J. Computational and Applied Math., **60** (1995)

Numerical Verification of Existence and Inclusion of Solutions for Nonlinear Operator Equations

Shin'ichi Oishi
Department of Computer and Information Sciences,
School of Science and Engineering, Waseda University,
Tokyo 169, Japan. e-mail:oishi@oishi.info.waseda.ac.jp

Abstract

Abstract nonlinear operator equations of the type

$$f(u) \equiv Lu + Nu = 0, \ u \in D(L)$$

is considered, where L is a densely defined closed linear operator from a Banach space X to an another Banach space Y and N a densely defined nonlinear operator from X to Y. A method is presented for numerical verification and inclusion of solutions for this equation.

1 Introduction

In this paper, we are concerned with abstract nonlinear operator equations of the type

$$f(u) \equiv Lu + Nu = 0, \ u \in D(L) \tag{1}$$

where L is a densely defined closed linear operator from a Banach space X to Y, and N a densely defined nonlinear operator from X to Y.This type of equations occur in a variety of situations in both pure and applied sciences. Eq. (1) is sometimes called a coincidence equation because one wants to find a point u for which the images under L and -N coincide. The purpose of the paper is to present a method for numerical verification and inclusion of solutions for Eq. (1). In the following, the domain of the definition of $L, D(L)$, and that for $N, D(N)$ is assumed to be Banach spaces satisfying $D(L) \subset D(N)$. For the sake of simplicity we will denote $D = D(L)$ and $B = D(N)$. The norms of D, B, X, and Y will be denoted by $\| \cdot \|_D$, $\| \cdot \|_B$, $\| \cdot \|_X$, and $\| \cdot \|_Y$, respectively.

In 1965, Urabe[13] has presented a method for numerical verification and inclusion of solutions for Eq. (1) for the case of $L = d/dt$. Then, he[14],[12] and his coresearchers[9],[10], [11] presented various results of numerical verifications for periodic and quasi-periodic solutions of ordinary differential equations. Urabe's method is based on his convergence theorem of a simplified Newton method for operator equations on suitable functional spaces. From the numerical analytic point of view, to apply Urabe's theorem, to obtain an estimate for the operator norm of the inverse of a certain linear operator becomes a key point. Urabe has presented a method in which this can be obtained by numerically obtaining a certain fundamental matrix. In 1972, Bouc[1] has shown that this kind of estimate can be accomplished without numerical integration by using functional analytic techniques for the case of $L = d/dt$. This paper is an extension of Urabe-Bouc's approach. That is, in this paper, we will treat the case in which L is a general closed operator L. By this, not only ordinary differential equations but also partial differential equations can be treated. Since mathematically rigorous bounds is required in obtaining such estimate, we have developed a rational arithmetic library. In this system using a continued fraction expansion of rational numbers, rounding errors during the numerical computations are completely taken into account.

Historically, several authors has presented different ways to use computers in proving the existence of solutions for nonlinear operator equations. Kantrovich[5] has presented a convergence theorem of Newton method on functional spaces and treated various kinds of functional equations. Kedem[6] has utilized this Newton-Kantrovich theorem to numerically prove the existence of solutions for certain two-point boundary problems. Cesari[2] presented also a method based on the alternative method. Collatz[3] and have presented methods based on the monotonicity or inverse-positivity. More recently, Kaucher-Miranker[4] presented a method using bases expansions. Nakao[7] has presented an infinite dimensional interval method and treated not only ordinary differential equations but also partial differential equations of various types. Plum[8] has also presented a method based on eigenvalue estimations. Our method of estimating the operator norm of a certain linear operator is completely different from these method.

2 Graph Norm Estimate

We consider here the graph norm introduced by L in $D(L)$:

$$\|u\|_L = \|u\|_X + \|Lu\|_Y \quad \text{for } u \in D(L)$$

Since L is closed, $D(L)$ becomes a Banach space with respect to the norm $\|u\|_L$. We denote this Banach space D_L. We assume that N is continuously Fréchet differentiable as a map from D_L to Y. For $u \in D_L$, we assume that the first derivative of $N, D_u N(u) = S(u)$, can be extended to a bounded linear map from X to Y. In order to numerically verify existence of solutions for Eq. (1), we introduce now a numerical framework. Let E and F be finite dimensional subspaces of D_L and Y, respectively, with $\dim E = \dim F$. Let P and Q be projections from D_L to E and Y to F, respectively. We assume that

$$\|u - Pu\|_X \leq c\|Lu\|_Y \quad \text{for } \forall u \in D_L \tag{2}$$

$$QLu = QLPu \quad \text{for } \forall\ u \in D_L \tag{3}$$

and

$$\|Q\|_{L(Y,Y)} \leq 1 \tag{4}$$

hold. Here, c is a constant independent of u. It should be noted here that for a choice of P, we usually suppose a situation in which the constant c can be chosen arbitrary small provided that $\dim E$ becomes sufficiently large.

Let $\{e_1, e_2, \cdots, e_m\}$ and $\{v_1, v_2, \cdots, v_m\}$ be bases of E and F, respectively. Then any element $e \in E$ and $v \in F$ can be represented as

$$e = \sum_{n=0}^{m} c_n(e) e_n \tag{5}$$

and

$$v = \sum_{n=0}^{m} d_n(v) v_n, \tag{6}$$

respectively. Here, $c_n(e)$'s and $d_n(v)$'s are suitable linear functionals. Thus maps $A_m : E \to E_m$ and $B_m : F \to F_m$ can be defined as

$$A_m e = (c_1(e), c_2(e), \cdots, c_m(e))^t \tag{7}$$

and

$$B_m v = (d_1(v), d_2(v), \cdots, d_m(v))^t, \tag{8}$$

respectively. Here, the superscript t denotes the transposition of vectors,

$$E_m = \{(c_1(e), c_2(e), \cdots, c_m(e))^t | e \in E\}$$

and

$$F_m = \{(d_1(v), d_2(v), \cdots, d_m(v))^t | v \in F\}.$$

For $c = (c_1, c_2, \cdots, c_m)^t$ and $d = (d_1, d_2, \cdots, d_m)^t$, define

$$\|c\|_{E_m} = \|\sum_{n=1}^{m} c_n e_n\|_X \tag{9}$$

and

$$\|d\|_{F_m} = \|\sum_{n=1}^{m} d_n v_n\|_Y. \tag{10}$$

Now, let $\tilde{u} \in E$ be an approximate solution of Eq. (1). Then, a linear transformation $J : E_m \to F_m$ can be defined as

$$J A_m P y = B_m\{Q(LPy + S(\tilde{u}))Py\}. \tag{11}$$

Since E_m and F_m are finite dimensional vector spaces, J can be considered as a matrix. If $\det J \neq 0$, we have

$$\|Py\|_Y \leq M\|Q(LPy + S(\tilde{u}))Py\|_X. \tag{12}$$

Here, M is a constant such that

$$\|J^{-1}\|_{L(F_m,E_m)} \leq M. \tag{13}$$

Then, one of our main results can be stated as follows:

Theorem 1 *Assume that* $\det J \neq 0$. *Let* K *and* M *be constants such that* $\|S(\tilde{u})\|_{L(X,Y)} \leq K$ *and* $\|J^{-1}\|_{L(F_m,E_m)} \leq M$. *If* $cK(1 + MK) < 1$, *then the map* $G(\tilde{u}) = L + S(\tilde{u}) : D_L \to Y$ *satisfies the following estimate for any* $y \in D_L$:

$$\|y\|_L \leq C\|G(\tilde{u})y\|_Y, \tag{14}$$

where

$$C = \frac{(1 + c)(1 + MK) + M}{1 - cK(1 + MK)}.$$

Now we define a residual

$$r = \|f(\tilde{u})\|_Y.$$

Moreover, we assume that

$$\|u\|_B \leq d\|u\|_L$$

holds for any $u \in D_L$, where d is constant independent of u. Let $U_p = B(\tilde{u}, p)$ be a closed ball in D_L centered at $\tilde{u}$ with a radius p. Here, if we assume that $S(u) = D_u N(u) : D_L \to Y$ is locally Lipschitz continuous:

$$\|S(u) - S(v)\|_{L(D_L,Y)} = a_U \|u - v\|_L \quad \text{for } u, v \in U \subset D_L.$$

Then we have

Theorem 2 *We assume that $G(\tilde{u}) : D_L \to Y$ has an inverse and $cK(1 + MK) < 1$ holds. Let $a = a_{U_p}$. If p satisfies*

1. $2Cr \leq p$ *and*
2. $aCp < 1$,

then there exists a solution u^* of Eq. (1) uniquely in U_p and satisfies

$$\|u^* - \tilde{u}\|_L \leq 2Cr.$$

Then, we also have

$$\|u^* - \tilde{u}\|_B \leq 2dCr.$$

3 Proof of Theorem 1

Put

$$G(\tilde{u})x = Lx + S(\tilde{u})x, \qquad G(\tilde{u}) : D_L \to Y. \tag{15}$$

For $x \in D_L$, we have

$$\begin{aligned} \|x\|_X &\leq \|x - Px\|_X + \|Px\|_X \\ &\leq c\|Lx\|_Y + \|Px\|_X. \end{aligned} \tag{16}$$

From the definition of (15), we have

$$\begin{aligned} \|Lx\|_Y &\leq \|G(\tilde{u})x\|_Y + \|S(\tilde{u})x\|_Y \\ &\leq \|G(\tilde{u})x\|_Y + K\|x - Px + Px\|_X \\ &\leq \|G(\tilde{u})x\|_Y + Kc\|Lx\|_Y + K\|Px\|_X. \end{aligned} \tag{17}$$

Moreover from (15) and (3), we have

$$QG(\tilde{u})x = QLx + QS(\tilde{u})(x - Px + Px) = QLPx + QS(\tilde{u})(x - Px + Px).$$

Here, if we put

$$s = QLPx + QS(\tilde{u})Px = Q[G(\tilde{u})x - S(\tilde{u})(x - Px)],$$

using (4) we have

$$\|s\|_Y \leq \|G(\tilde{u})x\|_Y + Kc\|Lx\|_Y. \tag{18}$$

Substituting the relation

$$\|Px\|_X \leq M\|s\|_Y \tag{19}$$

and (18) into (17), we have

$$\begin{array}{rcl} \|Lx\|_Y & \leq & \|G(\tilde{u})x\|_Y + Kc\|Lx\|_Y + MK\|s\| \\ & \leq & \|G(\tilde{u})x\|_Y + Kc\|Lx\|_Y + MK(\|G(\tilde{u})x\|_Y + Kc\|Lx\|_Y) \\ & = & (1+MK)\|G(\tilde{u})x\|_Y + Kc(1+MK)\|Lx\|_Y. \end{array}$$

Thus we have

$$\|Lx\|_Y \leq \frac{1+MK}{1-cK(1+MK)}\|G(\tilde{u})x\|_Y. \tag{20}$$

On the other hand, substituting (19) and (18) into (16), we have

$$\begin{array}{rcl} \|x\|_X & \leq & c\|Lx\|_Y + M\|s\|_Y \\ & \leq & c\|Lx\|_Y + M(\|G(\tilde{u})x\|_Y + Kc\|Lx\|_Y) \\ & = & c(1+MK)\|Lx\|_Y + M\|G(\tilde{u})x\|_Y. \end{array}$$

From this and (20), we have

$$\|x\|_X \leq \frac{c(1+MK)+M}{1-cK(1+MK)}\|G(\tilde{u})x\|_Y. \tag{21}$$

Summing up the above-mentioned discussions, we finally have@@@@@@@

$$\|x\|_L = \|x\|_X + \|Lx\|_Y \leq \frac{(1+c)(1+MK)+M}{1-cK(1+MK)}\|G(\tilde{u})x\|_Y$$

provided that $cK(1+MK) < 1$. This proves Theorem 1. *QED*

4 Proof of Theorem 2

We shall prove Theorem 2 by showing that the operator T defined in the below becomes a contraction mapping on U_p under the conditions of Theorem 2. Using $G(\tilde{u})^{-1}$, let us define an operator $T : D' \to D'$ by

$$Tu = G(\tilde{u})^{-1}(S(\tilde{u})u - Nu).$$

Since $G(\tilde{u})^{-1}$ exists, a fixed point of T is a solution of Eq. (1). In the first place, we shall show that $TU_p \subset U_p$. For any $u \in U_p$, we have

$$\begin{aligned} \|Tu - \tilde{u}\|_L &= \|G(\tilde{u})^{-1}(S(\tilde{u})u - Nu) - \tilde{u}\|_L \\ &= \|G^{-1}(S(\tilde{u})u - Nu - G(\tilde{u})\tilde{u})\|_L \\ &\leq C\|S(\tilde{u})u - Nu - G(\tilde{u})\tilde{u}\|_Y \\ &= C\|S(\tilde{u})u - Nu - L\tilde{u} - S(\tilde{u})\tilde{u}\|_Y \\ &= C(\| - Nu + N\tilde{u} - S(\tilde{u})(\tilde{u} - u)\|_Y + r). \end{aligned} \tag{22}$$

From the following estimate,

$$Nu = N\tilde{u} + S(\tilde{u})(u-\tilde{u}) + R,$$
$$\|R\|_Y \leq \frac{a}{2}\|u-\tilde{u}\|_Y,$$

we have

$$\begin{aligned} \|Tu-\tilde{u}\| &\leq C(\frac{a}{2}\|u-\tilde{u}\|_Y^2 + r) \\ &\leq C(\frac{a}{2}p^2 + r) < p. \end{aligned} \tag{23}$$

This implies $TU_p \subset U_p$.

We now show that T is contractive on U_p. For for $u, v \in U_p$, we have

$$\begin{aligned} \|Tu - Tv\|_L &\leq \|G(\tilde{u})^{-1}(S(v)u - Nu) - G(\tilde{u})^{-1}(S(v)v - Nv)\|_L \\ &= \|G(\tilde{u})^{-1}(S(v)(u-v) - (Nu - Nv))\|_L \\ &\leq C\|S(v)(u-v) - (Nu-Nv)\|_Y. \end{aligned}$$

Using the formula

$$Nu - Nv = \int_0^1 S(u + t(v-u))(v-u)dt,$$

it is seen that

$$\begin{aligned} &\|S(v)(u-v) - (Nu-Nv)\|_Y \\ &= \|\int_0^1 S(u+t(v-u)) - S(v)(v-u)dt\|_Y \\ &\leq \int_0^1 \sup_t \|S(u+t(v-u)) - S(v)(v-u)\|_{L(D',Y)}\|v-u\|_L dt \\ &= a\|v-u\|_L. \end{aligned}$$

Thus we have

$$\|Tu - Tv\|_L \leq aCp\|v-u\|_L. \tag{24}$$

This shows that T is contractive on U_p. Thus it follows that there exists unique a fixed point u^* of T in U_p. From the relation

$$\|u^* - \tilde{u}\|_L \leq (\frac{a}{2}Cp\|Tu^* - \tilde{u}\|_L + Cr),$$

we obtain an error bound

$$\|u^* - \tilde{u}\|_L \leq 2Cr.$$

This completes the proof. *QED*

5 Applications to Ordinary Differential Equations

In this section, taking a simple example, we study an application of the results in the previous sections to obtain a periodic solution of ordinary differential equations. For the sake of simplicity, we consider the following Duffing equation

$$x'' + Ax' + Bx^3 - C\cos t = 0, t \in J = (0, 2\pi).$$

Let $L_2(0, 2\pi)$ and $H_2(0, 2\pi)$ be the square integrable function's Lesbesgue space and the Sobolev space with norms

$$\|x\|_2 = \sqrt{\frac{1}{2\pi}\int_0^{2\pi} |x(t)|^2\, dt}$$

and

$$\|x\|_{H_2} = \sqrt{\|x\|^2 + \|x'\|^2 + \|x''\|^2},$$

respectively. Let $X = Y = L_2(0, \pi)$. Let us define operators $L : D(L) = \{x | x \in X, x(t) = -x(t + \pi)\} \to L_2(0, 2\pi)$ and $N : D(L) \to L_2(0, 2\pi)$ by

$$Lx = x'' + Ax'$$

and

$$Nx = Bx^3 - C\cos t,$$

respectively. Then, the graph norm associated with L is defined as

$$\|x\|_L = \|x\|_2 + \|x'' + Ax'\|_2.$$

We have

Lemma 1 *For $x \in D(L)$, we can expand x as*

$$x = \sqrt{2}\sum_{n=1}^{\infty}(a_n \cos(2n - 1)t + b_n \sin(2n - 1)t).$$

If we define a projection operator $P_m : D(L) \to \mathbb{R}^{2m+1}$ by

$$P_m x = \sqrt{2}\sum_{n=1}^{m}(a_n \cos(2n - 1)t + b_n \sin(2n - 1)t),$$

we have

$$\|x - P_m x\|_2 \leq \frac{1}{(2m + 1)^2}\sqrt{1 + \frac{A^2}{(2m + 1)^2}}\|Lx\|_2.$$

Proof

Let

$$x'(t) = \sqrt{2}\sum_{n=1}^{\infty}(a'_n \cos(2n - 1)t + b'_n \sin(2n - 1)t)$$

and

$$x'' = \sqrt{2}\sum_{n=1}^{\infty}(a''_n \cos(2n - 1)t + b''_n \sin(2n - 1)t).$$

Then, we have

$$a'_n = (2n - 1)b_n, b'_n = -(2n - 1)a_n,$$

and

$$a''_n = -(2n - 1)^2 a_n, b''_n = -(2n - 1)^2 b_n.$$

Thus if we put

$$x'' + Ax'(t) = \sqrt{2}\sum_{n=1}^{\infty}(\tilde{a}_n \cos(2n-1)t + \tilde{b}_n \sin(2n-1)t),$$

we have

$$\tilde{a}_n = -(2n-1)^2 a_n + (2n-1)Ab_n, \tilde{b}_n = -(2n-1)Aa_n - (2n-1)^2 b_n,$$

or

$$a_n = \frac{-(2n-1)^2\tilde{a}_n - (2n-1)A\tilde{b}_n}{(2n-1)^4 + (2n-1)^2A^2}$$

and

$$b_n = \frac{-(2n-1)^2\tilde{b}_n + (2n-1)A\tilde{a}_n}{(2n-1)^4 + (2n-1)^2A^2}.$$

Let us now consider $\|x - P_m x\|_2^2$. From the Perseval equality, we have

$$\begin{aligned}&\|x - P_m x\|_2^2\\ &\quad = \sum_{n=m+1}^{\infty}(a_n^2 + b_n^2)\\ &\quad \le \sum_{n=m+1}^{\infty}\frac{(2n-1)^4 + (2n-1)^2A^2}{(2n-1)^4 + A^2(2n-1)^2}(\tilde{a}_n^2 + \tilde{b}_n^2)\\ &\quad \le \frac{1}{(2m+1)^4}(1 + \frac{A^2}{(2m+1)^2})\|Lx\|_2^2.\end{aligned}$$

Thus we have the desired inequality. □

Moreover, we have

Lemma 2 *For $x \in H_2(0, 2\pi)$, we have*

$$b\|x\|_L \le \|x\|_{H_2} \le b'\|x\|_L,$$

where

$$b = \frac{1}{\max(1, A)}$$

and

$$b' = \sqrt{2(1 + A^2)}.$$

Proof

From the Perseval equality, we have

$$\begin{aligned}&\|x''\|_2^2\\ &\quad = \sum_{n=1}^{\infty}({a''_n}^2 + {b''_n}^2)\\ &\quad \le \sum_{n=1}^{\infty}\frac{(2n-1)^4((2n-1)^4 + A^2(2n-1)^2)}{[(2n-1)^4 + A^2(2n-1)^2]^2}(\tilde{a}_n^2 + \tilde{b}_n^2)\\ &\quad \le (1 + A^2)\|Lx\|_2^2.\end{aligned} \tag{25}$$

Similarly, we have

$$\|x'\|_2^2 \leq (1+A^2)\|Lx\|_2^2.$$

Thus we have

$$\begin{aligned}
&\|x''\|_2^2 + \|x'\|_2^2 + \|x\|_2^2 \\
&\quad \leq \|x\|_2^2 + 2(1+A^2)\|Lx\|_2^2 \\
&\quad \leq 2(1+A^2)\|Lx\|_L^2, \qquad (26)
\end{aligned}$$

which is the right half of the desired inequalities.

On the other hand, we have

$$\begin{aligned}
&\|x\|_L \\
&\quad = \|x\|_2 + \|x'' + Ax'\|_2 \\
&\quad \leq \|x\|_2 + \|x''\|_2 + A\|x'\|_2 \\
&\quad \leq \max(1, A)\|x\|_{H_2}.
\end{aligned} \qquad (27)$$

This is the desired result. □

Using these result, we have obtained an approximate 2π-periodic solution of the Duffing equation with A=0.1, B=1, and C=0.4464 as follows:

$$\begin{aligned}
x(t) \quad = \quad & 1.27737\cos t + 0.38447\sin t \\
& + 0.05362\cos 3t + 0.0628\sin 3t \\
& + 0.00061\cos 5t + 0.00483\sin 5t \\
& + 0.00013\cos 7t + 0.00000088\sin 9t \\
& + 0.0000010\cos 11t + 0.00000004\sin 11t \\
& + 0.00000005\cos 13t + 0.00000003\sin 13t.
\end{aligned}$$

For this approximate solution, as a result of estimation, we have

$$M = 3.118, r = 0.000000675, K = 6.8682, a = 10.41.$$

From these constants, we have

$$C = 81.17823, p = 0.00011, aCp \leq 0.0925.$$

References

[1] R. Bouc. "Sur la methode de Galerkin-Urabe pour les systemes differentierles periodiques". *Intern. J. Non-Linear Mech.*, 7:175–188, 1972.

[2] L. Cesari. "Functional analysis and periodic solutions of nonlinear equations". *Contributions to differential equations*, 1(2):149–187, 1963.

[3] L. Collatz. *"Functional analysis and numerical mathematics"*. Academic Press, 1966.

[4] Kaucher E.W. and W.L. Miranker. *"Self-validating numerics for function space problems"*. Academic Press, 1984.

[5] L.V. Kantorovich. “Functional analysis and applied mathematics”. *Uspeh. Math. Nauk*, 3:89–185, 1948.

[6] G. Kedem. “A posteriori bounds for two-point boundary value problems”. *SIAM J. Numer. Anal.*, 18(3):431–448, 1981.

[7] M. Nakao. “A numerical approach to the proof of existence of solutions for elliptic problems”. *Japan J. Appl. Math.*, 5:313–332, 1988.

[8] M. Plum. “Computer-assisted existence proofs for two point boundary value problems”. *Computing*, 46:19–34, 1991.

[9] Y. Shinohara. “A geometric method of numerical solutions of nonlinear equations and its application to nonlinear oscillations”. *Publ.RIMS, Kyoto Univ.*, 13, 1972.

[10] Y. Shinohara. “Numerical analysis of periodic solutions and their periods to autonomous differential systems”. *J. Math. Tokushima Univ.*, 11:11–32, 1972.

[11] Y. Shinohara and N. Yamamoto. “Galerkin approximation of periodic solution and its period to van der Pol equation”. *J. Math. Tokushima Univ.*, 12:19–42, 1978.

[12] M. Urabe. “Existence theorems of quasiperiodic solutions to nonlinear differential systems”. *Functional Ekvac.*, 15:75–100, 1972.

[13] M. Urabe. “Galerkin's procedure for nonlinear periodic systems”. *Arch. Rational Mech. Anal.*, 20:120–152, 1965.

[14] M. Urabe. “Numerical investigation of subharmonic solution to Duffing's equation”. *Publ. RIMS, Kyoto Univ.*, 5:79–112, 1969.