup_{\delta\to 0^+} \frac{\psi(x_0,\delta)}{\tilde{\psi}(x_0,\delta)} &\ge \liminf_{\delta\to 0^+} \frac{\psi(x_0,\delta)}{\tilde{\psi}(x_0,\delta)} \ge \inf_{\delta\in(0,r)} \frac{\psi(x_0,\delta)}{\tilde{\psi}(x_0,\delta)} \\
&= \left(\sup_{\delta\in(0,r)} \frac{\tilde{\psi}(x_0,\delta)}{\psi(x_0,\delta)} \right)^{-1} \ge \frac{1}{p} > 0.
\end{aligned}
$$

Por lo tanto, $\psi \overset{\{x_0\}}{\not\prec} \tilde{\psi}$, de lo cual se deduce que $\psi \overset{M}{\not\prec} \tilde{\psi}$. ■

Dado un método de regularización $\left(\{R_\alpha\}_{\alpha\in(0,\alpha_0)}, \alpha(\delta, y^\delta)\right)$ para el problema $Tx = y$ (ver Definición 2.1), en la Sección 2.2 definimos el error total que surge de comparar la solución regularizada $R_{\alpha(\delta,y^\delta)}y^\delta$ con la solución exacta $T^\dagger y$ (ver (2.19)). A continuación definiremos el error total de una familia de operadores de regularización $\{R_\alpha\}_{\alpha\in(0,\alpha_0)}$ en el sentido de la mayor discrepancia posible que puede obtenerse para una observación con nivel de ruido δ, con cualquier elección del parámetro de regularización α. Más precisamente, se tiene la siguiente definición.

Definición 6.3. Sean $\mathcal{X}, \mathcal{Y}$ espacios de Hilbert, $T \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$ y $\{R_\alpha\}_{\alpha\in(0,\alpha_0)}$ una familia de operadores de regularización para $T^\dagger$. Definimos el "*error total de* $\{R_\alpha\}_{\alpha\in(0,\alpha_0)}$ *en* $x \in \mathcal{X}$ *para un nivel de ruido* δ" como

$$\mathcal{E}^{\text{tot}}_{\{R_\alpha\}}(x,\delta) \doteq \inf_{\alpha\in(0,\alpha_0)} \sup_{y^\delta\in\overline{B_\delta(Tx)}} \left\| R_\alpha y^\delta - x \right\|,$$

donde $\overline{B_\delta(Tx)} \doteq \{y \in \mathcal{Y} : \|Tx - y\| \leq \delta\}$.

Observación 6.4. Sean $a > 0$, $M \subset \mathcal{X}$ y $\mathcal{E}^{\text{tot}}_{\{R_\alpha\}} : M \times (0, a) \to (0, \infty)$ el error total. Entonces $\mathcal{E}^{\text{tot}}_{\{R_\alpha\}} \in \mathcal{F}_M$. En efecto, para cada $x \in M$, $\mathcal{E}^{\text{tot}}_{\{R_\alpha\}}(x, \delta)$ es creciente como función de δ, y dado que $\{R_\alpha\}$ es una familia de operadores de regularización, resulta que $\mathcal{E}^{\text{tot}}_{\{R_\alpha\}}(x, \delta)$ es continuo como función de δ para cada $x \in M$ fijo y $\lim_{\delta\to 0^+} \mathcal{E}^{\text{tot}}_{\{R_\alpha\}}(x, \delta) = 0$ para todo $x \in M$.

Ahora estamos en condiciones de definir las cotas superiores de convergencia para el error total de una familia de operadores de regularización.

Definición 6.5. Sean $\mathcal{X}, \mathcal{Y}$ espacios de Hilbert, $T \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$ y $\{R_\alpha\}_{\alpha\in(0,\alpha_0)}$ una familia de operadores de regularización para $T^\dagger$, $M \subset \mathcal{X}$ y $\psi \in \mathcal{F}_M$.

(i) Decimos que ψ es una "*cota superior de convergencia para el error total de* $\{R_\alpha\}_{0<\alpha\leq\alpha_0}$ *en* M" si $\mathcal{E}^{\text{tot}}_{\{R_\alpha\}} \overset{M}{\preceq} \psi$.

(ii) Decimos que ψ es "*cota superior estricta de convergencia para el error total de* $\{R_\alpha\}_{0<\alpha\leq\alpha_0}$ *en* M" si $\mathcal{E}^{\text{tot}}_{\{R_\alpha\}} \overset{M}{\prec} \psi$.

(iii) Decimos que ψ es "*cota superior óptima de convergencia para el error total de* $\{R_\alpha\}_{0<\alpha\leq\alpha_0}$ *en* M" si $\mathcal{E}^{\text{tot}}_{\{R_\alpha\}} \overset{M}{\preceq} \psi$ y

$$\limsup_{\delta\to 0^+} \frac{\mathcal{E}^{\text{tot}}_{\{R_\alpha\}}(x,\delta)}{\psi(x,\delta)} > 0 \quad \text{para todo} \quad x \in M,$$

o equivalentemente, si $\mathcal{E}^{\text{tot}}_{\{R_\alpha\}}(x,\delta) \neq o(\psi(x,\delta))$ cuando $\delta \to 0^+$.

Denotaremos con $\mathcal{U}_M(\mathcal{E}^{\text{tot}}_{\{R_\alpha\}})$, $\mathcal{U}^{\text{est}}_M(\mathcal{E}^{\text{tot}}_{\{R_\alpha\}})$ y $\mathcal{U}^{\text{opt}}_M(\mathcal{E}^{\text{tot}}_{\{R_\alpha\}})$ al conjunto de todas las funciones $\psi \in \mathcal{F}_M$ que son, respectivamente, cotas superiores, cotas superiores estrictas y cotas superiores óptimas de convergencia para el error total de $\{R_\alpha\}_{0<\alpha\leq\alpha_0}$ en M. En virtud de la Observación 6.4, es claro que $\mathcal{E}^{\text{tot}}_{\{R_\alpha\}}(\cdot,\cdot) \in \mathcal{U}^{\text{opt}}_M(\mathcal{E}^{\text{tot}}_{\{R_\alpha\}})$ para todo $M \subset \mathcal{X}$.

De las definiciones previas se sigue inmediatamente que:

- Si $\psi \in \mathcal{F}_M$, entonces $\psi \in \mathcal{U}_M(\mathcal{E}^{\text{tot}}_{\{R_\alpha\}})$ si (y sólo si) $\mathcal{E}^{\text{tot}}_{\{R_\alpha\}}(x,\delta) = O(\psi(x,\delta))$ cuando $\delta \to 0^+$ para todo $x \in M$. Además, $\mathcal{U}^{\text{est}}_M(\mathcal{E}^{\text{tot}}_{\{R_\alpha\}})$ y $\mathcal{U}^{\text{opt}}_M(\mathcal{E}^{\text{tot}}_{\{R_\alpha\}})$ son subconjuntos disjuntos de $\mathcal{U}_M(\mathcal{E}^{\text{tot}}_{\{R_\alpha\}})$, aunque su unión no es todo $\mathcal{U}_M(\mathcal{E}^{\text{tot}}_{\{R_\alpha\}})$ (excepto cuando el conjunto M tiene un solo elemento).
- Si $\tilde{M} \subset M$, entonces $\mathcal{U}_M(\mathcal{E}^{\text{tot}}_{\{R_\alpha\}}) \subset \mathcal{U}_{\tilde{M}}(\mathcal{E}^{\text{tot}}_{\{R_\alpha\}})$, $\mathcal{U}^{\text{opt}}_M(\mathcal{E}^{\text{tot}}_{\{R_\alpha\}}) \subset \mathcal{U}^{\text{opt}}_{\tilde{M}}(\mathcal{E}^{\text{tot}}_{\{R_\alpha\}})$ y $\mathcal{U}^{\text{est}}_M(\mathcal{E}^{\text{tot}}_{\{R_\alpha\}}) \subset \mathcal{U}^{\text{est}}_{\tilde{M}}(\mathcal{E}^{\text{tot}}_{\{R_\alpha\}})$.
- Si $\psi \in \mathcal{U}_M(\mathcal{E}^{\text{tot}}_{\{R_\alpha\}})$, $\tilde{\psi} \in \mathcal{F}_M$ y $\psi \overset{M}{\preceq} \tilde{\psi}$, entonces $\tilde{\psi} \in \mathcal{U}_M(\mathcal{E}^{\text{tot}}_{\{R_\alpha\}})$.
- Si $\psi \in \mathcal{U}^{\text{opt}}_M(\mathcal{E}^{\text{tot}}_{\{R_\alpha\}})$, $\tilde{\psi} \in \mathcal{U}_M(\mathcal{E}^{\text{tot}}_{\{R_\alpha\}})$ y $\tilde{\psi} \overset{M}{\preceq} \psi$, entonces $\tilde{\psi} \in \mathcal{U}^{\text{opt}}_M(\mathcal{E}^{\text{tot}}_{\{R_\alpha\}})$.
- Si $\psi \in \mathcal{U}^{\text{est}}_M(\mathcal{E}^{\text{tot}}_{\{R_\alpha\}})$, $\tilde{\psi} \in \mathcal{F}_M$ y $\psi \overset{M}{\preceq} \tilde{\psi}$, entonces $\tilde{\psi} \in \mathcal{U}^{\text{est}}_M(\mathcal{E}^{\text{tot}}_{\{R_\alpha\}})$.

Definición 6.6. Sean $\psi, \tilde{\psi} \in \mathcal{F}_M$. Decimos que "$\psi$ *y* $\tilde{\psi}$ *son comparables sobre* M" si verifican $\psi \overset{M}{\preceq} \tilde{\psi}$ ó $\tilde{\psi} \overset{M}{\preceq} \psi$ (ó ambas).

Definición 6.7. Sean $\mathcal{A} \subset \mathcal{F}_M$ y $\psi^* \in \mathcal{A}$. Decimos que "ψ^* *es un elemento minimal de* $\left(\mathcal{A}, \overset{M}{\preceq}\right)$" si $\psi^* \overset{M}{\preceq} \psi$ para todo $\psi \in \mathcal{A}$ comparable con ψ^* sobre M. Equivalentemente, ψ^* es minimal de $\left(\mathcal{A}, \overset{M}{\preceq}\right)$ si para todo $\psi \in \mathcal{A}$, la condición $\psi \overset{M}{\preceq} \psi^*$ implica $\psi^* \overset{M}{\preceq} \psi$.

Lema 6.8. *Sean $\mathcal{A} \subset \mathcal{F}_M, \psi, \psi^* \in \mathcal{A}$ y ψ, ψ^* comparables sobre M. Si existe $M_0 \subset M$ tal que $\psi \overset{M_0}{\prec} \psi^*$ entonces ψ^* no es un elemento minimal de $\left(\mathcal{A}, \overset{M}{\preceq}\right)$.*

Demostración. Sean $\mathcal{A} \subset \mathcal{F}_M$ y $\psi, \psi^* \in \mathcal{A}$ comparables sobre M. Supongamos que existe $M_0 \subset M$ tal que $\psi \overset{M_0}{\prec} \psi^*$, entonces se sigue del Lema 6.2 que $\psi^* \overset{M_0}{\not\preceq} \psi$. Luego, $\psi^* \overset{M}{\not\preceq} \psi$ y como $\psi, \psi^* \in \mathcal{A}$ son comparables sobre M, se deduce de la Definición 6.7 que ψ^* no puede ser un elemento minimal de $\left(\mathcal{A}, \overset{M}{\preceq}\right)$. ■

Corolario 6.9. *Si $\psi^* \in \mathcal{U}_M(\mathcal{E}^{tot}_{\{R_\alpha\}})$ y existen $x_0 \in M$ y $\psi_0 \in \mathcal{U}_{\{x_0\}}(\mathcal{E}^{tot}_{\{R_\alpha\}})$ tales que $\psi_0 \overset{\{x_0\}}{\prec} \psi^*$ entonces ψ^* no es un elemento minimal de $\left(\mathcal{U}_M(\mathcal{E}^{tot}_{\{R_\alpha\}}), \overset{M}{\preceq}\right)$.*

Demostración. Este corolario es consecuencia inmediata del lema anterior con $\mathcal{A} = \mathcal{U}_M(\mathcal{E}^{\text{tot}}_{\{R_\alpha\}})$, $M_0 = \{x_0\}$ y

$$\psi(x, \delta) \doteq \begin{cases} \psi_0(x_0, \delta), & \text{si } x = x_0 \\ \psi^*(x, \delta), & \text{si } x \neq x_0. \end{cases}$$

Notar que esta función ψ así definida está en $\mathcal{U}_M(\mathcal{E}^{\text{tot}}_{\{R_\alpha\}})$ y es comparable con ψ^* sobre M (más aún $\psi \overset{M}{\preceq} \psi^*$). ■

A continuación probaremos que las cotas superiores óptimas de convergencia para el error total de $\{R_\alpha\}_{0<\alpha\leq\alpha_0}$ en M están caracterizadas por ser elementos minimales del conjunto parcialmente ordenado $\left(\mathcal{U}_M(\mathcal{E}^{\text{tot}}_{\{R_\alpha\}}), \overset{M}{\preceq}\right)$. Más precisamente, se tiene el siguiente resultado.

Teorema 6.10. *Sea $\psi \in \mathcal{U}_M(\mathcal{E}^{tot}_{\{R_\alpha\}})$. Entonces $\psi \in \mathcal{U}^{opt}_M(\mathcal{E}^{tot}_{\{R_\alpha\}})$ si y sólo si ψ es un elemento minimal de $\left(\mathcal{U}_M(\mathcal{E}^{tot}_{\{R_\alpha\}}), \overset{M}{\preceq}\right)$.*

Demostración. Sea $\psi \in \mathcal{U}^{\text{opt}}_M(\mathcal{E}^{\text{tot}}_{\{R_\alpha\}})$ y supongamos que ψ no es un elemento minimal de $\left(\mathcal{U}_M(\mathcal{E}^{\text{tot}}_{\{R_\alpha\}}), \overset{M}{\preceq}\right)$. Entonces existe $\psi_c \in \mathcal{U}_M(\mathcal{E}^{\text{tot}}_{\{R_\alpha\}})$ comparable con ψ sobre M para la cual no es cierto que $\psi \overset{M}{\preceq} \psi_c$. Luego, existe $x_0 \in M$ tal que

$$\limsup_{\delta\to 0^+} \frac{\psi(x_0, \delta)}{\psi_c(x_0, \delta)} = \infty. \tag{6.1}$$

Ahora, como $\psi \in \mathcal{U}_M^{\rm opt}(\mathcal{E}^{\rm tot}_{\{R_\alpha\}})$ y $x_0 \in M$, se tiene que

$$\limsup_{\delta\to 0^+} \frac{\mathcal{E}^{\rm tot}_{\{R_\alpha\}}(x_0,\delta)}{\psi(x_0,\delta)} > 0. \tag{6.2}$$

Puesto que

$$\limsup_{\delta\to 0^+} \frac{\mathcal{E}^{\rm tot}_{\{R_\alpha\}}(x_0,\delta)}{\psi_c(x_0,\delta)} = \limsup_{\delta\to 0^+} \frac{\mathcal{E}^{\rm tot}_{\{R_\alpha\}}(x_0,\delta)}{\psi(x_0,\delta)}\frac{\psi(x_0,\delta)}{\psi_c(x_0,\delta)},$$

se deduce utilizando (6.1) y (6.2) que $\limsup_{\delta\to 0^+} \frac{\mathcal{E}^{\rm tot}_{\{R_\alpha\}}(x_0,\delta)}{\psi_c(x_0,\delta)} = \infty$, lo cual implica que $\psi_c \notin \mathcal{U}_{\{x_0\}}(\mathcal{E}^{\rm tot}_{\{R_\alpha\}})$. Esto contradice el hecho que $\psi_c \in \mathcal{U}_M(\mathcal{E}^{\rm tot}_{\{R_\alpha\}})$. Luego, ψ es un elemento minimal de $\left(\mathcal{U}_M(\mathcal{E}^{\rm tot}_{\{R_\alpha\}}), \overset{M}{\preceq}\right)$.

Recíprocamente, supongamos que $\psi \in \mathcal{U}_M(\mathcal{E}^{\rm tot}_{\{R_\alpha\}})$ y $\psi \notin \mathcal{U}_M^{\rm opt}(\mathcal{E}^{\rm tot}_{\{R_\alpha\}})$. Entonces existe $x_0 \in M$ tal que $\psi \in \mathcal{U}^{\rm est}_{\{x_0\}}(\mathcal{E}^{\rm tot}_{\{R_\alpha\}})$, lo cual implica que $\mathcal{E}^{\rm tot}_{\{R_\alpha\}} \overset{\{x_0\}}{\prec} \psi$. En virtud del Lema 6.8 se deduce que ψ no es un elemento minimal de $\left(\mathcal{U}_M(\mathcal{E}^{\rm tot}_{\{R_\alpha\}}), \overset{M}{\preceq}\right)$. ■

Observar que ψ es un elemento minimal de $\left(\mathcal{U}_M(\mathcal{E}^{\rm tot}_{\{R_\alpha\}}), \overset{M}{\preceq}\right)$ si y sólo si es minimal de $\left(\mathcal{U}_{M^*}(\mathcal{E}^{\rm tot}_{\{R_\alpha\}}), \overset{M^*}{\preceq}\right)$ para todo $M^* \subset M$.

Corolario 6.11. *Sean $\mathcal{U}_M^{opt}(\mathcal{E}^{tot}_{\{R_\alpha\}})$ y $\mathcal{E}^{tot}_{\{R_\alpha\}}$ como antes.*

(i) *Si $\psi \in \mathcal{U}_M^{opt}(\mathcal{E}^{tot}_{\{R_\alpha\}})$ entonces $\psi \overset{M}{\approx} \mathcal{E}^{tot}_{\{R_\alpha\}}$.*

(ii) *Si $\psi, \tilde{\psi} \in \mathcal{U}_M^{opt}(\mathcal{E}^{tot}_{\{R_\alpha\}})$ entonces $\psi \overset{M}{\approx} \tilde{\psi}$.*

Demostración. ***(i)*** Si $\psi \in \mathcal{U}_M^{\rm opt}(\mathcal{E}^{\rm tot}_{\{R_\alpha\}})$ entonces $\mathcal{E}^{\rm tot}_{\{R_\alpha\}} \overset{M}{\preceq} \psi$, de lo cual resulta que $\mathcal{E}^{\rm tot}_{\{R_\alpha\}}$ y ψ son comparables sobre M. Luego, como $\mathcal{E}^{\rm tot}_{\{R_\alpha\}} \in \mathcal{U}_M(\mathcal{E}^{\rm tot}_{\{R_\alpha\}})$ y por el Teorema 6.10 ψ es un elemento minimal de $\left(\mathcal{U}_M(\mathcal{E}^{\rm tot}_{\{R_\alpha\}}), \overset{M}{\preceq}\right)$, se tiene que $\psi \overset{M}{\preceq} \mathcal{E}^{\rm tot}_{\{R_\alpha\}}$. Por lo tanto, $\psi \overset{M}{\approx} \mathcal{E}^{\rm tot}_{\{R_\alpha\}}$ como queríamos probar.

(ii) Es consecuencia inmediata de ***(i)*** y de las propiedades de la relación de equivalencia "$\overset{M}{\approx}$", pues toda $\psi \in \mathcal{U}_M^{\rm opt}(\mathcal{E}^{\rm tot}_{\{R_\alpha\}})$ es equivalente a $\mathcal{E}^{\rm tot}_{\{R_\alpha\}}$ sobre M. ■

Observar que si ψ es una cota superior óptima de convergencia en M para el error total de una familia de operadores de regularización, entonces en todo punto de M ψ tiende a cero, cuando el nivel de ruido tiende a cero, exactamente con la misma velocidad con que lo hace el error total.

Para introducir el concepto de saturación en la siguiente sección, necesitaremos además algunas definiciones y herramientas que nos permitirán comparar cotas de convergencia en diferentes conjuntos de $\mathcal{X}$.

Definición 6.12. Sean $M, \tilde{M} \subset \mathcal{X}$, $\psi \in \mathcal{F}_M$ y $\tilde{\psi} \in \mathcal{F}_{\tilde{M}}$.

(i) Decimos que "*ψ en M precede a $\tilde{\psi}$ en $\tilde{M}$*", y lo denotamos $\psi \overset{M,\tilde{M}}{\preceq} \tilde{\psi}$, si existen una constante $d > 0$ y $k : M \times \tilde{M} \to (0, \infty)$ tales que $\psi(x, \delta) \leq k(x, \tilde{x})\, \tilde{\psi}(\tilde{x}, \delta)$ para todo $x \in M$, para todo $\tilde{x} \in \tilde{M}$ y para todo $\delta \in (0, d)$.

(ii) Decimos que "*ψ en M es equivalente a $\tilde{\psi}$ en $\tilde{M}$*", y lo denotamos $\psi \overset{M,\tilde{M}}{\approx} \tilde{\psi}$, si $\psi \overset{M,\tilde{M}}{\preceq} \tilde{\psi}$ y $\tilde{\psi} \overset{\tilde{M},M}{\preceq} \psi$.

(iii) Decimos que "*ψ en M precede estrictamente a $\tilde{\psi}$ en $\tilde{M}$*", y lo denotamos $\psi \overset{M,\tilde{M}}{\prec} \tilde{\psi}$, si $\psi \overset{M,\tilde{M}}{\preceq} \tilde{\psi}$ y $\limsup_{\delta \to 0^+} \frac{\psi(x,\delta)}{\tilde{\psi}(\tilde{x},\delta)} = 0$ para todo $x \in M$ y para todo $\tilde{x} \in \tilde{M}$.

Observación 6.13. Las definiciones precedentes generalizan en un cierto sentido (aunque son ligeramente más fuertes que) las relaciones introducidas en la Definición 6.1. Observar por ejemplo que si $\psi \overset{M,M}{\prec} \tilde{\psi}$ entonces $\psi \overset{M}{\prec} \tilde{\psi}$, aunque el recíproco, en general, no es cierto.

Se sigue inmediatamente de la Definición 6.12 que si $\psi \overset{M,N}{\preceq} \tilde{\psi}$ entonces $\psi \overset{\tilde{M},\tilde{N}}{\preceq} \tilde{\psi}$ para todo $\tilde{M} \subset M$ y para todo $\tilde{N} \subset N$. Lo mismo ocurre para las relaciones "$\overset{M,N}{\approx}$" y "$\overset{M,N}{\prec}$".

A continuación extendemos la noción de "comparabilidad" dada en la Definición 6.6, ahora para el caso de subconjuntos distintos de $\mathcal{X}$.

Definición 6.14. Sean $M, \tilde{M} \subset \mathcal{X}$, $\psi \in \mathcal{F}_M$ y $\tilde{\psi} \in \mathcal{F}_{\tilde{M}}$.

(i) Decimos que "*ψ sobre M es comparable con $\tilde{\psi}$ sobre $\tilde{M}$*" si $\psi \overset{M,\tilde{M}}{\preceq} \tilde{\psi}$ o bien $\tilde{\psi} \overset{\tilde{M},M}{\preceq} \psi$.

(ii) Decimos que "*ψ es invariante sobre M*" si $\psi \overset{M,M}{\approx} \psi$.

Observación 6.15. Es inmediato que la condición $\psi \overset{M,M}{\approx} \psi$ es equivalente a $\psi \overset{M,M}{\preceq} \psi$.

Esta última noción de invariancia establece, esencialmente, que si ψ es invariante sobre M entonces los órdenes de convergencia de ψ como función de δ para $\delta \to 0^+$, en dos puntos cualesquiera de M, son equivalentes.

El siguiente resultado está relacionado con una cierta propiedad de "transitividad" de la relación de invariancia.

Lema 6.16. *Sean $M \subset \mathcal{X}$, $\psi, \tilde{\psi} \in \mathcal{F}_M$, tales que $\tilde{\psi} \overset{M}{\approx} \psi$ y $\psi \overset{M,M}{\approx} \psi$. Entonces:*

(i) *$\tilde{\psi} \overset{M,M}{\approx} \psi$ y*

(ii) *$\tilde{\psi} \overset{M,M}{\approx} \tilde{\psi}$ (i.e. $\tilde{\psi}$ es también invariante sobre M).*

Demostración. Sean $M \subset \mathcal{X}$, $\psi, \tilde{\psi} \in \mathcal{F}_M$ y $x, \tilde{x} \in M$. Supongamos que $\psi \overset{M,M}{\approx} \psi$ y $\tilde{\psi} \overset{M}{\approx} \psi$.

(i) Como $\tilde{\psi} \overset{M}{\approx} \psi$, existen constantes positivas d, k_x y $k_{\tilde{x}}$ tales que para todo $\delta \in (0, d)$,

$$\tilde{\psi}(x, \delta) \leq k_x \psi(x, \delta) \quad \text{y} \quad \psi(\tilde{x}, \delta) \leq k_{\tilde{x}} \tilde{\psi}(\tilde{x}, \delta). \tag{6.3}$$

De la invariancia de ψ sobre M resulta que existen constantes positivas d^* y $k^*_{x,\tilde{x}}$ tales que $\psi(x, \delta) \leq k^*_{x,\tilde{x}} \psi(\tilde{x}, \delta)$ para todo $\delta \in (0, d^*)$, lo cual junto a (6.3) implica que para todo $\delta \in (0, \text{mín}\{d, d^*\})$,

$$\tilde{\psi}(x, \delta) \leq k_x \psi(x, \delta) \leq k_x k^*_{x,\tilde{x}} \psi(\tilde{x}, \delta) \quad \text{y} \quad \psi(x, \delta) \leq k^*_{x,\tilde{x}} \psi(\tilde{x}, \delta) \leq k^*_{x,\tilde{x}} k_{\tilde{x}} \tilde{\psi}(\tilde{x}, \delta). \tag{6.4}$$

Luego, $\tilde{\psi} \overset{M,M}{\preceq} \psi$ y $\psi \overset{M,M}{\preceq} \tilde{\psi}$, es decir, $\tilde{\psi} \overset{M,M}{\approx} \psi$.

(ii) De la primera desigualdad de (6.4) y de la segunda desigualdad de (6.3) se sigue inmediatamente que $\tilde{\psi} \overset{M,M}{\preceq} \tilde{\psi}$ y por lo tanto, $\tilde{\psi}$ es invariante sobre M. ∎

El siguiente resultado es análogo al Lema 6.2 para este caso de comparación de cotas de convergencia sobre conjuntos distintos.

Lema 6.17. *Sean* $M, N \subset \mathcal{X}$*,* $\psi \in \mathcal{F}_M$ *y* $\tilde{\psi} \in \mathcal{F}_N$*. Si* $\psi \overset{M,N}{\prec} \tilde{\psi}$ *entonces* $\forall\, \tilde{M} \subset M, \forall\, \tilde{N} \subset N$ *se tiene que* $\tilde{\psi} \overset{\tilde{N},\tilde{M}}{\not\preceq} \psi$*.*

Demostración. Por el contrarrecíproco. Supongamos que existen $\tilde{M} \subset M$ y $\tilde{N} \subset N$ tales que $\tilde{\psi} \overset{\tilde{N},\tilde{M}}{\preceq} \psi$. Entonces existen una constante $d > 0$ y $k : \tilde{N} \times \tilde{M} \to (0, \infty)$ tales que $\tilde{\psi}(\tilde{x}, \delta) \leq k(\tilde{x}, x)\, \psi(x, \delta)$ para todos $\tilde{x} \in \tilde{N}$, $x \in \tilde{M}$ y $\delta \in (0, d)$. Sean $x_0 \in \tilde{M}$ y $\tilde{x}_0 \in \tilde{N}$, entonces $\tilde{\psi}(\tilde{x}_0, \delta) \leq k(\tilde{x}_0, x_0)\, \psi(x_0, \delta)$ para todo $\delta \in (0, d)$. Luego, $\sup\limits_{\delta \in (0,d)} \frac{\tilde{\psi}(\tilde{x}_0,\delta)}{\psi(x_0,\delta)} \leq k(\tilde{x}_0, x_0) < \infty$. Entonces,

$$\begin{aligned}
\limsup_{\delta \to 0^+} \frac{\psi(x_0, \delta)}{\tilde{\psi}(\tilde{x}_0, \delta)} &\geq \liminf_{\delta \to 0^+} \frac{\psi(x_0, \delta)}{\tilde{\psi}(\tilde{x}_0, \delta)} \geq \inf_{\delta \in (0,d)} \frac{\psi(x_0, \delta)}{\tilde{\psi}(\tilde{x}_0, \delta)} \\
&= \left(\sup_{\delta \in (0,r)} \frac{\tilde{\psi}(\tilde{x}_0, \delta)}{\psi(x_0, \delta)} \right)^{-1} \geq \frac{1}{k(\tilde{x}_0, x_0)} > 0.
\end{aligned}$$

Por lo tanto, $\psi \overset{\{x_0\},\{\tilde{x}_0\}}{\not\prec} \tilde{\psi}$, de lo cual se sigue que $\psi \overset{M,N}{\not\prec} \tilde{\psi}$. ∎

6.2. Saturación global

A continuación procedemos a formalizar el concepto de saturación global de una familia de operadores de regularización $\{R_\alpha\}$ sobre un subconjunto de $\mathcal{X}$.

Definición 6.18. Sean $M_S \subset \mathcal{X}$ y $\psi_S \in \mathcal{U}_{M_S}(\mathcal{E}^{\text{tot}}_{\{R_\alpha\}})$. Decimos que ψ_S es "*función de saturación global de* $\{R_\alpha\}$ *sobre* M_S" si ψ_S satisface las siguientes tres condiciones:

(*S1*) Para todos $x^* \in \mathcal{X}$, $x^* \neq 0$, $x \in M_S$, $\limsup\limits_{\delta \to 0^+} \frac{\mathcal{E}^{\text{tot}}_{\{R_\alpha\}}(x^*,\delta)}{\psi_S(x,\delta)} > 0$.

(*S2*) ψ_S es invariante sobre M_S.

(*S3*) No existe ninguna cota superior de convergencia para el error total de $\{R_\alpha\}$ que sea extensión propia de ψ_S (en la variable x) y satisfaga (*S1*) y (*S2*), es decir, no existen $\tilde{M} \supsetneqq M_S$ y $\tilde{\psi} \in \mathcal{U}_{\tilde{M}}(\mathcal{E}^{\mathrm{tot}}_{\{R_\alpha\}})$ tales que $\tilde{\psi}$ satisfaga (*S1*) y (*S2*) con $M_S = \tilde{M}$ y $\psi_S = \tilde{\psi}$.

Observación 6.19. Notar que la condición (*S1*) implica que para todo $M \subset \mathcal{X}$ y para todo $\psi \in \mathcal{U}_M(\mathcal{E}^{\mathrm{tot}}_{\{R_\alpha\}})$, $\limsup_{\delta\to 0^+} \frac{\psi(x^*,\delta)}{\psi_S(x,\delta)} > 0$ para todos $x^* \in M$, $x \in M_S$ (esto es consecuencia inmediata de (*S1*) y del hecho que $\mathcal{E}^{\mathrm{tot}}_{\{R_\alpha\}} \overset{M}{\preceq} \psi \;\forall\, \psi \in \mathcal{U}_M(\mathcal{E}^{\mathrm{tot}}_{\{R_\alpha\}})$). Por lo tanto, no puede suceder que $\psi \overset{M,M_S}{\prec} \psi_S$. Por otro lado, si $\psi \in \mathcal{U}_M(\mathcal{E}^{\mathrm{tot}}_{\{R_\alpha\}})$ entonces no es necesariamente cierto que $\psi_S \overset{M_S,M}{\preceq} \psi$ aún cuando ψ sobre M sea comparable con ψ_S sobre M_S, pues en este caso puede suceder que $\liminf_{\delta\to 0^+} \frac{\psi(x,\delta)}{\psi_S(x_S,\delta)} = 0$ para algún $x \in M$ y algún $x_S \in M_S$ (lo cual implica que $\psi_S \overset{M_S,M}{\npreceq} \psi$), y todavía tener $\limsup_{\delta\to 0^+} \frac{\psi(x,\delta)}{\psi_S(x_S,\delta)} > 0$. Sin embargo, si ψ sobre M es comparable con ψ_S sobre M_S y existe $\lim_{\delta\to 0^+} \frac{\psi(x,\delta)}{\psi_S(x_S,\delta)}$ para todo $x \in M$ y para todo $x_S \in M_S$, entonces sí es cierto que $\psi_S \overset{M_S,M}{\preceq} \psi$.

Observar además que la condición (*S1*) puede reemplazarse por

$$\limsup_{\delta\to 0^+} \frac{\psi(x^*,\delta)}{\psi_S(x,\delta)} > 0 \quad \forall\, \psi \in \mathcal{U}_{\{x^*\}}(\mathcal{E}^{\mathrm{tot}}_{\{R_\alpha\}}), \forall\, x^* \in \mathcal{X}, x^* \neq 0, x \in M_S.$$

Esta noción de saturación global establece esencialmente que en ningún punto $x^* \in \mathcal{X}$, $x^* \neq 0$, puede existir una cota superior de convergencia para el error total de una familia de operadores de regularización que sea “estrictamente mejor” que la función de saturación ψ_S en cualquier punto de M_S.

Probaremos ahora que una función de saturación es siempre una cota superior óptima de convergencia.

Lema 6.20. *Sea $\psi_S \in \mathcal{U}_{M_S}(\mathcal{E}^{tot}_{\{R_\alpha\}})$. Si ψ_S satisface la condición (S1) sobre M_S, entonces $\psi_S \in \mathcal{U}^{opt}_{M_S}(\mathcal{E}^{tot}_{\{R_\alpha\}})$. En particular, si ψ_S es una función de saturación de $\{R_\alpha\}$ sobre M_S entonces $\psi_S \in \mathcal{U}^{opt}_{M_S}(\mathcal{E}^{tot}_{\{R_\alpha\}})$.*

Demostración. La condición (*S1*) implica en particular que $\limsup_{\delta\to 0^+} \frac{\mathcal{E}^{\mathrm{tot}}_{\{R_\alpha\}}(x,\delta)}{\psi_S(x,\delta)} > 0$ para todo $x \in M_S$. Puesto que además por definición $\psi_S \in \mathcal{U}_{M_S}(\mathcal{E}^{\mathrm{tot}}_{\{R_\alpha\}})$, se sigue que $\psi_S \in \mathcal{U}^{\mathrm{opt}}_{M_S}(\mathcal{E}^{\mathrm{tot}}_{\{R_\alpha\}})$, como queríamos probar. ∎

Un corolario inmediato de este lema es la equivalencia entre la función de saturación y el error total.

Corolario 6.21. *Si ψ_S es una función de saturación de $\{R_\alpha\}$ sobre M_S entonces $\psi_S \overset{M_S}{\approx} \mathcal{E}^{tot}_{\{R_\alpha\}}$. Más aún, se tiene la equivalencia más fuerte $\psi_S \overset{M_S,M_S}{\approx} \mathcal{E}^{tot}_{\{R_\alpha\}}$.*

Demostración. La primera parte del corolario es una consecuencia inmediata del Lema 6.20 y del Corolario 6.11 ***(i)***. La segunda parte se deduce de la primera y del hecho que $\psi_S \overset{M_S,M_S}{\approx} \psi_S$ via el Lema 6.16 ***(i)***. ∎

Observación 6.22. Una consecuencia de la primer parte de este corolario y del Lema 6.16 ***(ii)*** es que si ψ_S es función de saturación de $\{R_\alpha\}$ sobre M_S, entonces $\mathcal{E}^{\mathrm{tot}}_{\{R_\alpha\}} \overset{M_S,M_S}{\approx} \mathcal{E}^{\mathrm{tot}}_{\{R_\alpha\}}$, es decir, el error total debe ser invariante sobre M_S. Echaremos más luz sobre el particular en el Teorema 6.25.

Definición 6.23. Sean $M \subset \mathcal{X}$ y $\psi \in \mathcal{U}_{\mathcal{X}}(\mathcal{E}^{\mathrm{tot}}_{\{R_\alpha\}})$. Decimos que *"$M$ es óptimo para ψ"*, y lo denotamos $M \in \mathcal{O}(\psi)$, si satisface la siguiente condición:

(*C2*) Para todos $x \in M$, $x_c \in M^c$ ni $\psi \overset{\{x_c\},\{x\}}{\prec} \psi$ ni $\psi \overset{\{x_c\},\{x\}}{\approx} \psi$.

Que un conjunto M sea óptimo para ψ significa esencialmente que en cualquier punto del complemento de M, el orden de convergencia de ψ como función de δ para $\delta \to 0^+$ no puede ser mejor ni siquiera equivalente al orden de convergencia de ψ en cualquier punto de M; es decir, en cualquier punto fuera de M, el orden de convergencia de ψ debe ser "peor" que en cualquier punto de M. No obstante, veremos a continuación que esta condición de optimalidad impone una restricción bastante precisa. Como veremos más adelante (en el Teorema 6.25), es justamente esta restricción sobre el error total, junto con su invariancia, lo que permite caracterizar a las familias de operadores de regularización que poseen saturación.

Sean $\psi \in \mathcal{U}_{\mathcal{X}}(\mathcal{E}^{\mathrm{tot}}_{\{R_\alpha\}})$, $M \subset \mathcal{X}$ y consideremos las siguientes condiciones:

(*C1*) $\quad \psi \overset{M,M^c}{\prec} \psi$.

(*C3*) $\quad \psi \overset{M^c,M}{\nprec} \psi$ y $\psi \overset{M^c,M}{\not\approx} \psi$.

Observar que la condición (*C2*) (de conjunto óptimo) es estrictamente más fuerte que la condición (*C3*), y estrictamente más débil que la condición (*C1*). En efecto, si M es óptimo para ψ entonces para todos $x \in M$, $x_c \in M^c$ se tiene que $\psi \overset{\{x_c\},\{x\}}{\nprec} \psi$ y $\psi \overset{\{x_c\},\{x\}}{\not\approx} \psi$, de lo cual se sigue inmediatamente que $\psi \overset{M^c,M}{\nprec} \psi$ y $\psi \overset{M^c,M}{\not\approx} \psi$, es decir, se satisface (*C3*). Sin embargo, para que (*C3*) se satisfaga es suficiente que existan $x \in M$ y $x_c \in M^c$ tales que $\psi \overset{\{x_c\},\{x\}}{\nprec} \psi$ y $\psi \overset{\{x_c\},\{x\}}{\not\approx} \psi$, lo cual obviamente no implica la condición (*C2*). Por otra parte, si se satisface (*C1*) entonces se sigue del Lema 6.17 que para todos $x \in M$, $x_c \in M^c$ se tiene que $\psi \overset{\{x_c\},\{x\}}{\npreceq} \psi$ y en consecuencia, $\psi \overset{\{x_c\},\{x\}}{\nprec} \psi$ y $\psi \overset{\{x_c\},\{x\}}{\not\approx} \psi$ para todos $x \in M$, $x_c \in M^c$, es decir, se satisface (*C2*). No obstante, (*C2*) no implica (*C1*) pues puede ocurrir que M sea óptimo para ψ y que existan $x \in M$ y $x_c \in M^c$ tales que $\psi \overset{\{x\},\{x_c\}}{\npreceq} \psi$ y por lo tanto, $\psi \overset{M,M^c}{\nprec} \psi$.

Para caracterizar las familias de operadores de regularización que poseen función de saturación necesitaremos previamente del siguiente resultado.

Lema 6.24. *Supongamos que* $\{R_\alpha\}$ *posee función de saturación sobre* $M \subset \mathcal{X}$ *y para todos* $x \in M$, $x_c \in M^c$ *se tiene que* $\mathcal{E}^{tot}_{\{R_\alpha\}} \overset{\{x_c\},\{x\}}{\nprec} \mathcal{E}^{tot}_{\{R_\alpha\}}$. *Entonces* $\mathcal{E}^{tot}_{\{R_\alpha\}} \overset{\{x_c\},\{x\}}{\not\approx} \mathcal{E}^{tot}_{\{R_\alpha\}}$ *para todos* $x \in M$, $x_c \in M^c$.

Demostración. Puesto que $\{R_\alpha\}$ posee función de saturación sobre M, se deduce de la Observación 6.22 que $\mathcal{E}^{\mathrm{tot}}_{\{R_\alpha\}}$ es invariante sobre M. Supongamos que $\mathcal{E}^{\mathrm{tot}}_{\{R_\alpha\}} \overset{\{x_c\},\{x\}}{\nprec} \mathcal{E}^{\mathrm{tot}}_{\{R_\alpha\}}$ para todos $x \in M$, $x_c \in M^c$ y que existen $\tilde{x} \in M$, $\tilde{x}_c \in M^c$ tales que

$$\mathcal{E}^{\mathrm{tot}}_{\{R_\alpha\}} \overset{\{\tilde{x}_c\},\{\tilde{x}\}}{\approx} \mathcal{E}^{\mathrm{tot}}_{\{R_\alpha\}}. \tag{6.5}$$

Entonces,

$$\limsup_{\delta \to 0^+} \frac{\mathcal{E}^{\mathrm{tot}}_{\{R_\alpha\}}(\tilde{x}, \delta)}{\mathcal{E}^{\mathrm{tot}}_{\{R_\alpha\}}(\tilde{x}_c, \delta)} > 0. \tag{6.6}$$

Definimos $\tilde{M} \doteq M \cup \{\tilde{x}_c\}$ y

$$\tilde{\psi}(x,\delta) \doteq \begin{cases} \psi(x,\delta), & \text{si } x \in M \\ \mathcal{E}^{\text{tot}}_{\{R_\alpha\}}(x,\delta), & \text{si } x = \tilde{x}_c, \end{cases}$$

donde ψ es una función de saturación de $\{R_\alpha\}$ sobre M. A continuación probaremos que $\tilde{\psi}$ es saturación sobre $\tilde{M}$. Claramente, $\tilde{\psi}$ es cota superior de convergencia para el error total en $\tilde{M}$, i.e., $\tilde{\psi} \in \mathcal{U}_{\tilde{M}}(\mathcal{E}^{\text{tot}}_{\{R_\alpha\}})$ y como ψ es saturación sobre M, resulta que $\tilde{\psi}(x,\delta)$ satisface la condición (*S1*) para todo $x \in M$.

Veamos ahora que $\tilde{\psi}(\tilde{x}_c,\delta)$ también satisface (*S1*). Como $\tilde{x} \in M$ se tiene que

$$\limsup_{\delta\to 0^+} \frac{\mathcal{E}^{\text{tot}}_{\{R_\alpha\}}(x^*,\delta)}{\mathcal{E}^{\text{tot}}_{\{R_\alpha\}}(\tilde{x},\delta)} > 0 \quad \forall\, x^* \in \mathcal{X},\ x^* \neq 0. \tag{6.7}$$

Si $x^* \in M$, esta desigualdad se deduce del hecho que $\mathcal{E}^{\text{tot}}_{\{R_\alpha\}}$ es invariante sobre M y si $x^* \in M^c$, es consecuencia de que $\mathcal{E}^{\text{tot}}_{\{R_\alpha\}} \overset{\{x^*\},\{\tilde{x}\}}{\not\prec} \mathcal{E}^{\text{tot}}_{\{R_\alpha\}}$.

Entonces, para todo $x^* \in \mathcal{X}$, $x^* \neq 0$ se tiene que

$$\limsup_{\delta\to 0^+} \frac{\mathcal{E}^{\text{tot}}_{\{R_\alpha\}}(x^*,\delta)}{\tilde{\psi}(\tilde{x}_c,\delta)} = \limsup_{\delta\to 0^+} \frac{\mathcal{E}^{\text{tot}}_{\{R_\alpha\}}(x^*,\delta)}{\mathcal{E}^{\text{tot}}_{\{R_\alpha\}}(\tilde{x},\delta)} \frac{\mathcal{E}^{\text{tot}}_{\{R_\alpha\}}(\tilde{x},\delta)}{\tilde{\psi}(\tilde{x}_c,\delta)} > 0$$

en virtud de (6.6) y (6.7). Así, $\tilde{\psi}(x,\delta)$ satisface (*S1*) para todo $x \in \tilde{M}$.

Veamos ahora que $\tilde{\psi}$ satisface (*S2*) sobre $\tilde{M}$. Puesto que ψ es función de saturación de $\{R_\alpha\}$ sobre M, se tiene que $\tilde{\psi}$ es invariante sobre M. Entonces sólo resta ver que $\tilde{\psi} \overset{\{\tilde{x}_c\},M}{\approx} \tilde{\psi}$, es decir que $\mathcal{E}^{\text{tot}}_{\{R_\alpha\}} \overset{\{\tilde{x}_c\},M}{\approx} \psi$. Pero esto es una consecuencia inmediata de (6.5), del Corolario 6.21 que implica que $\psi \overset{M}{\approx} \mathcal{E}^{\text{tot}}_{\{R_\alpha\}}$ y del hecho que ψ es invariante sobre M.

Hemos demostrado así que $\tilde{\psi}$ es una extensión propia de ψ que satisface (*S1*) y (*S2*) sobre $\tilde{M}$, lo cual implica entonces que ψ no satisface la condición (*S3*). Esto contradice el hecho que ψ es función de saturación de $\{R_\alpha\}$ sobre M. Por lo tanto, para todos $x \in M$, $x_c \in M^c$ se tiene que $\mathcal{E}^{\text{tot}}_{\{R_\alpha\}} \overset{\{x_c\},\{x\}}{\not\approx} \mathcal{E}^{\text{tot}}_{\{R_\alpha\}}$ como queríamos probar. ∎

Teorema 6.25. *(Condición necesaria y suficiente para la existencia de saturación.) Una familia de operadores de regularización $\{R_\alpha\}$ tiene función de saturación si y sólo* *si existe $M \subset \mathcal{X}$ ($M \neq \{0\}, M \neq \emptyset$) tal que $\mathcal{E}^{tot}_{\{R_\alpha\}}$ es invariante sobre M y M es óptimo para $\mathcal{E}^{tot}_{\{R_\alpha\}}$. En tal caso, $\mathcal{E}^{tot}_M(x,\delta) \doteq \mathcal{E}^{tot}_{\{R_\alpha\}}(x,\delta)$ para $x \in M$ y $\delta > 0$ es función de saturación de $\{R_\alpha\}$ sobre M.*

Demostración. Supongamos que $\{R_\alpha\}$ tiene función de saturación ψ sobre M. Entonces se sigue de la Observación 6.22 que $\mathcal{E}^{\mathrm{tot}}_{\{R_\alpha\}}$ es invariante sobre M.

Veamos ahora que M es óptimo para $\mathcal{E}^{\mathrm{tot}}_{\{R_\alpha\}}$. Sean $x \in M$ y $x_c \in M^c$. Probemos primero que $\mathcal{E}^{\mathrm{tot}}_{\{R_\alpha\}} \overset{\{x_c\},\{x\}}{\not\preceq} \mathcal{E}^{\mathrm{tot}}_{\{R_\alpha\}}$. Como $\psi \in \mathcal{U}_M(\mathcal{E}^{\mathrm{tot}}_{\{R_\alpha\}})$ y $x \in M$, existen constantes positivas d y k_x tales que $\mathcal{E}^{\mathrm{tot}}_{\{R_\alpha\}}(x,\delta) \leq k_x \psi(x,\delta)$ para todo $\delta \in (0,d)$. Entonces

$$\limsup_{\delta \to 0^+} \frac{\mathcal{E}^{\mathrm{tot}}_{\{R_\alpha\}}(x_c,\delta)}{\mathcal{E}^{\mathrm{tot}}_{\{R_\alpha\}}(x,\delta)} \geq \limsup_{\delta \to 0^+} \frac{\mathcal{E}^{\mathrm{tot}}_{\{R_\alpha\}}(x_c,\delta)}{k_x \psi(x,\delta)} > 0,$$

donde la última desigualdad se sigue del hecho que ψ satisface la condición (*S1*) en M. Por lo tanto $\forall\, x \in M$, $\forall\, x_c \in M^c$, $\mathcal{E}^{\mathrm{tot}}_{\{R_\alpha\}} \overset{\{x_c\},\{x\}}{\not\preceq} \mathcal{E}^{\mathrm{tot}}_{\{R_\alpha\}}$. Esta condición junto con el hecho que $\mathcal{E}^{\mathrm{tot}}_{\{R_\alpha\}}$ es invariante sobre M implica que, además, $\forall\, x \in M$, $\forall\, x_c \in M^c$, $\mathcal{E}^{\mathrm{tot}}_{\{R_\alpha\}} \overset{\{x_c\},\{x\}}{\not\approx} \mathcal{E}^{\mathrm{tot}}_{\{R_\alpha\}}$, lo cual se sigue en virtud del Lema 6.24. Hemos probado así que M es óptimo para $\mathcal{E}^{\mathrm{tot}}_{\{R_\alpha\}}$.

Recíprocamente, supongamos que existe $M \subset \mathcal{X}$ ($M \neq \{0\}, M \neq \emptyset$) tal que $\mathcal{E}^{\mathrm{tot}}_{\{R_\alpha\}}$ es invariante sobre M y M es óptimo para $\mathcal{E}^{\mathrm{tot}}_{\{R_\alpha\}}$ y definamos $\mathcal{E}^{\mathrm{tot}}_M(x,\delta) \doteq \mathcal{E}^{\mathrm{tot}}_{\{R_\alpha\}}(x,\delta)$ para $x \in M$ y $\delta > 0$. Probaremos que $\mathcal{E}^{\mathrm{tot}}_M$ es función de saturación de $\{R_\alpha\}$ sobre M. Claramente, $\mathcal{E}^{\mathrm{tot}}_M \in \mathcal{U}_M(\mathcal{E}^{\mathrm{tot}}_{\{R_\alpha\}})$ y como por hipótesis $\mathcal{E}^{\mathrm{tot}}_M$ es invariante sobre M, sólo resta probar que $\mathcal{E}^{\mathrm{tot}}_M$ satisface las condiciones (*S1*) y (*S3*).

Para probar (*S1*), sean $x^* \in \mathcal{X}$, $x^* \neq 0$ y $x \in M$. Si $x^* \in M$, entonces la invariancia de $\mathcal{E}^{\mathrm{tot}}_M$ sobre M implica que $\mathcal{E}^{\mathrm{tot}}_M \overset{\{x^*\},\{x\}}{\approx} \mathcal{E}^{\mathrm{tot}}_M$ y por lo tanto

$$\limsup_{\delta \to 0^+} \frac{\mathcal{E}^{\mathrm{tot}}_{\{R_\alpha\}}(x^*,\delta)}{\mathcal{E}^{\mathrm{tot}}_M(x,\delta)} = \limsup_{\delta \to 0^+} \frac{\mathcal{E}^{\mathrm{tot}}_M(x^*,\delta)}{\mathcal{E}^{\mathrm{tot}}_M(x,\delta)} > 0. \tag{6.8}$$

Por otro lado, si $x^* \in M^c$, el límite anterior también es positivo debido a que $\mathcal{E}^{\mathrm{tot}}_{\{R_\alpha\}} \overset{\{x^*\},\{x\}}{\not\preceq} \mathcal{E}^{\mathrm{tot}}_M$ (condición (*C2*)) por ser M óptimo para $\mathcal{E}^{\mathrm{tot}}_{\{R_\alpha\}}$. Luego, $\mathcal{E}^{\mathrm{tot}}_M$ satisface la condición (*S1*).

Finalmente, supongamos que $\mathcal{E}_M^{\text{tot}}$ no satisface la condición (*S3*), es decir, que existen $\tilde{M} \supsetneqq M$ y $\tilde{\psi} \in \mathcal{U}_{\tilde{M}}(\mathcal{E}_{\{R_\alpha\}}^{\text{tot}})$ tales que $\tilde{\psi}$ es una extensión propia de $\mathcal{E}_M^{\text{tot}}$ que satisface las condiciones (*S1*) y (*S2*) sobre $\tilde{M}$. Sea $\tilde{x} \in \tilde{M} \setminus M$, entonces $\tilde{\psi} \overset{\{\tilde{x}\},M}{\approx} \tilde{\psi}$ por la invariancia de $\tilde{\psi}$ sobre $\tilde{M}$ y como $\tilde{\psi}$ coincide con $\mathcal{E}_M^{\text{tot}}$ sobre M, resulta que

$$\tilde{\psi} \overset{\{\tilde{x}\},M}{\approx} \mathcal{E}_M^{\text{tot}}. \tag{6.9}$$

Puesto que $\tilde{\psi} \in \mathcal{U}_{\tilde{M}}(\mathcal{E}_{\{R_\alpha\}}^{\text{tot}})$ satisface (*S1*) sobre $\tilde{M}$, el Lema 6.20 implica que $\tilde{\psi} \in \mathcal{U}_{\tilde{M}}^{opt}(\mathcal{E}_{\{R_\alpha\}}^{\text{tot}})$. Luego, en virtud del Corolario 6.11 ***(i)*** se tiene que $\mathcal{E}_{\{R_\alpha\}}^{\text{tot}} \overset{\tilde{M}}{\approx} \tilde{\psi}$. En particular, $\mathcal{E}_{\{R_\alpha\}}^{\text{tot}} \overset{\{\tilde{x}\}}{\approx} \tilde{\psi}$, lo cual junto a (6.9) implica que $\mathcal{E}_{\{R_\alpha\}}^{\text{tot}} \overset{\{\tilde{x}\},M}{\approx} \mathcal{E}_M^{\text{tot}}$, es decir, $\mathcal{E}_{\{R_\alpha\}}^{\text{tot}} \overset{\{\tilde{x}\},M}{\approx} \mathcal{E}_{\{R_\alpha\}}^{\text{tot}}$. Dado que $\tilde{x} \in M^c$, esta equivalencia contradice el hecho que M es óptimo para $\mathcal{E}_{\{R_\alpha\}}^{\text{tot}}$. Por lo tanto, $\mathcal{E}_M^{\text{tot}}$ debe satisfacer la condición (*S3*) y en consecuencia, es función de saturación de $\{R_\alpha\}$ sobre M. ■

Observación 6.26. A partir del teorema anterior conluímos entonces que una función de saturación de una familia de operadores de regularización es una cota superior óptima de convergencia para el error total, invariante y sin extensión propia.

Observar que una función de saturación es "doblemente optimal" en el sentido que si ψ es saturación sobre M, entonces M es óptimo para ψ y ψ es (cota superior) óptima para el error total de $\{R_\alpha\}$ sobre M. Además, M y ψ (módulo M, M equivalencia) están unívocamente determinados. En efecto, si se modifica el dominio M, este deja de ser óptimo para ψ y si se modifica la función ψ, aunque sea sólo en un punto de M, de manera tal que ψ no sea invariante sobre M, entonces deja de ser cota superior óptima. Supongamos que en el punto $x \in M$, se redefine $\psi(x,\delta) \doteq \tilde{\psi}(x,\delta)$, donde $\tilde{\psi} \in \mathcal{F}_M$. Si $\tilde{\psi} \overset{\{x\}}{\prec} \psi$, entonces ψ deja de ser cota superior para el error total de $\{R_\alpha\}$ en M y si $\psi \overset{\{x\}}{\prec} \tilde{\psi}$ entonces ψ es cota superior pero no es óptima.

6.3. Saturación de métodos de regularización espectrales

El objetivo de esta Sección es aplicar la teoría desarrollada anteriormente al caso de métodos de regularización espectrales (ver Definición 3.1). Además, mostraremos que esta teoría es consistente con resultados previamente existentes acerca de convergencia óptima de métodos de regularización espectrales.

Sea $\{g_\alpha\}_{\alpha\in(0,\alpha_0)}$ un MRE. Entonces satisface las hipótesis (*H1*)-(*H3*) dadas en la Definición 3.1. Cuando sea necesario, pediremos que $\{g_\alpha\}_{\alpha\in(0,\alpha_0)}$ satisfaga además la siguiente hipótesis:

(*H4*) $G_\alpha \doteq \|g_\alpha(\cdot)\|_\infty = O\left(\frac{1}{\sqrt{\alpha}}\right)$ cuando $\alpha \to 0^+$.

De aquí en adelante, denotaremos con $\{R_\alpha\}$ a la familia de operadores de regularización definida en (3.3) a partir del MRE $\{g_\alpha\}$.

6.3.1. Métodos espectrales con calificación clásica

En el siguiente resultado probaremos que si un MRE posee calificación clásica de orden μ_0 (ver Definición 3.6) entonces la función $\delta^{\frac{2\mu_0}{2\mu_0+1}}$ es cota superior de convergencia para el error total de la familia de operadores de regularización en el conjunto fuente $\mathcal{X}_{\mu_0} \doteq Ran((T^*T)^{\mu_0}) \setminus \{0\}$ (observar que, a diferencia de (2.27) en este capítulo definimos los conjuntos fuente de tal manera que no contienen al cero).

Lema 6.27. *Sean $\mathcal{X}$, $\mathcal{Y}$ espacios de Hilbert, $T \in \mathcal{L}(\mathcal{X},\mathcal{Y})$, $\{g_\alpha\}$ un MRE que satisface la hipótesis (H4) y posee calificación clásica de orden μ_0. Entonces $\psi_{\mu_0}(x,\delta) \doteq \delta^{\frac{2\mu_0}{2\mu_0+1}}$, para $x \in \mathcal{X}_{\mu_0} \doteq Ran((T^*T)^{\mu_0}) \setminus \{0\}$ y $\delta > 0$, es cota superior de convergencia para el error total de $\{R_\alpha\}_{\alpha\in(0,\alpha_0)}$ en $\mathcal{X}_{\mu_0}$, es decir, $\psi_{\mu_0} \in \mathcal{U}_{\mathcal{X}_{\mu_0}}(\mathcal{E}^{tot}_{\{R_\alpha\}})$.*

Demostración. Como $\{g_\alpha\}$ satisface la hipótesis (*H4*), se tiene que $G_\alpha = O\left(\frac{1}{\sqrt{\alpha}}\right)$ cuando $\alpha \to 0^+$, y entonces satisface (4.2). De esto y del hecho que $\{g_\alpha\}$ es un MRE que posee calificación clásica de orden μ_0, se deduce de la Proposición 4.1 que existe una regla de elección de parámetros *a-priori* $\alpha^* : \mathbb{R}^+ \to (0, \alpha_0)$ tal que el método de regularización $(\{R_\alpha\}, \alpha^*(\delta))$ es de orden óptimo en $\mathcal{X}_{\mu_0}$, es decir, para todo $x \in \mathcal{X}_{\mu_0}$ existe $k(x) > 0$ tal que para todo $\delta > 0$,

$$\sup_{y^\delta \in \overline{B_\delta(Tx)}} \left\| R_{\alpha^*(\delta)} y^\delta - x \right\| \leq k(x)\, \delta^{\frac{2\mu_0}{2\mu_0+1}}.$$

Luego, para todo $x \in \mathcal{X}_{\mu_0}$ y para todo $\delta > 0$

$$\inf_{\alpha \in (0,\alpha_0)} \sup_{y^\delta \in \overline{B_\delta(Tx)}} \left\| R_\alpha y^\delta - x \right\| \leq k(x)\, \delta^{\frac{2\mu_0}{2\mu_0+1}},$$

es decir, $\mathcal{E}^{\text{tot}}_{\{R_\alpha\}} \overset{\mathcal{X}_{\mu_0}}{\preceq} \psi_{\mu_0}$, lo cual implica que $\psi_{\mu_0} \in \mathcal{U}_{\mathcal{X}_{\mu_0}}(\mathcal{E}^{\text{tot}}_{\{R_\alpha\}})$. ∎

Teorema 6.28. *(Saturación para métodos de regularización espectrales con calificación clásica).*

Sean $\mathcal{X}$, $\mathcal{Y}$ espacios de Hilbert, $T \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$ tal que Ran(T) no es cerrado, $\{g_\alpha\}$ un MRE que satisface la hipótesis (H4). Supongamos además que:

i) *Existen constantes positivas $\lambda_1 \leq \|T\|^2$, γ_1, γ_2 y $c_1 > 1$ tales que*

a) $0 \leq r_\alpha(\lambda) \leq 1$, $\alpha > 0$, $0 \leq \lambda \leq \lambda_1$;

b) $r_\alpha(\lambda) \geq \gamma_1$, $0 \leq \lambda < \alpha \leq \lambda_1$;

c) $|r_\alpha(\lambda)|$ *es monótona creciente con respecto a α para* $\lambda \in (0, \|T\|^2]$;

d) $g_\alpha(c_1\alpha) \geq \frac{\gamma_2}{\alpha}$, $0 < c_1\alpha \leq \lambda_1$;

e) $g_\alpha(\lambda) \geq g_\alpha(\tilde{\lambda})$, *para* $0 < \alpha \leq \lambda \leq \tilde{\lambda} \leq \lambda_1$; *y*

ii) $\{g_\alpha\}_{\alpha \in (0,\alpha_0)}$, *donde* $\alpha_0 \doteq \min\{\lambda_1, \frac{\lambda_1}{c}\}$, *posee calificación clásica de orden* μ_0.

iii) *El orden μ_0 satisface la condición (3.17), es decir, existen constantes $\gamma, c > 0$ tales que*

$$\left(\frac{\lambda}{\alpha}\right)^{\mu_0} |r_\alpha(\lambda)| \geq \gamma, \quad \text{para todo } 0 < c\alpha \leq \lambda \leq \|T\|^2, \tag{6.10}$$

donde $r_\alpha(\lambda) \doteq 1 - \lambda g_\alpha(\lambda)$, es decir, como se definió en (3.7).

Entonces $\psi_{\mu_0}(x, \delta) \doteq \delta^{\frac{2\mu_0}{2\mu_0+1}}$ para $x \in \mathcal{X}_{\mu_0} \doteq$ Ran$((T^*T)^{\mu_0}) \setminus \{0\}$ y $\delta > 0$, es función de saturación de $\{R_\alpha\}_{\alpha \in (0,\alpha_0)}$ sobre $\mathcal{X}_{\mu_0}$.

Para probar este teorema necesitaremos de dos lemas previos.

Lema 6.29. *Sean $\mathcal{X}$, $\mathcal{Y}$ espacios de Hilbert, $T \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$, $\{g_\alpha\}_{\alpha>0}$ un MRE que satisface la hipótesis (H4). Supongamos que además se verifican las hipótesis **i.b)**, **i.c)**, **ii)** y **iii)** del Teorema 6.28. Entonces para todo $\alpha \in (0, \alpha_0)$ el operador $r_\alpha(T^*T)$ es invertible, donde $\alpha_0 \doteq \min\{\lambda_1, \frac{\lambda_1}{c}\}$.*

Demostración. Probaremos que para todo $\alpha \in (0, \alpha_0)$ y para todo $x \in \mathcal{X}$, la función $r_\alpha^{-2}(\lambda)$ es integrable con respecto a la medida $d\left\|E_\lambda x\right\|^2$. Sea $\alpha \in (0, \alpha_0)$ arbitrario pero fijo. Como $\alpha_0 \leq \lambda_1$, se sigue de la hipótesis ***i.b)*** que $r_\alpha(\lambda) \geq \gamma_1 > 0$ para todo $\lambda \in [0, \alpha)$. Entonces

$$\int_0^\alpha \frac{1}{r_\alpha^2(\lambda)}\, d\left\|E_\lambda x\right\|^2 \leq \frac{\|x\|^2}{\gamma_1^2} < +\infty. \tag{6.11}$$

Restaría probar que $\int_\alpha^{\|T\|^2+} \frac{1}{r_\alpha^2(\lambda)}\, d\left\|E_\lambda x\right\|^2 < +\infty$. Para ello consideraremos dos casos.

Caso I: $c \leq 1$. En este caso, para todo $\lambda \in [\alpha, \|T\|^2]$ se tiene que $\lambda \geq \alpha \geq c\,\alpha > 0$ y de (6.10) resulta que $|r_\alpha(\lambda)| \geq \gamma\left(\frac{\alpha}{\lambda}\right)^{\mu_0}$ para todo $\lambda \in [\alpha, \|T\|^2]$. Por lo tanto

$$\int_\alpha^{\|T\|^2+} \frac{1}{r_\alpha^2(\lambda)}\, d\left\|E_\lambda x\right\|^2 \leq \int_\alpha^{\|T\|^2+} \frac{\lambda^{2\mu_0}}{(\alpha^{\mu_0}\gamma)^2}\, d\left\|E_\lambda x\right\|^2 \leq \frac{\left\|(T^*T)^{\mu_0}x\right\|^2}{(\alpha^{\mu_0}\gamma)^2} < +\infty.$$

Caso II: $c > 1$. En este caso escribimos

$$\int_\alpha^{\|T\|^2+} \frac{1}{r_\alpha^2(\lambda)}\, d\left\|E_\lambda x\right\|^2 = \int_\alpha^{c\,\alpha} \frac{1}{r_\alpha^2(\lambda)}\, d\left\|E_\lambda x\right\|^2 + \int_{c\,\alpha}^{\|T\|^2+} \frac{1}{r_\alpha^2(\lambda)}\, d\left\|E_\lambda x\right\|^2. \tag{6.12}$$

Como en el caso anterior, en virtud de (6.10), la segunda integral en el lado derecho de (6.12) está acotada por arriba por $\frac{\|(T^*T)^{\mu_0}x\|^2}{(\alpha^{\mu_0}\gamma)^2} < +\infty$. Para la primera integral en el lado derecho de (6.12), en virtud de la hipótesis ***i.c)*** se tiene que

$$r_\alpha^2(\lambda) \geq r_{\alpha/c}^2(\lambda), \tag{6.13}$$

pues $\frac{\alpha}{c} < \alpha$. Por otro lado, nuevamente en virtud de (6.10), y dado que $0 < c(\frac{\alpha}{c}) \leq \lambda$ se tiene que

$$\left(\frac{\lambda}{\alpha/c}\right)^{2\mu_0} r_{\alpha/c}^2(\lambda) \geq \gamma^2. \tag{6.14}$$

De (6.13) y (6.14) se tiene que para todo $\lambda \in [\alpha, c\,\alpha]$, $r_\alpha^2(\lambda) \geq \gamma^2 \left(\frac{\alpha}{c\,\lambda}\right)^{2\mu_0}$. Así, para la primera integral en el lado derecho de (6.12) se tiene la acotación

$$\int_\alpha^{c\,\alpha} \frac{1}{r_\alpha^2(\lambda)}\, d\left\|E_\lambda x\right\|^2 \leq \int_\alpha^{c\,\alpha} \frac{c^{2\mu_0}}{\alpha^{2\mu_0}\gamma^2}\lambda^{2\mu_0}\, d\left\|E_\lambda x\right\|^2 \leq \frac{c^{2\mu_0}}{\alpha^{2\mu_0}\gamma^2}\left\|(T^*T)^{\mu_0}x\right\|^2 < \infty.$$

Por lo tanto, $r_\alpha(T^*T)$ es un operador invertible. ■

Lema 6.30. *Sean $\mathcal{X}$, $\mathcal{Y}$ espacios de Hilbert, $T \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$, $\{g_\alpha\}_{\alpha>0}$ un MRE que satisface la hipótesis (H4) y supongamos que se satisfacen además las hipótesis **i.b)**, **i.c)**, **ii)** y **iii)** del Teorema 6.28. Sea $\varphi : [0, \|T\|^2] \to \mathbb{R}^+$ continua, estrictamente creciente con $\varphi(0) = 0$. Si para algún $x^* \in \mathcal{X}$, $x^* \neq 0$ se tiene que $\mathcal{E}^{tot}_{\{R_\alpha\}}(x^*, \delta) = o(\varphi(\delta))$ cuando $\delta \to 0^+$, entonces existe una regla de elección del parámetro a-priori $\tilde{\alpha}(\delta)$ tal que*

$$\sup_{y^\delta \in \overline{B_\delta(Tx^*)}} \left\|R_{\tilde{\alpha}(\delta)} y^\delta - x^*\right\| = o(\varphi(\delta)) \quad \textit{cuando } \delta \to 0^+.$$

Lo mismo vale cambiando $o(\varphi(\delta))$ por $O(\varphi(\delta))$.

Demostración. Supongamos que existe $x^* \in \mathcal{X}$, $x^* \neq 0$ tal que $\mathcal{E}^{\mathrm{tot}}_{\{R_\alpha\}}(x^*, \delta) = o(\varphi(\delta))$ cuando $\delta \to 0^+$. Entonces por definición de $\mathcal{E}^{\mathrm{tot}}_{\{R_\alpha\}}$,

$$\lim_{\delta\to 0^+} \frac{\displaystyle\inf_{\alpha\in(0,\alpha_0)} \sup_{y^\delta \in \overline{B_\delta(Tx^*)}} \left\|R_\alpha y^\delta - x^*\right\|}{\varphi(\delta)} = \lim_{\delta\to 0^+} \inf_{\alpha\in(0,\alpha_0)} \frac{\displaystyle\sup_{y^\delta \in \overline{B_\delta(Tx^*)}} \left\|R_\alpha y^\delta - x^*\right\|}{\varphi(\delta)} = 0. \quad (6.15)$$

Para simplificar denotemos con

$$f(\alpha, \delta) \doteq \frac{\displaystyle\sup_{y^\delta \in \overline{B_\delta(Tx^*)}} \left\|R_\alpha y^\delta - x^*\right\|}{\varphi(\delta)} \quad \text{y} \quad h(\delta) \doteq \inf_{\alpha\in(0,\alpha_0)} f(\alpha, \delta).$$

Entonces $h(\delta) > 0$ para todo $\delta \in (0, \infty)$ y (6.15) puede escribirse simplemente como $\lim_{\delta\to 0^+} h(\delta) = 0$. A continuación, para $n \in \mathbb{N}$ definimos

$$\delta_n \doteq \sup\left\{\delta > 0 : h(\delta) \leq \frac{1}{n}\right\},\ \forall\, n \in \mathbb{N}.$$

Claramente, $\delta_n \downarrow 0$ y $h(\delta) = \inf_{\alpha\in(0,\alpha_0)} f(\alpha,\delta) \leq \frac{1}{n}$ para todo $\delta \in (0,\delta_n]$ para todo $n \in \mathbb{N}$. Luego, existe $\alpha_n = \alpha_n(\delta_n) \in (0,\alpha_0)$ tal que

$$f(\alpha_n,\delta) \leq \frac{2}{n} \quad \forall\, \delta \in (0,\delta_n],\ \forall\, n \in \mathbb{N}. \tag{6.16}$$

Definimos entonces $\alpha(\delta) \doteq \alpha_n$ para todo $\delta \in (\delta_{n+1},\delta_n]$ para todo $n \in \mathbb{N}$. De esto, del hecho que $\delta_n \downarrow 0$ y de (6.16) se deduce que $\lim_{\delta\to 0^+} f(\alpha(\delta),\delta) = \lim_{n\to+\infty} f(\alpha_n,\delta_n) = 0$.

No podemos garantizar la existencia del límite de $\alpha(\delta)$ cuando $\delta \to 0^+$. Sin embargo, veremos a continuación que $\alpha(\delta)$ se puede reemplazar por una función $\tilde{\alpha} : \mathbb{R}^+ \to (0,\alpha_0)$ tal que $\lim_{\delta\to 0^+} \tilde{\alpha}(\delta) = 0$ (i.e, de modo que $\tilde{\alpha}(\delta)$ sea una regla de elección del parámetro admisible) manteniendo $\lim_{\delta\to 0^+} f(\tilde{\alpha}(\delta),\delta) = 0$. En efecto, como $\{\alpha_n\}_{n\in\mathbb{N}} \subset (0,\alpha_0)$ es una sucesión acotada de números reales, contiene una subsucesión convergente $\{\alpha_{n_k}\}_{k\in\mathbb{N}}$, con $\alpha_{n_k} \to \alpha^*$ para $k \to +\infty$, $\alpha^* \in [0,\alpha_0]$. Definimos $\tilde{\alpha}(\delta) \doteq \alpha_{n_k}$ para todo $\delta \in (\delta_{n_{k+1}},\delta_{n_k}]$, para todo $k \in \mathbb{N}$. Entonces,

$$\lim_{\delta\to 0^+} \tilde{\alpha}(\delta) = \lim_{k\to+\infty} \alpha_{n_k} = \alpha^*. \tag{6.17}$$

Puesto que $\{\alpha_{n_k}\}_{k\in\mathbb{N}}$ y $\{\delta_{n_k}\}_{k\in\mathbb{N}}$ son subsucesiones de $\{\alpha_n\}_{n\in\mathbb{N}}$ y $\{\delta_n\}_{n\in\mathbb{N}}$, $\lim_{\delta\to 0^+} f(\tilde{\alpha}(\delta),\delta) = \lim_{k\to+\infty} f(\alpha_{n_k},\delta_{n_k}) = 0$. Entonces, por definición de f,

$$\lim_{\delta\to 0^+} \frac{\sup\limits_{y^\delta\in\overline{B_\delta(Tx^*)}} \left\| R_{\tilde{\alpha}(\delta)} y^\delta - x^* \right\|}{\varphi(\delta)} = 0,$$

es decir,

$$\sup_{y^\delta\in\overline{B_\delta(Tx^*)}} \left\| R_{\tilde{\alpha}(\delta)} y^\delta - x^* \right\| = o(\varphi(\delta)) \quad \text{cuando } \delta \to 0^+. \tag{6.18}$$

Sólo resta probar que $\alpha^* = 0$. Si $\alpha^* > 0$, entonces se sigue de (6.17) que existe $\delta_0 > 0$ tal que $\tilde{\alpha}(\delta) > \frac{\alpha^*}{2}$ para todo $\delta \in (0,\delta_0)$. La hipótesis ***i.c)*** del Teorema 6.28 implica entonces que para todo $\delta \in (0,\delta_0)$, $\left| r_{\tilde{\alpha}(\delta)}(\lambda) \right| \geq \left| r_{\frac{\alpha^*}{2}}(\lambda) \right|$ para todo $\lambda \in (0,\|T\|^2]$. Luego, para todo $\delta \in (0,\delta_0)$,

$$\left\| r_{\tilde{\alpha}(\delta)}(T^*T)x^* \right\|^2 \;=\; \int_0^{\|T\|^2+} r^2_{\tilde{\alpha}(\delta)}(\lambda)\, d\left\| E_\lambda x^* \right\|^2$$

$$\begin{aligned} &\geq \int_0^{\|T\|^2+} r^2_{\frac{\alpha^*}{2}}(\lambda)\, d\, \|E_\lambda x^*\|^2 \\ &= \left\| r_{\frac{\alpha^*}{2}}(T^*T)x^* \right\|^2. \end{aligned}$$

Ahora bien, para todo $\delta \in (0, \delta_0)$,

$$\begin{aligned} \sup_{y^\delta \in B_\delta(Tx^*)} \left\| R_{\tilde{\alpha}(\delta)} y^\delta - x^* \right\| &\geq \left\| R_{\tilde{\alpha}(\delta)} T x^* - x^* \right\| = \left\| (I - g_{\tilde{\alpha}(\delta)}(T^*T)T^*T)x^* \right\| \\ &= \left\| r_{\tilde{\alpha}(\delta)}(T^*T)x^* \right\| \geq \left\| r_{\frac{\alpha^*}{2}}(T^*T)x^* \right\|. \end{aligned}$$

Tomando límite cuando $\delta \to 0^+$, resulta que

$$\lim_{\delta \to 0^+} \sup_{y^\delta \in B_\delta(Tx^*)} \left\| R_{\tilde{\alpha}(\delta)} y^\delta - x^* \right\| \geq \left\| r_{\frac{\alpha^*}{2}}(T^*T)x^* \right\|,$$

lo cual junto a (6.18) implica que $\left\| r_{\frac{\alpha^*}{2}}(T^*T)x^* \right\| = 0$. Pero como $\frac{\alpha^*}{2} < \alpha_0$, se sigue del Lema 6.29 que $r_{\frac{\alpha^*}{2}}(T^*T)$ es invertible y por lo tanto $x^* = 0$, lo cual es una contradicción pues x^* era no nulo. Por lo tanto α^* debe ser cero, como queríamos probar.

Procedemos ahora a probar la segunda parte del Lema. Supongamos que existe $x^* \in \mathcal{X}$, $x^* \neq 0$ tal que $\mathcal{E}^{\text{tot}}_{\{R_\alpha\}}(x^*, \delta) = O(\varphi(\delta))$ cuando $\delta \to 0^+$. Entonces existen constantes positivas k y d tales que $\inf\limits_{\alpha \in (0, \alpha_0)} f(\alpha, \delta) \leq k$ para todo $\delta \in (0, d)$. Sea $\{\delta_n\}_{n\in\mathbb{N}} \subset (0, d)$ tal que $\delta_n \downarrow 0$ y $\alpha_n = \alpha_n(\delta_n) \in (0, \alpha_0)$ tal que

$$f(\alpha_n, \delta) \leq k + \delta_n, \ \forall\, \delta \in (0, d), \ \forall\, n \in \mathbb{N}.$$

Definimos (como lo hicimos anteriormente para el caso "o") $\alpha(\delta) \doteq \alpha_n$ para todo $\delta \in (\delta_{n+1}, \delta_n]$ para todo $n \in \mathbb{N}$. Puesto que $\delta_n \downarrow 0$ se sigue que $f(\alpha(\delta), \delta) \leq k + \delta_1$ para todo $\delta \in (0, d)$ y por lo tanto

$$\sup_{y^\delta \in B_\delta(Tx^*)} \left\| R_{\alpha(\delta)} y^\delta - x^* \right\| = O(\varphi(\delta)) \quad \text{cuando } \delta \to 0^+. \tag{6.19}$$

Exactamente de la misma manera en que se procedió en la primera parte de la demostración, definiendo la función $\tilde{\alpha}(\delta)$ (a partir de una subsucesión convergente de $\{\alpha_n\}_{n\in\mathbb{N}}$), se demuestra (6.19) con $\tilde{\alpha}(\delta)$ en lugar de $\alpha(\delta)$. Por último, y también de igual forma, se prueba que $\tilde{\alpha}(\delta)$ converge a cero cuando $\delta \to 0^+$, es decir que $\tilde{\alpha}(\delta)$ es una regla de elección del parámetro admisible. Puesto que los pasos son idénticos a los del caso anterior, no escribimos los detalles aquí. ■

Estamos ahora en condiciones de probar el Teorema 6.28.

Demostración del Teorema 6.28. Probaremos que $\psi_{\mu_0}(x,\delta) \doteq \delta^{\frac{2\mu_0}{2\mu_0+1}}$ para $x \in \mathcal{X}_{\mu_0}$ y $\delta > 0$, es función de saturación de $\{R_\alpha\}_{\alpha\in(0,\alpha_0)}$ sobre $\mathcal{X}_{\mu_0}$.

En virtud del Lema 6.27 se tiene que $\psi_{\mu_0} \in \mathcal{U}_{\mathcal{X}_{\mu_0}}(\mathcal{E}^{\text{tot}}_{\{R_\alpha\}})$. Veamos ahora que ψ_{μ_0} satisface la condición (*S1*) (ver Definición 6.18) de saturación sobre $\mathcal{X}_{\mu_0}$. Supongamos que no, es decir, que existen $x^* \in \mathcal{X}$, $x^* \neq 0$ y $x \in \mathcal{X}_{\mu_0}$ tales que $\limsup\limits_{\delta\to 0^+} \frac{\mathcal{E}^{\text{tot}}_{\{R_\alpha\}}(x^*,\delta)}{\psi(x,\delta)} = 0$. Entonces $\mathcal{E}^{\text{tot}}_{\{R_\alpha\}}(x^*,\delta) = o\left(\delta^{\frac{2\mu_0}{2\mu_0+1}}\right)$ cuando $\delta \to 0^+$ y por el Lema 6.30 se sigue que existe una regla de elección del parámetro *a-priori* $\alpha(\delta)$ tal que

$$\sup_{y^\delta \in \overline{B_\delta(Tx^*)}} \left\| R_{\alpha(\delta)} y^\delta - x^* \right\| = o(\delta^{\frac{2\mu_0}{2\mu_0+1}}) \quad \text{cuando } \delta \to 0^+. \tag{6.20}$$

Notar que la hipótesis (*H4*) implica que existe una constante positiva finita β tal que $\sqrt{\lambda}\,|g_\alpha(\lambda)| \leq \frac{\beta}{\sqrt{\alpha}}$, para todo $\alpha \in (0,\alpha_0)$ y para todo $\lambda \in [0, \|T\|^2]$ y entonces se satisface la hipótesis ***f)*** del Teorema 3.18. De esto junto con (6.20), el hecho que *Ran*(T) no es cerrado y dado que $\{g_\alpha\}$ es un MRE que satisface la hipótesis (*H4*) además de ***i)-iii)***, se sigue del Teorema 3.19 que $x^* = 0$, lo cual contradice que x^* era no nulo. Luego, ψ_{μ_0} satisface la condición (*S1*) sobre $\mathcal{X}_{\mu_0}$.

Como ψ_{μ_0} es independiente de x, se tiene además que ψ_{μ_0} es (trivialmente) invariante sobre $\mathcal{X}_{\mu_0}$, es decir, satisface la condición (*S2*).

Solo resta probar que ψ_{μ_0} satisface la condición (*S3*), es decir que el conjunto $\mathcal{X}_{\mu_0}$ es óptimo para ψ_{μ_0}. Supongamos que no. Entonces existen $M \supsetneqq \mathcal{X}_{\mu_0}$ y $\tilde{\psi} \in \mathcal{U}_M(\mathcal{E}^{\text{tot}}_{\{R_\alpha\}})$ tales que $\tilde{\psi}\,|_{\mathcal{X}_{\mu_0}} = \psi_{\mu_0}$ y $\tilde{\psi}$ satisface (*S1*) y (*S2*) sobre M. Sea $x^* \in M \setminus \mathcal{X}_{\mu_0}$, $x^* \neq 0$.

Como $\tilde{\psi} \in \mathcal{U}_M(\mathcal{E}^{\text{tot}}_{\{R_\alpha\}})$ se tiene que

$$\mathcal{E}^{\text{tot}}_{\{R_\alpha\}} \overset{\{x^*\}}{\preceq} \tilde{\psi}. \tag{6.21}$$

Puesto que $\tilde{\psi}$ es invariante sobre M, se tiene que $\tilde{\psi} \overset{\{x^*\},\mathcal{X}_{\mu_0}}{\preceq} \tilde{\psi}$ y como $\tilde{\psi}$ coincide con ψ_{μ_0} sobre $\mathcal{X}_{\mu_0}$, resulta que $\tilde{\psi} \overset{\{x^*\},\mathcal{X}_{\mu_0}}{\preceq} \psi_{\mu_0}$. Esto junto a (6.21) implica que $\mathcal{E}^{\text{tot}}_{\{R_\alpha\}} \overset{\{x^*\},\mathcal{X}_{\mu_0}}{\preceq} \psi_{\mu_0}$ y por lo tanto $\mathcal{E}^{\text{tot}}_{\{R_\alpha\}}(x^*, \delta) = O\left(\delta^{\frac{2\mu_0}{2\mu_0+1}}\right)$ cuando $\delta \to 0^+$. El Lema 6.30 implica entonces que existe una regla de elección del parámetro *a-priori* $\alpha(\delta)$ tal que

$$\sup_{y^\delta \in \overline{B_\delta(Tx^*)}} \left\| R_{\alpha(\delta)} y^\delta - x^* \right\| = O\left(\delta^{\frac{2\mu_0}{2\mu_0+1}}\right) \quad \text{cuando } \delta \to 0^+.$$

Como $\mu_0 < +\infty$ se sigue que $x^* \in Ran(T^*T)^{\mu_0}$ (ver [49], Corolario 2.6) y puesto que $x^* \neq 0$, se tiene que $x^* \in \mathcal{X}_{\mu_0}$, lo cual contradice que $x^* \in M \setminus \mathcal{X}_{\mu_0}$. Así, ψ_{μ_0} satisface la condición (*S3*) y ψ_{μ_0} es función de saturación de $\{R_\alpha\}$ sobre $\mathcal{X}_{\mu_0}$, como queríamos probar. ∎

6.3.2. Métodos espectrales con calificación máxima

A continuación veremos que, bajo hipótesis generales, es posible también caracterizar la saturación de los métodos de regularización espectrales que poseen calificación máxima (ver Definición 3.11). Para ello, necesitaremos previamente de la siguiente definición.

Definición 6.31. Sea $\rho : [0, a) \to (0, +\infty)$ una función continua no decreciente tal que $\lim\limits_{t\to 0^+} \rho(t) = 0$ y $\beta \geq 0$. Diremos que "*ρ es de tipo superior local β*" si existe una constante positiva d tal que $\rho(t) \leq d\, s^{-\beta} \rho(st)$ para todos $s \in (0, 1]$, $t \in (0, a]$.

Si una función es de tipo superior finito, se dice que satisface una condición Δ_2.

Teorema 6.32. *(Saturación para métodos de regularización espectrales con calificación máxima.)*

Sean $\mathcal{X}$, $\mathcal{Y}$ espacios de Hilbert, $T : \mathcal{X} \to \mathcal{Y}$ un operador lineal y compacto, $\{g_\alpha\}$ un MRE que satisface la hipótesis (H4). Supongamos que se satisfacen además las siguientes hipótesis:

i) *Existen $\{\tilde{\lambda}_k\}_{k=1}^{\infty} \subset \sigma_p(TT^*)$ y $c \geq 1$ tales que $\tilde{\lambda}_k \downarrow 0$ y $\frac{\tilde{\lambda}_k}{\tilde{\lambda}_{k+1}} \leq c$ para todo $k \in \mathbb{N}$;*

ii) *Existen constantes positivas $\lambda_1 \leq \|T\|^2$, γ_1, γ_2 y $c_1 > 1$ tales que*

a) $0 \leq r_\alpha(\lambda) \leq 1$, $\alpha > 0$, $0 \leq \lambda \leq \lambda_1$;

b) $r_\alpha(\lambda) \geq \gamma_1$, $0 \leq \lambda < \alpha \leq \lambda_1$;

c) $|r_\alpha(\lambda)|$ *es monótona creciente con respecto a α para $\lambda \in (0, \|T\|^2]$;*

d) $g_\alpha(c_1\alpha) \geq \frac{\gamma_2}{\alpha}$, $0 < c_1\alpha \leq \lambda_1$ *y*

e) $g_\alpha(\lambda) \geq g_\alpha(\tilde{\lambda})$, *para* $0 < \alpha \leq \lambda \leq \tilde{\lambda} \leq \lambda_1$.

iii) *Existe $\rho : (0, \|T\|^2] \to (0, +\infty)$, estrictamente creciente y de tipo superior β, para algún $\beta \geq 0$, tal que*

a) *existen constantes positivas a y k tales que*

$$\frac{\rho(\lambda)\,|r_\alpha(\lambda)|}{\rho(\alpha)} \geq a, \quad \text{para todos } \alpha, \lambda \text{ tales que } 0 < k\,\alpha \leq \lambda \leq \|T\|^2 .$$

b) *el MRE $\{g_\alpha\}$ posee calificación máxima ρ.*

iv) *Existe una constante $b > 0$ tal que*

$$\sup_{\lambda \in (0, \|T\|^2]} \sqrt{\lambda}\, |g_\alpha(\lambda)| \geq \frac{b}{\sqrt{\alpha}} \quad \text{para todo } \alpha \in (0, \alpha_0).$$

v) *Para todo $\alpha \in (0, \alpha_0)$ la función $\lambda \to |r_\alpha(\lambda)|^2$, $\lambda \in (0, \|T\|^2]$ es convexa.*

Sea $\Theta(t) \doteq \sqrt{t}\rho(t)$ para $t \in (0, \|T\|^2]$. Entonces $\psi(x, \delta) \doteq (\rho \circ \Theta^{-1})(\delta)$ para $x \in \mathcal{X}^\rho \doteq Ran(\rho(T^*T)) \setminus \{0\}$ y $\delta \in (0, \Theta(\alpha_0))$, es función de saturación de $\{R_\alpha\}_{\alpha \in (0,\alpha_0)}$ sobre $\mathcal{X}^\rho$.

Para probar este teorema necesitaremos previamente de dos resultados recíprocos que establecemos en los siguientes dos Lemas.

Lema 6.33. *Sean $\mathcal{X}$, $\mathcal{Y}$ espacios de Hilbert, $T : \mathcal{X} \to \mathcal{Y}$ un operador lineal y continuo, $\{g_\alpha\}_{\alpha \in (0,\alpha_0)}$ un MRE y $\rho : (0, \|T\|^2] \to \mathbb{R}^+$ una función continua estrictamente creciente* *que satisface la hipótesis* ***iii.b)*** *del Teorema 6.32. Si para algún* $x \in \mathcal{X}$*,* $\|R_\alpha Tx - x\| = O(\rho(\alpha))$ *cuando* $\alpha \to 0^+$*, entonces* $x \in Ran(\rho(T^*T))$*.*

Demostración. De la hipótesis ***iii.b)*** del Teorema 6.32 se sigue que

$$\|R_\alpha Tx - x\|^2 = \int_0^{\|T\|^2+} r_\alpha^2(\lambda)\, d\,\|E_\lambda x\|^2 \geq a^2\, \rho^2(\alpha) \int_{k\,\alpha}^{\|T\|^2+} \rho^{-2}(\lambda)\, d\,\|E_\lambda x\|^2 . \quad (6.22)$$

Puesto que $\|R_\alpha Tx - x\| = O(\rho(\alpha))$ cuando $\alpha \to 0^+$, se sigue entonces que existen constantes $C > 0$ y α^*, $0 < \alpha^* \leq \alpha_0$ tales que

$$\int_{k\,\alpha}^{\|T\|^2+} \rho^{-2}(\lambda)\, d\,\|E_\lambda x\|^2 \leq \frac{\|R_\alpha Tx - x\|^2}{a^2 \rho^2(\alpha)} \leq \frac{C^2}{a^2} \quad \text{para todo } \alpha \in (0, \alpha^*).$$

Tomando límite para $\alpha \to 0^+$ resulta que $\int_0^{\|T\|^2+} \rho^{-2}(\lambda)\, d\,\|E_\lambda x\|^2 < +\infty$, de lo cual se sigue que $w \doteq \int_0^{\|T\|^2+} \rho^{-1}(\lambda)\, dE_\lambda x \in \mathcal{X}$. Luego,

$$\rho(T^*T)w = \int_0^{\|T\|^2+} \rho(\lambda)\rho^{-1}(\lambda)\, dE_\lambda x = x$$

y por lo tanto $x \in Ran(\rho(T^*T))$. ■

Lema 6.34. *Bajo las mismas hipótesis del Teorema 6.32, si para algún* $x \in \mathcal{X}$ *se tiene que*

$$\sup_{y^\delta \in \overline{B_\delta(Tx)}} \; \inf_{\alpha \in (0,\alpha_0)} \left\| R_\alpha y^\delta - x \right\| = O(\rho(\Theta^{-1}(\delta))) \quad \textit{cuando } \delta \to 0^+, \quad (6.23)$$

entonces $x \in Ran(\rho(T^*T))$*.*

Demostración. Sin pérdida de generalidad supondremos que $\alpha_0 \doteq \text{mín}\{\frac{\lambda_1}{c_1}, \frac{\lambda_1}{k}\}$ y que $x \neq 0$ (si $x = 0$ el resultado es trivial).

Sea $\bar{\lambda} \in \sigma_p(TT^*)$ tal que $0 < c_1\, \bar{\lambda} \leq \lambda_1$ (la compacidad de T garantiza la existencia de tal $\bar{\lambda}$), y definimos

$$\bar{\delta} \doteq \frac{\bar{\lambda}^{1/2}}{\gamma_2} \|R_{\bar{\lambda}} Tx - x\| .$$

Entonces, claramente la ecuación

$$\|R_\alpha Tx - x\|^2 = \frac{(\gamma_2\, \bar{\delta})^2}{\alpha} \quad (6.24)$$

tiene como solución $\alpha = \bar{\lambda}$. Más aún, de la hipótesis ***ii.c)*** y dado que $x \neq 0$, se sigue que $\alpha = \bar{\lambda}$ es la única solución de (6.24). Notar además que $\bar{\delta} \to 0^+$ si y sólo si $\bar{\lambda} \to 0^+$.

Ahora dados $\delta > 0$ y $\bar{\lambda} \in \sigma_p(TT^*)$ tal que $c_1\bar{\lambda} \leq \lambda_1$, definimos

$$\bar{y}^{\,\delta} \doteq Tx - \delta G_{\bar{\lambda}} z, \quad \forall\, \delta > 0, \tag{6.25}$$

con

$$z \doteq \begin{cases} Tx \left\| G_{\bar{\lambda}} Tx \right\|^{-1}, & \text{si } G_{\bar{\lambda}} Tx \neq 0 \\ \text{arbitrario con } \| G_{\bar{\lambda}} z \| = 1, & \text{en otro caso} \end{cases}$$

donde $G_{\bar{\lambda}} \doteq F_{c_1 \bar{\lambda}} - F_{\bar{\lambda}}$ y $\{F_\lambda\}$ es la familia espectral asociada al operador autoadjunto TT^*. Como $\bar{\lambda} \in \sigma_p(TT^*)$, esta definición tiene sentido pues $G_{\bar{\lambda}}$ no es el operador nulo. Notar que $\left\| \bar{y}^{\,\delta} - Tx \right\| = \delta$, lo cual implica que $\bar{y}^{\,\delta} \in \overline{B_\delta(Tx)}$.

Ahora, de (3.3), (6.25) y (1.35) se sigue que para todo $\alpha \in (0, \alpha_0)$ y para todo $\delta > 0$,

$$\begin{aligned} \left\langle R_\alpha Tx - x \ , R_\alpha(\bar{y}^{\,\delta} - Tx) \right\rangle &= \langle g_\alpha(T^*T)T^*Tx - x, -g_\alpha(T^*T)T^*\,\delta G_{\bar{\lambda}} z \rangle \\ &= \delta \,\langle g_\alpha(T^*T)T^*Tx - x, -T^* g_\alpha(TT^*) G_{\bar{\lambda}} z \rangle \\ &= \delta \,\langle T g_\alpha(T^*T)T^*Tx - Tx, -g_\alpha(TT^*) G_{\bar{\lambda}} z \rangle \\ &= \delta \,\langle (TT^* g_\alpha(TT^*) - I)Tx, -g_\alpha(TT^*) G_{\bar{\lambda}} z \rangle \\ &= \delta \,\langle -r_\alpha(TT^*)Tx, -g_\alpha(TT^*) G_{\bar{\lambda}} z \rangle \\ &= \delta \int_0^{\|T\|^2+} r_\alpha(\lambda) g_\alpha(\lambda)\, d\,\langle F_\lambda Tx, G_{\bar{\lambda}} z \rangle\,. \end{aligned} \tag{6.26}$$

Puesto que $c_1\bar{\lambda} \leq \lambda_1$, se sigue de la hipótesis ***ii.a)*** que tanto $g_\alpha(\lambda)$ como $r_\alpha(\lambda)$ son no negativas para todo $\lambda \in [0, c_1\bar{\lambda}]$. Por otro lado, de las definiciones de $G_{\bar{\lambda}}$ y z se sigue inmediatamente que la función $h(\lambda) \doteq \langle F_\lambda Tx, G_{\bar{\lambda}} z \rangle$ es real y no decreciente y por lo tanto

$$\int_0^{c_1\bar{\lambda}+} r_\alpha(\lambda) g_\alpha(\lambda)\, d\,\langle F_\lambda Tx, G_{\bar{\lambda}} z \rangle \geq 0. \tag{6.27}$$

Por otra parte, como $h(\lambda) = \langle Tx, F_\lambda G_{\bar{\lambda}} z \rangle$ y $F_\lambda G_{\bar{\lambda}} = G_{\bar{\lambda}}$ para todo $\lambda \geq c_1\bar{\lambda}$, resulta que $h(\lambda)$ es constante para todo $\lambda \geq c_1\bar{\lambda}$ y entonces

$$\int_{c_1\bar{\lambda}+}^{\|T\|^2+} r_\alpha(\lambda) g_\alpha(\lambda)\, d\,\langle F_\lambda Tx, G_{\bar{\lambda}} z \rangle = 0. \tag{6.28}$$

De (6.27) y (6.28) se tiene que

$$\int_0^{\|T\|^2+} r_\alpha(\lambda) g_\alpha(\lambda)\, d\left\langle F_\lambda Tx, G_{\bar\lambda} z\right\rangle \geq 0,$$

lo cual, en virtud de (6.26), implica que

$$\left\langle R_\alpha Tx - x, R_\alpha(\bar y^{\,\delta} - Tx)\right\rangle \geq 0. \tag{6.29}$$

Utilizando nuevamente (3.3) y (6.25) junto con (6.29) se sigue entonces que para todo $\alpha \in (0, \alpha_0)$, para todo $\bar\lambda \in \sigma_p(TT^*)$ tal que $c_1\bar\lambda \leq \lambda_1$ y para todo $\delta > 0$,

$$\begin{aligned}
\left\|R_\alpha \bar y^{\,\delta} - x\right\|^2 &= \|R_\alpha Tx - x\|^2 + \left\|R_\alpha(\bar y^{\,\delta} - Tx)\right\|^2 + 2\left\langle R_\alpha Tx - x, R_\alpha(\bar y^{\,\delta} - Tx)\right\rangle \\
&= \|R_\alpha Tx - x\|^2 + \delta^2 \left\|g_\alpha(T^*T)T^* G_{\bar\lambda} z\right\|^2 + 2\left\langle R_\alpha Tx - x, R_\alpha(\bar y^{\,\delta} - Tx)\right\rangle \\
&\geq \|R_\alpha Tx - x\|^2 + \delta^2 \left\|g_\alpha(T^*T)T^* G_{\bar\lambda} z\right\|^2 \\
&= \|R_\alpha Tx - x\|^2 + \delta^2 \int_0^{+\infty} \lambda\, g_\alpha^2(\lambda)\, d\left\|F_\lambda G_{\bar\lambda} z\right\|^2 \\
&\geq \|R_\alpha Tx - x\|^2 + \delta^2 \int_{\bar\lambda}^{c_1\bar\lambda} \lambda\, g_\alpha^2(\lambda)\, d\left\|F_\lambda G_{\bar\lambda} z\right\|^2.
\end{aligned} \tag{6.30}$$

Consideraremos ahora dos casos posibles diferentes.

Caso I: $\alpha \leq \bar\lambda$. Como $c_1\bar\lambda \leq \lambda_1$ y $c_1 > 1$, se sigue de la hipótesis ***ii.e)*** que

$$g_\alpha(\lambda) \geq g_\alpha(c_1\bar\lambda) \geq g_\alpha(\lambda_1) \quad \text{para todo} \quad \lambda \in [\bar\lambda, c_1\bar\lambda]. \tag{6.31}$$

Por otro lado, de la hipótesis ***ii.a)*** resulta que $r_\alpha(\lambda_1) \leq 1$, lo cual implica que $\lambda_1\, g_\alpha(\lambda_1) \geq 0$ y por lo tanto, $g_\alpha(\lambda_1) \geq 0$. Se sigue entonces de (6.31) que $g_\alpha^2(\lambda) \geq g_\alpha^2(c_1\bar\lambda)$ para todo $\lambda \in [\bar\lambda, c_1\bar\lambda]$. Luego,

$$\begin{aligned}
\int_{\bar\lambda}^{c_1\bar\lambda} \lambda\, g_\alpha^2(\lambda)\, d\left\|F_\lambda G_{\bar\lambda} z\right\|^2 &\geq \bar\lambda\, g_\alpha^2(c_1\bar\lambda) \int_{\bar\lambda}^{c_1\bar\lambda} d\left\|F_\lambda G_{\bar\lambda} z\right\|^2 \\
&= \bar\lambda\, g_\alpha^2(c_1\bar\lambda),
\end{aligned} \tag{6.32}$$

donde la última igualdad se sigue del hecho que $\int_{\bar\lambda}^{c_1\bar\lambda} d\left\|F_\lambda G_{\bar\lambda} z\right\|^2 = 1$, lo cual es consecuencia de que $\int_{\bar\lambda}^{c_1\bar\lambda} d\left\|F_\lambda G_{\bar\lambda} z\right\|^2 = \left\|F_{c_1\bar\lambda} G_{\bar\lambda} z\right\|^2 - \left\|F_{\bar\lambda} G_{\bar\lambda} z\right\|^2$, de la definición de $G_{\bar\lambda}$, del hecho que $F_\lambda F_\mu = F_{\min\{\lambda,\mu\}}$ para todos $\lambda, \mu \in \mathbb{R}$ y de $\|G_{\bar\lambda} z\| = 1$.

A su vez, las hipótesis ***ii.a)*** e ***ii.c)*** implican que $g_\alpha(\lambda)$ es monótona decreciente con respecto a α para todo $\lambda \in [0, \lambda_1]$. Como $\alpha \leq \bar{\lambda}$ y $c_1\,\bar{\lambda} \leq \lambda_1$, se tiene entonces que

$$g_\alpha(c_1\,\bar{\lambda}) \geq g_{\bar{\lambda}}(c_1\,\bar{\lambda}), \tag{6.33}$$

y de la hipótesis ***ii.d)*** que

$$g_{\bar{\lambda}}(c_1\,\bar{\lambda}) \geq \gamma_2/\bar{\lambda} > 0. \tag{6.34}$$

De (6.33) y (6.34) concluimos que

$$g_\alpha^2(c_1\,\bar{\lambda}) \geq \left(\gamma_2/\bar{\lambda}\right)^2. \tag{6.35}$$

Sustituyendo (6.35) en (6.32) se tiene que

$$\int_{\bar{\lambda}}^{c_1\,\bar{\lambda}} \lambda\, g_\alpha^2(\lambda)\, d\left\|F_\lambda G_{\bar{\lambda}} z\right\|^2 \geq \gamma_2^2/\bar{\lambda},$$

lo cual junto a (6.30) implica que si $\alpha \leq \bar{\lambda}$, entonces $\left\|R_\alpha \bar{y}^{\,\delta} - x\right\|^2 \geq (\gamma_2\,\delta)^2/\bar{\lambda}$.

<u>Caso 2</u>: $\alpha > \bar{\lambda}$. En este caso se sigue de la hipótesis ***ii.c)*** que $r_\alpha^2(\lambda) \geq r_{\bar{\lambda}}^2(\lambda)$ para todo $\lambda \in (0, \|T\|^2]$. Luego,

$$\|R_\alpha T x - x\|^2 = \int_0^{+\infty} r_\alpha^2(\lambda)\, d\left\|E_\lambda x\right\|^2 \geq \int_0^{+\infty} r_{\bar{\lambda}}^2(\lambda)\, d\left\|E_\lambda x\right\|^2 = \left\|R_{\bar{\lambda}} T x - x\right\|^2,$$

lo cual junto a (6.30) implica que $\left\|R_\alpha \bar{y}^{\,\delta} - x\right\|^2 \geq \|R_{\bar{\lambda}} T x - x\|^2$.

Resumiendo los resultados obtenidos en los casos I y II, podemos escribir que

$$\begin{aligned} \left\|R_\alpha \bar{y}^{\,\delta} - x\right\|^2 &\geq \begin{cases} \left\|R_{\bar{\lambda}} T x - x\right\|^2, & \text{si } \alpha > \bar{\lambda} \\ (\gamma_2\,\delta)^2/\bar{\lambda}, & \text{si } \alpha \leq \bar{\lambda}. \end{cases} \\ &\geq \ \text{mín}\left\{\left\|R_{\bar{\lambda}} T x - x\right\|^2, (\gamma_2\,\delta)^2/\bar{\lambda}\right\}, \end{aligned} \tag{6.36}$$

lo cual es válido para todos $\alpha \in (0, \alpha_0)$, $\bar{\lambda} \in \sigma_p(TT^*)$ tal que $c_1\bar{\lambda} \leq \lambda_1$ y para todo $\delta > 0$. Entonces

$$\text{mín}\left\{\left\|R_{\bar{\lambda}} T x - x\right\|, \gamma_2\,\delta/\sqrt{\bar{\lambda}}\right\} = \left(\text{mín}\{\left\|R_{\bar{\lambda}} T x - x\right\|^2, (\gamma_2\,\delta)^2/\bar{\lambda}\}\right)^{1/2}$$

$$
\begin{aligned}
&\leq \inf_{\alpha\in(0,\alpha_0)} \left\| R_\alpha \bar{y}^\delta - x \right\| && \text{(por (6.36))}\\
&\leq \sup_{y^\delta \in \overline{B_\delta(Tx)}} \inf_{\alpha\in(0,\alpha_0)} \left\| R_\alpha y^\delta - x \right\| && \text{(pues } \bar{y}^\delta \in \overline{B_\delta(Tx)})\\
&= O(\rho(\Theta^{-1}(\delta))) \quad \text{cuando} \quad \delta \to 0^+ && \text{(por hipótesis).}
\end{aligned}
$$

Ahora, puesto que $\bar{\lambda} = \alpha(\bar{\delta})$ resuelve la ecuación (6.24), de la desigualdad anterior se tiene que

$$\left\| R_{\alpha(\bar{\delta})} Tx - x \right\| = \gamma_2\, \bar{\delta}/\sqrt{\bar{\lambda}} = O(\rho(\Theta^{-1}(\bar{\delta}))) \quad \text{cuando} \quad \bar{\delta} \to 0^+, \tag{6.37}$$

lo cual implica que

$$\frac{\bar{\delta}}{\rho(\Theta^{-1}(\bar{\delta}))} = O\left(\sqrt{\alpha(\bar{\delta})}\right) \quad \text{cuando} \quad \bar{\delta} \to 0^+. \tag{6.38}$$

Como $\delta = \Theta(\Theta^{-1}(\delta))$ resulta de la definición de Θ que $\delta = \sqrt{\Theta^{-1}(\delta)}\rho(\Theta^{-1}(\delta))$. Luego, se sigue de (6.38) que $\sqrt{\Theta^{-1}(\bar{\delta})} = O(\sqrt{\alpha(\bar{\delta})})$ cuando $\bar{\delta} \to 0^+$. De esto y de (6.37) se deduce entonces que

$$\left\| R_{\alpha(\bar{\delta})} Tx - x \right\| = O(\rho(\alpha(\bar{\delta}))) \quad \text{cuando } \bar{\delta} \to 0^+ \ \forall\, \alpha(\bar{\delta}) \in \sigma_p(TT^*) : c_1\, \alpha(\bar{\delta}) \leq \lambda_1. \tag{6.39}$$

Ahora, sea $\alpha \in \mathbb{R}^+$ tal que $\alpha \leq \max\limits_{k\in\mathbb{N}}\{\tilde{\lambda}_k : \tilde{\lambda}_k \leq \frac{\lambda_1}{c_1}\}$. Entonces existen $\tilde{\lambda}_k, \tilde{\lambda}_{k+1} \in \sigma_p(TT^*)$ tales que $\frac{\tilde{\lambda}_k}{\tilde{\lambda}_{k+1}} \leq c$, $\tilde{\lambda}_{k+1} < \alpha \leq \tilde{\lambda}_k$ y $(\tilde{\lambda}_{k+1}, \tilde{\lambda}_k) \cap \sigma_p(TT^*) = \emptyset$.

De la hipótesis ***ii.c)*** y del hecho que $\tilde{\lambda}_k \in \sigma_p(TT^*)$ y $\tilde{\lambda}_k \leq \frac{\lambda_1}{c_1}$ se sigue que

$$
\begin{aligned}
\| R_\alpha Tx - x \|^2 &= \int_0^{+\infty} r_\alpha^2(\lambda)\, d\, \| E_\lambda x \|^2\\
&\leq \int_0^{+\infty} r_{\tilde{\lambda}_k}^2(\lambda)\, d\, \| E_\lambda x \|^2\\
&= \left\| R_{\tilde{\lambda}_k} Tx - x \right\|^2\\
&= O(\rho^2(\tilde{\lambda}_k)), \quad \text{(en virtud de (6.39)).}
\end{aligned} \tag{6.40}
$$

De la hipótesis ***i)*** se tiene que $\tilde{\lambda}_k \leq c\, \tilde{\lambda}_{k+1}$ y como ρ es estrictamente creciente y no negativa resulta que

$$\rho^2(\tilde{\lambda}_k) \leq \rho^2(c\, \tilde{\lambda}_{k+1}). \tag{6.41}$$

Puesto que $c \geq 1$ y ρ es de tipo superior local β para algún $\beta \geq 0$ (hipótesis ***iii)***), existe una constante positiva d tal que

$$\rho(c\,\tilde{\lambda}_{k+1}) \leq d\,c^{\beta}\rho\left(\frac{1}{c}\,c\,\tilde{\lambda}_{k+1}\right) = d\,c^{\beta}\rho(\tilde{\lambda}_{k+1}). \tag{6.42}$$

De (6.40), (6.41), (6.42) y el hecho que $\rho^2(\tilde{\lambda}_{k+1}) < \rho^2(\alpha)$ resulta que $\|R_\alpha Tx - x\| = O(\rho(\alpha))$ cuando $\alpha \to 0^+$. Por lo tanto, el Lema 6.33 implica que $x \in Ran(\rho(T^*T))$, como queríamos probar. ∎

De la definición de calificación se sigue que

$$\|R_\alpha Tx - x\|^2 \leq \gamma^2\,\rho^2(\alpha) \int_0^{+\infty} \rho^{-2}(\lambda)\,d\,\|E_\lambda x\|^2 .$$

Por lo tanto, en el Lema anterior, la hipótesis ***i)*** y el hecho que ρ sea de tipo superior local β para algún $\beta \geq 0$ pueden sustituirse por el requerimiento que $\rho(T^*T)$ sea invertible, o equivalentemente, que $\rho^{-2}(\lambda)$ sea integrable con respecto a la medida $d\,\|E_\lambda x\|^2$ para todo $x \in \mathcal{X}$.

Demostración del Teorema 6.32.

Sin pérdida de generalidad suponemos que $\alpha_0 \doteq \text{mín}\{\frac{\lambda_1}{c_1}, \frac{\lambda_1}{k}\}$. Primero probaremos que $\psi(x,\delta) \doteq (\rho \circ \Theta^{-1})(\delta)$ para $x \in \mathcal{X}^\rho$ y $\delta \in (0, \Theta(\alpha_0))$, es cota superior de convergencia para el error total de $\{R_\alpha\}_{\alpha\in(0,\alpha_0)}$ en $\mathcal{X}^\rho$, es decir, probaremos que $\psi \in \mathcal{U}_{\mathcal{X}^\rho}(\mathcal{E}^{\text{tot}}_{\{R_\alpha\}})$. Para todo $r \geq 0$ definimos los conjuntos fuente $\mathcal{X}^{\rho,r} \doteq \{x \in \mathcal{X} : x = \rho(T^*T)\xi, \|\xi\| \leq r\}$. Sea $x \in \mathcal{X}^\rho$, entonces existe $r \geq 1$ tal que $x \in \mathcal{X}^{\rho,r}$. Como Θ es una función continua y estrictamente creciente en $(0,\alpha_0)$ existe un único $\tilde{\alpha} \in (0,\alpha_0)$ tal que $x \in \mathcal{X}^{\rho,r}$ y $\Theta(\tilde{\alpha}) = \frac{\delta}{r}$. Por lo tanto,

$$\begin{aligned}
\mathcal{E}^{\text{tot}}_{\{R_\alpha\}}(x,\delta) &= \inf_{\alpha\in(0,\alpha_0)} \sup_{y^\delta \in \overline{B_\delta(Tx)}} \left\|R_\alpha\,y^\delta - x\right\| \\
&\leq \sup_{y^\delta \in \overline{B_\delta(Tx)}} \left\|R_{\tilde{\alpha}}\,y^\delta - x\right\|
\end{aligned}$$

$$\leq \sup_{x \in \mathcal{X}^{\rho,r}} \sup_{y^\delta \in \overline{B_\delta(Tx)}} \left\| R_{\tilde{\alpha}}\, y^\delta - x \right\|. \tag{6.43}$$

Por otro lado, dado que $\{g_\alpha\}$ es un MRE que satisface la hipótesis $(H4)$ y tiene calificación ρ, junto al hecho que ρ trivialmente cubre ("*covers*") ρ con constante igual a 1 (ver [44], Definición 2) y puesto que $\Theta(\tilde{\alpha}) = \frac{\delta}{r}$, se sigue en virtud del Teorema 2 de [44] que existe una constante positiva K independiente de δ tal que

$$\sup_{x \in \mathcal{X}^{\rho,r}} \sup_{y^\delta \in \overline{B_\delta(Tx)}} \left\| R_{\tilde{\alpha}}\, y^\delta - x \right\| \leq K\, \rho\left(\Theta^{-1}\left(\frac{\delta}{r}\right)\right), \text{ para } 0 < \delta \leq r\, \Theta(\|T\|^2). \tag{6.44}$$

De (6.43) y (6.44) se sigue que para todo $\delta \in (0, \Theta(\alpha_0))$,

$$\mathcal{E}^{\mathrm{tot}}_{\{R_\alpha\}}(x,\delta) \leq K\, \rho\left(\Theta^{-1}\left(\frac{\delta}{r}\right)\right) \leq K\, \rho(\Theta^{-1}(\delta)) = K\, \psi(x,\delta),$$

donde la última desigualdad se sigue del hecho que $r \geq 1$ y tanto ρ como Θ^{-1} son funciones crecientes. Esto prueba que $\psi \in \mathcal{U}_{\mathcal{X}^\rho}(\mathcal{E}^{\mathrm{tot}}_{\{R_\alpha\}})$.

A continuación veremos que ψ satisface la condición $(S1)$ de saturación sobre $\mathcal{X}^\rho$. Del hecho que $\{g_\alpha\}$ es un MRE que satisface la hipótesis $(H4)$ y posee calificación máxima ρ y puesto que se satisfacen además las hipótesis ***iv)*** y ***v)*** se sigue, en virtud del Teorema 2.3 y de la Definición 2.2 de [42], que para todos $x^* \in \mathcal{X}$, $x^* \neq 0$ y $x \in \mathcal{X}^\rho$ existen constantes positivas $a \doteq a(x,x^*)$ y $d = d(x,x^*)$ tales que

$$\frac{\mathcal{E}^{\mathrm{tot}}_{\{R_\alpha\}}(x^*,\delta)}{\psi(x,\delta)} \geq a \quad \forall\, \delta \in (0,d).$$

Luego, $\limsup_{\delta \to 0^+} \frac{\mathcal{E}^{\mathrm{tot}}_{\{R_\alpha\}}(x^*,\delta)}{\psi(x,\delta)} > 0$ para todos $x^* \in \mathcal{X}$, $x^* \neq 0$ y $x \in \mathcal{X}^\rho$, es decir, ψ satisface la condición $(S1)$ sobre $\mathcal{X}^\rho$.

Además, como ψ es independiente de x, es invariante sobre $\mathcal{X}^\rho$, es decir, ψ satisface la condición $(S2)$ de saturación.

Resta probar que ψ satisface la condición $(S3)$. Supongamos que no. Entonces existen $M \supsetneqq \mathcal{X}^\rho$ y $\tilde{\psi} \in \mathcal{U}_M(\mathcal{E}^{\mathrm{tot}}_{\{R_\alpha\}})$ tales que $\tilde{\psi}\,|_{\mathcal{X}^\rho} = \psi$ y $\tilde{\psi}$ satisface $(S1)$ y $(S2)$ sobre M. Sea $x^* \in M \setminus \mathcal{X}^\rho$, $x^* \neq 0$. Como $\tilde{\psi} \in \mathcal{U}_M(\mathcal{E}^{\mathrm{tot}}_{\{R_\alpha\}})$ se tiene que

$$\mathcal{E}^{\mathrm{tot}}_{\{R_\alpha\}} \overset{\{x^*\}}{\preceq} \tilde{\psi}. \tag{6.45}$$

Puesto que $\tilde{\psi}$ es invariante sobre M y $\mathcal{X}^\rho \subset M$, se sigue que $\tilde{\psi} \overset{\{x^*\},\mathcal{X}^\rho}{\preceq} \tilde{\psi}$ y como $\tilde{\psi}$ coincide con ψ sobre $\mathcal{X}^\rho$, resulta que $\tilde{\psi} \overset{\{x^*\},\mathcal{X}^\rho}{\preceq} \psi$. Esto junto a (6.45) implica que $\mathcal{E}^{\rm tot}_{\{R_\alpha\}} \overset{\{x^*\},\mathcal{X}^\rho}{\preceq} \psi$ y por lo tanto $\mathcal{E}^{\rm tot}_{\{R_\alpha\}}(x^*,\delta) = O(\rho(\Theta^{-1}(\delta)))$ cuando $\delta \to 0^+$. El Lema 6.30 implica entonces que existe una regla de elección de parámetros *a-priori* $\tilde{\alpha} : \mathbb{R}^+ \to (0,\alpha_0)$ tal que

$$\sup_{y^\delta \in \overline{B_\delta(Tx^*)}} \left\| R_{\tilde{\alpha}(\delta)} y^\delta - x^* \right\| = O(\rho(\Theta^{-1}(\delta))) \quad \text{cuando } \delta \to 0^+.$$

Luego,

$$\sup_{y^\delta \in \overline{B_\delta(Tx^*)}} \inf_{\alpha \in (0,\alpha_0)} \left\| R_\alpha y^\delta - x^* \right\| = O(\rho(\Theta^{-1}(\delta))) \quad \text{cuando } \delta \to 0^+.$$

Finalmente, el Lema 6.34 implica que $x^* \in Ran(\rho(T^*T))$ y como $x^* \neq 0$, se tiene que $x^* \in \mathcal{X}^\rho$, lo cual contradice que $x^* \in M \setminus \mathcal{X}^\rho$. Así, ψ satisface la condición (*S3*) y por lo tanto, ψ es función de saturación de $\{R_\alpha\}$ sobre $\mathcal{X}^\rho$. ■

Observar que los Lemas 6.29 y 6.30 siguen siendo válidos si se reemplazan las hipótesis ***iii)*** y ***iv)*** del Teorema 6.28 por el hecho que exista $\rho : (0, \|T\|^2] \to (0,\infty)$ que sea calificación de $\{R_\alpha\}_{\alpha \in (0,\alpha_0)}$ y satisfaga la hipótesis ***iii.a)*** del Teorema 6.32.