Convergence of the Neumann series in higher norms

Charles L. Epstein*
Department of Mathematics, University of Pennsylvania

Version 1.0   August 1, 2003

Abstract

Natural conditions on an operator A are given so that the Neumann series for $(\mathrm{Id}+A)^{-1}$ converges in higher norm topologies.

1 The general case

In applications one often considers equations of the form

$$(\mathrm{Id}+A)u = v \tag{1}$$

where A is a bounded linear operator. In this note we consider the convergence of the Neumann series solution for this type of equation. We first prove a fairly abstract result and then consider a variety of special cases. Suppose that $\{X_k\}$ is a nested collection of Banach spaces, $X_0 \supset X_1 \supset \dots X_k \supset X_{k+1} \supset \dots$, with norms $\{\|\cdot\|_k\}$. Let $|||A|||$ denote the operator norm of $A : X_0 \to X_0$,

$$|||A||| = \sup_{X_0 \ni u \neq 0} \frac{\|Au\|_0}{\|u\|_0}. \tag{2}$$

If $|||A||| < \alpha_0 < 1$ then it is well known that the Neumann series for $(\mathrm{Id}+A)^{-1}$ converges, see [1]. Indeed, if we define the sequence

$$\begin{aligned} u_0 &= v \\ u_j &= v - Au_{j-1} \end{aligned} \tag{3}$$

then u_j is the jth partial sum of the Neumann series. A simple induction argument shows that

$$\|u_{j+1} - u_j\|_0 \leq \alpha_0^j \|u_1 - u_0\|_0. \tag{4}$$

*Research partially supported by NSF grant DMS02–03705 and the Francis J. Carey chair. Address: Department of Mathematics, University of Pennsylvania, Philadelphia, PA. E-mail: cle@math.upenn.edu
Keywords: Neumann series, higher norms, Marchenko equation

If we further suppose that $A : X_k \to X_k$ is a bounded operator for $k \leq K$, and $v \in X_k$, for such a k, then the iterates $\{u_j\}$, defined in equation (3), belong to X_k. Our main theorem gives conditions under which the solution u to (1) also belongs to X_k, and the iterates defined in (3) converge to u in the X_k-topology.

Theorem 1. *Suppose that $A : X_0 \to X_0$ is a bounded linear operator and K is a positive integer or ∞. Assume that*

1. $\|A\| = \alpha_0 < 1$
2. $A : X_k \to X_k$ *boundedly for each* $k \leq K$
3. *For each $k \leq K$ there are constants $\alpha_k < 1$ and $C_k < \infty$ so that, for $u \in X_k$, we have the estimate*
$$\|Au\|_k \leq \alpha_k \|u\|_k + C_k \|u\|_{k-1}. \tag{5}$$

If $v \in X_k$, for a $k \leq K$, then the sequence $\{u_j\}$ defined in (3) *converges to u in the X_k-topology and therefore the solution to* (1) *belongs to X_k.*

***Remark* 1.** The estimate in (5) is natural from the perspective of pseudodifferential operators. Suppose that M is a compact manifold and $X_k = H^k(M)$. If A is a pseudodifferential operator on M, of order zero, whose principal symbol satisfies
$$\sup_{x\in M} \limsup_{\xi\to\infty} |\sigma_0(A)(x,\xi)| = \alpha < 1, \tag{6}$$
then A satisfies the estimates in (5), with $\alpha_k = \alpha$. See [2].

Proof. The proof of the theorem is a small extension of the proof for the $k = 0$ case. By induction assume that we have shown, for $l < k$, that there are constants, $\{\beta_l\}$ less than 1, and constants $\{C'_l\}$ so that, for all j, we have:
$$\|u_{j+1} - u_j\|_l \leq C'_l \beta_l^j \|u_1 - u_0\|_l. \tag{7}$$
The estimate in (5), and the definition of the sequence, $\{u_j\}$, imply that
$$\|u_{j+1} - u_j\|_k \leq \alpha_k \|u_j - u_{j-1}\|_k + C_k \|u_j - u_{j-1}\|_{k-1} \tag{8}$$
Applying (7) gives the estimate
$$\|u_{j+1} - u_j\|_k \leq \alpha_k \|u_j - u_{j-1}\|_k + C_k C'_{k-1} \beta_{k-1}^{j-1} \|u_1 - u_0\|_{k-1} \tag{9}$$
Using this estimate recursively we obtain
$$\|u_{j+1} - u_j\|_k \leq \alpha_k^j \|u_1 - u_0\|_k + C_k C'_{k-1} \left[\sum_{m=1}^{j} \beta_{k-1}^{j-m} \alpha_k^{m-1}\right] \|u_1 - u_0\|_{k-1}. \tag{10}$$
From this estimate it is immediate that there exists a constant $\beta_k < 1$ and $C'_k < \infty$ so that
$$\|u_{j+1} - u_j\|_k \leq C'_k \beta_k^j \|u_1 - u_0\|_k. \tag{11}$$
This completes the proof of the induction hypothesis. The proof of the theorem follows easily from (11) via the $k = 0$ argument. □

2 Positive self adjoint operators

In this section we assume that X_0 is a Hilbert space and that $B : X_0 \to X_0$ is a bounded, positive, self adjoint operator. For every $v \in X_0$ the equation

$$(\mathrm{Id} + B)u = v \tag{12}$$

has a unique solution. By slightly modifying the equation, the solution can be found by summing a Neumann series. In particular, we observe that, if γ is a constant, then u is a solution to (12) if and only if u is a solution to

$$[\mathrm{Id} + \frac{2}{2+\gamma}(B - \frac{\gamma}{2}\,\mathrm{Id})]u = \frac{2}{2+\gamma}v. \tag{13}$$

Lemma 1. *Let $B : X_0 \to X_0$ be a bounded, positive, self adjoint operator. If $\gamma = \|B\|$ then the Neumann sequence for equation* (13)*:*

$$\begin{aligned} u_0 &= \frac{2}{2+\gamma}v, \\ u_{j+1} &= \frac{2}{2+\gamma}v - \frac{2}{2+\gamma}(B - \frac{\gamma}{2}\,\mathrm{Id})u_j. \end{aligned} \tag{14}$$

converges, in the X_0-topology, to the unique solution of (12)*.*

Proof. Since B is a positive operator, the spectrum of B lies in the interval $[0, \gamma]$, and, therefore

$$\left\| \frac{2}{2+\gamma}(B - \frac{\gamma}{2}\,\mathrm{Id}) \right\| = \frac{\gamma}{2+\gamma} < 1. \tag{15}$$

This in turn implies that the sequence defined in (14) converges in the X_0-topology to the solution of equation (13). □

To apply Theorem 1 to equation (12) requires further hypotheses on B. First we assume that $B : X_k \to X_k$ is bounded for every $k \leq K$. A simple natural assumption is that B is a “smoothing”operator, so that there are constants $\{C_k\}$ such that, we have the estimates

$$\|Bu\|_k \leq C_k \|u\|_{k-1}. \tag{16}$$

In fact somewhat less is needed to apply the theorem. It suffices to assume that there are constants $\{\alpha_k\}$ and $\{C_k\}$ so that, for every k, we have

$$\begin{aligned} \alpha_k &< 1, \\ \|Bu\|_k &\leq \alpha_k \|u\|_k + C_k \|u\|_{k-1}. \end{aligned} \tag{17}$$

Proposition 1. *Let K be a positive integer, or infinity. Assume that $B : X_k \to X_k$ is a bounded operator for $k \leq K$, which is moreover positive and self adjoint, when $k = 0$. Suppose there are constants $\{\alpha_k\}$ and $\{C_k\}$ so that the estimates in* (17) *hold for $k \leq K$. If $v \in X_k$, for a $k \leq K$, then the iterates defined in* (14) *converge to u in the X_k-topology.*

Proof. We only need to verify the third hypothesis of Theorem 1 for the operator $\frac{2}{2+\gamma}(B - \frac{\gamma}{2}\operatorname{Id})$. Let $u \in X_k$, then the triangle inequality, and our assumptions on B imply that

$$\begin{aligned}\|\frac{2}{2+\gamma}(B - \frac{\gamma}{2}\operatorname{Id})u\|_k &\leq \|\frac{2}{2+\gamma}Bu\|_k + \frac{\gamma}{2+\gamma}\|u\|_k \\ &\leq \left[\frac{2\alpha_k + \gamma}{2+\gamma}\right]\|u\|_k + \frac{2C_k}{2+\gamma}\|Bu\|_{k-1}.\end{aligned} \tag{18}$$

The assumption $\alpha_k < 1$ implies that

$$\frac{2\alpha_k + \gamma}{2+\gamma} < 1. \tag{19}$$

Hence we can apply Theorem 1 to complete the proof of the Proposition. □

***Remark* 2.** Suppose that B is a pseudodifferential operator of order zero, on a compact manifold M with a nonnegative principal symbol, $\sigma_0(B)$. If

$$\sup_{x\in M} \limsup_{\xi\to\infty} |\sigma_0(B)(x,\xi)| = \tilde{\gamma}_1, \tag{20}$$

then $\frac{2}{2+\tilde{\gamma}_1}(B - \frac{\tilde{\gamma}_1}{2}\operatorname{Id})$ is an order zero pseudodifferential operator with principal symbol $\frac{2}{2+\tilde{\gamma}_1}(\sigma_0(B) - \frac{\tilde{\gamma}_1}{2})$. The sup-norm of this symbol is bounded by $\frac{\tilde{\gamma}_1}{2+\tilde{\gamma}_1}$. Using for $\{X_k\}$ the standard Sobolev spaces $\{H^k(M)\}$, Garding's inequality implies that, for each $k > 0$, there is a constant C_k, so that:

$$\|\frac{2}{2+\tilde{\gamma}_1}(B - \frac{\tilde{\gamma}_1}{2}\operatorname{Id})u\|_k \leq \frac{\tilde{\gamma}_1}{2+\tilde{\gamma}_1}\|u\|_k + C_k\|u\|_{k-1}. \tag{21}$$

Suppose that B is also a positive self adjoint operator, with L^2-norm $\tilde{\gamma}_2$. If $\gamma = \max\{\tilde{\gamma}_1, \tilde{\gamma}_2\}$ then, for $v \in L^2(M)$, Lemma 1 applies to show that the sequence in (14) converges in $L^2(M)$ to the solution, u, of $(\operatorname{Id} + B)u = v$. If $v \in H^k(M)$, then the estimates in (21), with $\tilde{\gamma}_1$ replaced by γ allow Theorem 1 to be applied to conclude that this sequence also converges in $H^k(M)$.

***Remark* 3.** If B is a positive operator then there are many other iteration schemes which converge to the solution of (12). For example, the conjugate gradient method provides a different sequence. It seems quite an interesting question whether, under hypotheses like those in the proposition, these schemes provide sequences which converge in the X_k-topology.

3 A Marchenko equation

We finish by considering a concrete example which arises in the inverse scattering theory of the Zakharov-Shabat 2×2-system. Suppose that f is function defined on $\mathbb{R}$

which belongs to $L^1([2t,\infty))$, for every finite t. For each such t, define the operator, F_t, on $L^2([t,\infty))$

$$F_t h(s) = \int_t^\infty f(s+y)h(y)dy. \tag{22}$$

Lemma 2. *If f belongs to $L^1([2t,\infty))$, then the F_t is a bounded operator on $L^2([t,\infty))$ with*

$$\|F_t\| = \int_{2t}^\infty |f(x)|dx. \tag{23}$$

Proof. The proof is a straightforward application of the Cauchy-Schwarz inequality, which we leave to the interested reader. □

A simple calculation shows that the adjoint of F_t, as a operator on $L^2([t,\infty))$, is given by

$$F_t^* h(s) = \int_t^\infty \bar{f}(s+y)h(y)dy. \tag{24}$$

The Marchenko equation for the ZS-2 $\times$ 2 system can be written

$$[(\mathrm{Id} + F_t^* F_t)k_t](s) = \bar{f}(s+t). \tag{25}$$

For each t, this evidently satisfies the hypotheses of Lemma 1 with $\gamma_t = \|F_t^* F_t\|$, where

$$\gamma_t = \left[\int_{2t}^\infty |f(x)|dx\right]^2.$$

For applications, it is of considerable interest to know when a sequence converging to a solution of the Marchenko equation also converges in a stronger topology . For this case we take $X_k = H^k([t,\infty))$, with the norms

$$\|h\|_k^2 = \sum_{j=0}^k \|\partial_x^j h\|^2_{L^2([t,\infty))}.$$

Lemma 3. *Suppose that the functions $\{f, \partial_x f, \ldots, \partial_x^k f\}$ belong to $L^1([2t,\infty))$. Then F_t defines a bounded map from X_0 to X_k, for $j = 1, \ldots, k$.*

Proof. This follows immediately from Lemma 2 and the observation that

$$\partial_x^j [F_t h](s) = \int_t^\infty (\partial_x^j f)(s+y)h(y)dy. \tag{26}$$

□

Using integration by parts one easily proves the following:

Lemma 4. *Suppose that the functions* $\{f, \partial_x f, \dots, \partial_x^{k-1} f\}$ *belong to* $L^1([2t,\infty)) \cap L^2([2t,\infty))$. *Then* F_t *defines a bounded map from* X_j *to* X_j *for each* $j \leq k$.

Proof. We sketch the $k = 1$ case. Integrating by parts we see that

$$\partial_x [F_t h](s) = -f(s+t)h(t) - \int_t^\infty f(s+y) h_x(y) dy. \tag{27}$$

The claim follows from the hypotheses of the lemma, Lemma 2 and the elementary estimate:

$$|h(t)| \leq \sqrt{\|h_x\|^2_{L^2([t,\infty))} + \|h\|^2_{L^2([t,\infty))}}. \tag{28}$$

The general case follows by repeatedly differentiating (27) and integrating by parts. □

As $F_t^* F_t$ is a positive self adjoint operator on $L^2([t,\infty))$, we can apply Proposition 1 to the Marchenko equation to obtain:

Proposition 2. *Suppose that the functions* $\{f, \partial_x f, \dots, \partial_x^k f\}$ *belong to* $L^1([2t,\infty))$ *and* $\{f, \partial_x f, \dots, \partial_x^{k-1} f\}$ *to* $L^2([2t,\infty))$. *If* $v \in H^k([t,\infty))$ *then the sequence defined by*

$$\begin{aligned} h_0 &= \frac{2}{2+\gamma_t} v, \\ h_{j+1} &= \frac{2}{2+\gamma_t} v - \frac{2}{2+\gamma_t} [F_t^* F_t - \frac{\gamma_t}{2} \operatorname{Id}] h_j \end{aligned} \tag{29}$$

converges in $H^k([t,\infty))$ *to the unique solution of*

$$[\operatorname{Id} + F_t^* F_t] h = v. \tag{30}$$

Proof. Lemma 3, and the hypotheses imply that for each $j \leq k$, there is a constant C_j so that, for all $u \in L^2([t,\infty))$, we have the estimate

$$\|F_t^* F_t u\|_j \leq C_j \|F_t u\|_0 \leq \sqrt{\gamma_t} C_j \|u\|_0. \tag{31}$$

Hence $F_t^* F_t$ satisfies satisfies the hypotheses of Proposition 1 with $\alpha_k \equiv 0$. □

The $k = 1$ case is of particular interest in applications. In this case, the image of the unit ball in $L^2([t,\infty))$ under $F_t^* F_t$ consists of uniformly bounded, uniformly equicontinuous functions. Hence the iterates defined in (18) are also uniformly bounded and uniformly equicontinuous. It is therefore an easy consequence of the Arzela-Ascoli theorem that they converge locally uniformly to the solution of the Marchenko equation.

Acknowledgment

I would like to that Carlos Tomei And Jeremy Magland for helpful remarks and corrections.

References

[1] F. RIESZ AND B. SZ.-NAGY, *Functional Analysis*, Frederick Ungar Publishing Co., New York, 1978. Translated from the 2nd French Edition by Leo F. Boron.

[2] M. E. TAYLOR, *Pseudodiferential Operators*, Princeton Univ. Press, Princeton, 1981.