JOURNAL OF MATHEMATICAL ANALYSIS AND APPLICATIONS 126, 209–222 (1987)

Approximate Solutions of Problems Involving Normal Operators

W. LAMB, J. MIKA, AND G. F. ROACH

Department of Mathematics, University of Strathclyde, Glasgow, G1 1XH, Scotland, United Kingdom

Submitted by R. P. Boas

Received September 2, 1985

The spectral theory for unbounded normal operators is used to develop a systematic method of approximating functions of operators with other, more easily computable functions, leading to a priori error estimates in the operator norm. In particular, polynomial approximations are obtained for resolvents and semigroups in terms of inverses and resolvents, respectively. © 1987 Academic Press, Inc.

1. INTRODUCTION

A central objective in the analysis of problems in mathematical physics is that of establishing an operator realisation of a given problem in the form

$$Tu = f \in H, \qquad u \in D(T) \subset H \tag{1.1}$$

say, where H is, for instance, a complex Hilbert space. Of crucial importance in any such representation is the requirement that T^{-1} should exist and moreover can be obtained relatively easily. Results in this direction are virtually complete when T has the particular form

$$T = \lambda I - A, \qquad \lambda \in \mathbb{C}, \tag{1.2}$$

where A is a bounded operator of sufficiently small norm. In this case (1.1) yields

$$u = T^{-1}f = (\lambda I - A)^{-1} f = R(\lambda; A) f = \lambda^{-1} \sum_{k=0}^{\infty} \left(\frac{A}{\lambda}\right)^k f, \tag{1.3}$$

where $R(\lambda; A)$ is the resolvent of A. It is well known that a sufficient condition for the convergence of the series in (1.3) is that

$$\|A\| < |\lambda|. \tag{1.4}$$

0022-247X/87 $3.00

Constructive methods based on this approach face two potential difficulties. First, the condition (1.4) might not be satisfied in which case the series in (1.3) might not converge. This can often be overcome by renorming H in terms of suitably weighted norms [1, 7, 8]. Second, the condition (1.4) is satisfied but the parameter λ may have a value near an eigenvalue of A in which case the convergence of the series in (1.3) could be extremely slow. This difficulty can be resolved by using either an eigenvalue shifting technique or a modified Green's function approach [2, 3, 6]. However, an approximation method involving the series in (1.3) fails completely when A is an unbounded operator. This coupled with the fact that not all operators of interest have the form (1.2) indicates the need for alternative approximation procedures.

In this article an approximation method is established which is based on an unexpectedly simple application of the spectral theorem which seems not to have been fully exploited, although similar ideas have already been employed in [5]. The idea behind the method stems from the failure of the convergence of the series in (1.3) when A is unbounded. If A is unbounded and normal then $\sigma(A)$ is unbounded. If furthermore the spectral theorem can be applied to A then we can infer, under suitable conditions, that $\sigma(A^{-1})$ is bounded and that in turn the operator A^{-1} is bounded. This reasoning suggested the present approximation procedure which will be seen to be based on a polynomial involving powers of A^{-1} rather than A as in (1.3). It will be seen that the method presented here reduces, in its zero order form, to that developed in [5] for approximating the inverse of an unbounded normal operator. However, in the method given here it is possible to consider more general functions of an operator than its inverse and to obtain, quite readily, a priori error estimates. In particular we use the method for approximating semigroups generated by unbounded operators in terms of resolvents.

2. Notation and Preliminary Results

Let H be a complex Hilbert space equipped with an inner product $(\cdot, \cdot)$ and induced norm $\|\cdot\|_H$. The sets of bounded operators and of densely defined normal operators in H will be denoted by $B(H)$ and $N(H)$, respectively. For any $A \in N(H)$, there exists a unique resolution of the identity E_A in terms of which A can be expressed uniquely in the form

$$Ax = \int_{\sigma(A)} \lambda dE_A(\lambda)\, x, \qquad x \in D(A),\ \lambda \in \sigma(A), \tag{2.1}$$

where $\sigma(A)$ denotes the spectrum of A and $D(A)$, the domain of A, is given by

$$D(A) = \left\{ x \in H: \int_{\sigma(A)} |\lambda|^2 \, d\|E_A(\lambda)\, x\|_H^2 < \infty \right\}$$

[9]. Strictly (2.1) should be interpreted as

$$(Ax, y) = \int_{\sigma(A)} \lambda d(E_A(\lambda)\, x, y), \qquad x \in D(A),\ y \in H,$$

but we shall adopt the standard practice of writing either (2.1) or, when $A \in B(H)$, the abbreviated version

$$A = \int_{\sigma(A)} \lambda dE_A(\lambda). \tag{2.2}$$

Unless otherwise stated, any operator A will be assumed to be normal on H and consequently it will possess a unique spectral decomposition of the form (2.1). We note in passing that A will be bounded whenever $\sigma(A)$ is a bounded set and, since $A \in N(H)$,

$$\|A\| = \sup\{|\lambda|: \lambda \in \sigma(A)\},$$

where $\|\cdot\|$ denotes the usual operator norm [9].

It is well known that the resolution of the identity E_A associated with a given operator A on H can be used to construct other normal operators $f(A)$ on H for a wide class of scalar functions f. This is done in a natural way by defining $f(A)$ to be the operator satisfying

$$f(A)\, x = \int_{\sigma(A)} f(\lambda)\, dE_A(\lambda)\, x, \qquad x \in D(f(A)), \tag{2.3}$$

where [9],

$$D(f(A)) = \left\{ x \in H: \int_{\sigma(A)} |f(\lambda)|^2 \, d\|E_A(\lambda)\, x\|_H^2 < \infty \right\}. \tag{2.4}$$

In this paper we shall restrict attention to functions f which are continuous and bounded on $\sigma(A)$, so that $f(A) \in B(H)$ and

$$\|f(A)\| = \sup\{|f(\lambda)|: \lambda \in \sigma(A)\}. \tag{2.5}$$

Of particular importance later are the functions $\theta(t; \lambda) = e^{\lambda t}$, $t \geqslant 0$. If $\sigma(A) \subseteq \{\lambda: \operatorname{Re} \lambda \leqslant \omega\}$, $\omega < \infty$, then for each $t \geqslant 0$, $\theta(t; \lambda)$ is a continuous

function of λ on $\sigma(A)$ and therefore a family of bounded normal operators $\{T(t; A); t \geqslant 0\}$ can be defined by

$$T(t; A) = \int_{\sigma(A)} \theta(t; \lambda)\, dE_A(\lambda), \qquad t \geqslant 0. \tag{2.6}$$

Routine calculations establish that $\{T(t; A); t \geqslant 0\}$ is a (C_0) semigroup of type ω with the operator A being the infinitesimal generator [9]. A converse to this result also holds in that any (C_0) semigroup of bounded normal operators $\{T(t; A); t \geqslant 0\}$ has a normal infinitesimal generator A with $\sigma(A) \subseteq \{\lambda: \operatorname{Re} \lambda \leqslant \omega\}$, for some finite real constant ω, and, [9], can be expressed by (2.6).

Finally, we state the following result concerning the spectral representation of the resolvent of an operator $A \in N(H)$.

LEMMA 2.1. *If $\mu \in \rho(A)$, the resolvent set of A, then the resolvent operator $R(\mu; A) = (\mu I - A)^{-1}$ can be written as*

$$R(\mu; A) = \int_{\sigma(A)} r(\mu; \lambda)\, dE_A(\lambda), \tag{2.7}$$

where $r(\mu; \lambda) = (\mu - \lambda)^{-1}$.

3. An Approximation Procedure

In applications, the problem of determining an operator $f(A)$, associated with a given operator A and some function f, often arises. Commonly occurring examples of f are the functions $r(\mu; \lambda) = (\mu - \lambda)^{-1}$, $\theta(t; \lambda) = e^{\lambda t} (t \geqslant 0)$ and $f_\alpha(\lambda) = (-\lambda)^\alpha$ which lead, respectively, to the operators $R(\mu; A) = (\mu I - A)^{-1}$ (the resolvent operator), $\theta(t; A) = e^{tA}$, $t \geqslant 0$, (semigroups) and $f_\alpha(A) = (-A)^\alpha$ (fractional powers).

Frequently it is impossible to obtain an analytic representation of $f(A)$ which can be used for calculating $f(A)x$, for a given $x \in H$, and consequently approximations for $f(A)$ have to be constructed.

In this section we introduce a method which reduces the problem of finding a suitable approximation of $f(A)$ to the more familiar problem of constructing approximations of scalar-valued functions. As will be seen, this particular method has the additional advantage of automatically providing a priori error estimates.

The problem we shall consider can be stated as follows: given an operator $f(A) \in B(H)$ and some constant $\varepsilon > 0$, find $\phi(A) \in B(H)$, an approximation to $f(A)$, which satisfies $\|f(A) - \phi(A)\| \leqslant \varepsilon$.

Throughout we shall assume that the functions f and ϕ are continuous and bounded on $\sigma(A)$. Therefore, from (2.5), we have

$$\|f(A)-\phi(A)\| = \sup\{|f(\lambda)-\phi(\lambda)|: \lambda \in \sigma(A)\}. \tag{3.1}$$

In applications, the function ϕ will depend often upon several parameters (i.e., $\phi = \phi(\lambda; a_1,..., a_n)$) and be constrained by the fact that it should be more suited to computational purposes than the original function f. These factors usually lead to the reformulation of the above approximation problem as a minimisation problem of the type

$$\min\{\sup\{|f(\lambda)-\phi(\lambda; a_1,..., a_n)|: \lambda \in \sigma(A)\}: a_1,..., a_n \in \Lambda\} \tag{3.2}$$

where Λ denotes an admissible set of parameters.

When A is a bounded operator, an obvious choice for ϕ would be a polynomial in λ of the form $\phi(\lambda) = \sum_{k=0}^{n} a_k \lambda^k$, in which case the minimisation problem becomes

$$\min\left\{\sup\left\{\left|f(\lambda)-\sum_{k=0}^{n} a_k\lambda^k\right|: \lambda \in \sigma(A)\right\}: a_0, a_1,..., a_n \in \mathbb{C}\right\}. \tag{3.3}$$

For self-adjoint operators, the spectrum $\sigma(A)$ lies on the real axis and the problem (3.3) becomes a well-known problem of contructing an optimal polynomial approximating a given continuous function, in the maximum norm. It is well known that there exist such polynomials for any fixed degree n although for $n \geqslant 2$ the general algorithm for obtaining such a polynomial is not known [10], and a numerical approach is found to be necessary.

In the more general case when the operator is normal but not self-adjoint, the situation becomes more complicated since the techniques associated with polynomial approximations of functions of a complex variable are less well developed. However, we would mention that in [5] a normal operator having a spectrum lying off the real axis is considered. Since, for unbounded operators, the method presented here reduces, in a zero order form, to that developed in [5] it is felt that it offers good prospects for dealing with quite general normal operators.

The minimisation problem (3.3) together with the spectral theorem indicates that a suitable approximation for $f(A)$ can be obtained in terms of powers of A, provided A is bounded. In the case when A is unbounded any polynomial in A will also be unbounded unless the polynomial is zero order [5]. However, the spectral theorem indicates that for a given normal operator A and f a function of a complex variable, analytic in a neighbourhood of $\sigma(A)$ we have the result

$$\sigma(f(A)) = f(\sigma(A)).$$

This suggests that a means of avoiding the above difficulties associated with unbounded operators is to seek polynomial approximations in terms of the resolvent operator $R(\mu; A)$, $\mu \in \rho(A)$ rather than A. In this case (3.2) becomes

$$\min \left\{ \sup \left\{ \left| f(\lambda) - \sum_{k=0}^{n} a_k (\mu - \lambda)^{-k} \right| : \lambda \in \sigma(A) \right\} : a_0, a_1, \ldots, a_n \in \mathbb{C} \right\} \tag{3.4}$$

with the further possibility of minimising with respect to $\mu \in \rho(A)$.

As an example of the use of (3.4) consider the case when f has the particular form defined by $\theta(t; \lambda) = e^{\lambda t}$ and A is the generator of a (C_0) semigroup. The minimisation problem presented by (3.4) may now be regarded as that of determining an approximation $\tilde{u}_n(x, t)$ to the solution $u(x, t)$ of the initial-boundary value problem

$$\frac{\partial}{\partial t} u(x, t) = Au(x, t), \qquad x \in (a, b) \subset \mathbb{R}, \; t \in (0, \infty) \tag{3.5}$$

$$u(x, 0) = u_0(x) \tag{3.6a}$$

$$u(a, t) = g_1(t) \tag{3.6b}$$

$$u(b, t) = g_2(t). \tag{3.6c}$$

The approximate solution is sought in the form

$$\tilde{u}_n(x, t) = \sum_{k=0}^{n} a_k(t)\, u_k(x), \tag{3.7}$$

where the $u_k(x)$ satisfy the time-independent problems

$$(\mu I - A)\, u_k(x) = u_{k-1}(x), \qquad k = 1, 2, \ldots, n. \tag{3.8}$$

The error at any stage of approximation can be obtained by first noting that formally (3.5) has a solution of the form

$$u(x, t) = T(t; A)\, u_0(x).$$

An approximation to this solution will be given by

$$\tilde{u}_n(x, t) = \sum_{k=0}^{n} a_k(t)(\mu I - A)^{-k}\, u_0(x) = \phi_n(t; A)\, u_0(x).$$

Hence, from (2.3), (2.6), and (3.1)

$$\|u-\tilde{u}_n\| \leqslant \|T(t;A)-\phi_n(t;A)\|\,\|u_0(x)\|$$
$$= \|u_0(x)\| \min\left\{\sup\left\{\left|e^{\lambda t}-\sum_{k=0}^{n} a_k(t)(\mu-\lambda)^{-k}\right| : \lambda\in\sigma(A)\right\} : a_k\in\mathbb{C}\right\}. \tag{3.9}$$

This procedure offers a means of approximating a time-dependent problem by an associated time-independent problem. Alternatively it provides a means of approximating a time-dependent Green's function by a time-independent one, the latter being often much easier to calculate, relatively speaking.

4. Approximations of Resolvents

As a first illustration of the method described in the previous section, the problem of approximating the resolvent $(I+A)^{-1}$ of a self-adjoint operator A is considered. This problem has already been examined in [5] for the case when A is bounded with spectrum confined to the interval $[0, M]$, $M<1$. For such operators A, the zero and first order optimal polynomial approximations of $(I+A)^{-1}$ were found to be

$$\begin{aligned} p_0^{\text{opt}}(A) &= \tfrac{1}{2}(2+M)/(1+M)\,I \\ p_1^{\text{opt}}(A) &= [(1+M)^{-1/2}+\tfrac{1}{2}M(1+M)^{-1}]\,I-(1+M)^{-1}\,A, \end{aligned} \tag{4.1}$$

the latter being obtained by means of a simple application of the following lemma [10].

Lemma 4.1. *Let f be a twice continuously differentiable function defined on the interval $[a, b]$ and assume that f'' does not change sign in $[a, b]$. Then the first order polynomial of best approximation of f over the interval $[a, b]$ is given by*

$$p_1^{\text{opt}}(\lambda) = \alpha_0+\alpha_1\lambda,$$

where

$$\alpha_0 = \frac{1}{2(b-a)}\,[(b+c)\,f(a)-(c+a)\,f(b)+(b-a)\,f(c)]$$

$$\alpha_1 = \frac{1}{b-a}\,[f(b)-f(a)]$$

and $c \in (a, b)$ satisfies the equation $f'(c) = \alpha_1$. The error is given by

$$\varepsilon^{(1)} = \frac{1}{2(b-a)} |(b-c) f(a) + (c-a) f(b) + (a-b) f(c)|. \tag{4.2}$$

From (2.5), it follows that the error associated with the linear approximation is

$$\|(I+A)^{-1} - p_1^{\text{opt}}(A)\| = \sup\{|(1+\lambda)^{-1} - p_1^{\text{opt}}(\lambda)|; 0 \leqslant \lambda \leqslant M\} = \varepsilon^{(1)},$$

where $\varepsilon^{(1)}$, given by (4.2) with $a = 0$, $b = M$, $f(\lambda) = (1+\lambda)^{-1}$ and $c = (1+M)^{1/2} - 1$, can be expressed as

$$\varepsilon^{(1)} = \tfrac{1}{2}[1 - (1+M)^{-1/2}]^2. \tag{4.3}$$

The error $\varepsilon^{(0)}$ involved in the zero order approximation is easily calculated as

$$\varepsilon^{(0)} = \|(I+A)^{-1} - p_0^{\text{opt}}(A)\| = \tfrac{1}{2} M/(1+M).$$

If we now remove the boundedness restriction on A and consider instead the problem of approximating $(I+A)^{-1}$ for a self-adjoint operator A with $\sigma(A) = [1/M, \infty)$, $0 < M < 1$, then similar arguments lead to an optimal zero order polynomial approximation given by $\frac{1}{2}MI/(1+M)$ with associated error $\varepsilon^{(0)} = \frac{1}{2}M/(1+M)$. However, for higher order approximations the method fails since polynomials in A are unbounded whenever A is. This difficulty can be overcome by first noting that in this case A^{-1} exists as a bounded operator since, by assumption, $0 \notin \sigma(A)$, and then seeking higher order approximations in terms of polynomials in A^{-1}. In the first order case, this means that we must attempt to minimise the error

$$\begin{aligned} \varepsilon^{(1)} &= \|(I+A)^{-1} - a_0 I - a_1 A^{-1}\| \\ &= \sup\{|(1+\lambda)^{-1} - a_0 - a_1/\lambda|; 1/M \leqslant \lambda < \infty\}. \end{aligned}$$

A simple change of variable reduces this problem to one of finding the optimal first order polynomial approximating the function $\lambda/(1+\lambda)$ over the interval $[0, M]$. This can be constructed either by a direct application of Lemma 4.1 or by noting that

$$\begin{aligned} &\sup\{|\lambda/(1+\lambda) - a_0 - a_1\lambda|; 0 \leqslant \lambda \leqslant M\} \\ &\quad = \sup\{|(1-a_0) - a_1\lambda - (1+\lambda)^{-1}|; 0 \leqslant \lambda \leqslant M\} \end{aligned}$$

and therefore, from (4.1) and (4.3),

$$p_1^{\text{opt}}(A^{-1}) = [\tfrac{1}{2}(2+M)/(1+M) - (1+M)^{-1/2}]\, I + (1+M)^{-1} A^{-1} \tag{4.4}$$

with the error $\varepsilon^{(1)}$ given by (4.3).

TABLE I

M	0.9	0.7	0.5	0.2	0.1	0.05
$\varepsilon^{(1)}$	0.03768	0.02715	0.01684	0.003796	0.001083	0.0002904
$\varepsilon^{(\mathrm{Neu})}$	0.3837	0.2018	0.0833	0.006667	0.0009091	0.0001190
$\varepsilon^{(1)}/\varepsilon^{(\mathrm{Neu})}$	0.0982	0.1345	0.2022	0.5694	1.191	2.440

In Table I, a list of errors is given for different values of M. A comparison is also made with the corresponding errors $\varepsilon^{(\mathrm{Neu})}$ obtained by writing $(I+A)^{-1}=(I+A^{-1})^{-1}A^{-1}$ and using the first two terms in the Neumann expansion of $(I+A^{-1})^{-1}$ to obtain an approximation of $(I+A)^{-1}$ in the form $A^{-1}-A^{-2}$. This error is given by $\varepsilon^{(\mathrm{Neu})}=M^3(1+M)^{-1}$.

It is seen at once from Table I that, for a range of values of M, the optimal polynomial (4.4) is a better approximation of $(I+A)^{-1}$ than that obtained from the Neumann series, with a substantially smaller error being achieved for values of M which lie close to one. Furthermore, at values of M where the Neumann approximation is superior, little advantage is to be gained by its use since the errors involved in either approximation are small. More significantly, the small values of $\varepsilon^{(\mathrm{Neu})}$ are attained at the expense of using a non-optimal approximating polynomial of *second* order. The simple device of replacing $(I+A^{-1})^{-1}$ by $p_1^{\mathrm{opt}}(A^{-1})$, constructed using (4.1), produces, without any extra expense, a greatly improved second order approximation of $(I+A)^{-1}$ in the form $p_1^{\mathrm{opt}}(A^{-1})A^{-1}$ with associated error being at most $M\varepsilon^{(1)}$, where $\varepsilon^{(1)}$ is given by (4.3).

Examples where the above procedure may prove useful are provided by taking A to be proportional to either the Laplace or wave operators. In these cases, $I+A$ would be related to the Helmholtz and Klein–Gordon equations, respectively.

5. Approximations of Semigroups

To investigate the suitability of the above method for approximating semigroups of operators, we consider again a self-adjoint operator A in a Hilbert space H defined by

$$\begin{aligned} Ax &= \int_{-\infty}^{\omega} \lambda dE_A(\lambda)\, x, \qquad x \in D(A), \quad \omega < \infty \\ D(A) &= \left\{ x \in H: \int_{-\infty}^{\omega} |\lambda|^2\, d\|E_A(\lambda)\, x\|_H^2 < \infty \right\}. \end{aligned} \tag{5.1}$$

Obviously, A is an infinitesimal generator of a (C_0) semigroup $\{T(t; A); t \geqslant 0\}$ with index ω. Moreover, if we define $\tilde{A}$ by

$$\tilde{A}x = (A - \omega I)\, x, \qquad x \in D(A), \tag{5.2}$$

then $\tilde{A}$ generates the semigroup $\{\tilde{T}(t; A) = e^{-\omega t} T(t; A); t \geqslant 0\}$ with index 0. Therefore, without loss of generality, we assume that $\omega = 0$ in (5.1) so that

$$Ax = \int_{-\infty}^{0} \lambda dE_A(\lambda)\, x, \qquad x \in D(A)$$

and

$$T(t; A) = \int_{-\infty}^{0} e^{\lambda t}\, dE_A(\lambda).$$

In this case, neither A nor A^{-1} exists as a bounded operator and therefore we seek an approximation of $T(t; A)$, for fixed t, in the form of a polynomial in the resolvent operator $R(\mu; A)$, for some fixed $\mu > 0$. The error associated with this approximation is given by

$$\begin{aligned} \varepsilon^{(n)} &= \left\| T(t; A) - \sum_{k=0}^{n} \alpha_k (\mu I - A)^{-k} \right\| \\ &= \sup \left\{ \left| e^{\lambda t} - \sum_{k=0}^{n} \alpha_k (\mu - \lambda)^{-k} \right| ; \lambda \leqslant 0 \right\} \\ &= \sup \left\{ \left| e^{m(1 - 1/\omega)} - \sum_{k=0}^{n} \alpha_k \left(\frac{\omega}{\mu} \right)^k \right| ; 0 < \omega \leqslant 1 \right\}, \end{aligned} \tag{5.3}$$

where $m = \mu t$ and $\omega = \mu(\mu - \lambda)^{-1}$. This error will be a minimum, for a given μ and n, if $\sum_{k=0}^{n} \alpha_k (\omega/\mu)^k$ is the optimal approximating polynomial of order n of the function $f(\cdot; m)$, defined by

$$f(\omega; m) = \begin{cases} \exp(m[1 - 1/\omega]), & 0 < \omega \leqslant 1 \\ 0, & \omega = 0 \end{cases} \tag{5.4}$$

over the interval $[0, 1]$.

When $n = 0$, we obtain trivially

$$\alpha_0 = \tfrac{1}{2}, \qquad \varepsilon^{(0)} = \tfrac{1}{2},$$

the value in each case being independent of $\mu > 0$.

When $n = 1$, we note first that, for the function f defined by (5.4),

$$\frac{df}{d\omega} = mf/\omega^2 \qquad \text{and} \qquad \frac{d^2 f}{d\omega^2} = m(m - 2\omega)\, f/\omega^4.$$

The second derivative is clearly positive in $(0, 1]$ whenever μ and t are such that $m = \mu t > 2$ and consequently, for such values of m, the optimal linear polynomial in $R(\mu; A)$ approximating $T(t; A)$ can be found on applying Lemma 4.1. In addition, for $m \geqslant 1$, the equation

$$\frac{df}{d\omega}(\omega; m) = f(1; m) - f(0; m),$$

which reduces to

$$\frac{m}{\omega^2} e^{m(1-1/\omega)} = 1, \tag{5.5}$$

has only one solution in $(0, 1)$, since $(df/d\omega)\ (1; m) = m \geqslant 1$. This means that the algorithm presented in Lemma 4.1 can, in fact, be used to construct the optimal linear approximation for any $m \geqslant 1$, and yields the following formulae for α_0, α_1 and $\varepsilon^{(1)}$ in terms of m, t, and c, where $c \in (0, 1)$ is the solution of (5.5),

$$\alpha_0 = \tfrac{1}{2}(c^2/m - c)$$

$$\alpha_1 = \mu = m/t$$

$$\varepsilon^{(1)} = -\alpha_0 = \tfrac{1}{2}(c - c^2/m).$$

The results of calculations, listed in Table II, indicate that the minimum value of $\varepsilon^{(1)}$ is approximately 0.102 occurring when $m = 1$ and this implies that the optimal first order approximation formula is

$$T(t; A) \doteqdot -0.102I + t^{-1}R(t^{-1}; A) = -0.102I + (I - tA)^{-1}. \tag{5.6}$$

When $0 < m < 1$, Eq. (5.5) has two roots in $(0, 1)$ and Lemma 4.1 is no longer applicable. However, preliminary calculations suggest that the error values $\varepsilon^{(1)}$ for such m exceed the value of $\varepsilon^{(1)}$ at $m = 1$.

Two important features of the above approximation method should be noted. First, the remarkable fact that an error of the order of only 10% can be achieved using such a crude approximation as that represented by

TABLE II

m	1.0	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	2.0	2.5	3.0
$\varepsilon^{(1)}$[a]	0.102	0.112	0.121	0.130	0.140	0.147	0.155	0.163	0.170	0.176	0.183	0.211	0.233
m	3.5	4.0	4.5	5	6	7	8	9	10	15	20	25	30
$\varepsilon^{(1)}$	0.252	0.267	0.281	0.292	0.312	0.328	0.340	0.351	0.360	0.392	0.411	0.424	0.433

[a] $\varepsilon^{(1)} = \|T(t) - a_0 I - t^{-1}R(t^{-1}; A)\|$, $a_0 = -\varepsilon^{(1)}$.

(5.6). This raises expectations that the error would decrease rapidly for higher order approximations. Second, the coefficients which appear in the approximation formula (5.6), and in similar expressions involving higher order terms, are universally valid for *all* dissipative self-adjoint operators since indices ω which differ from zero can be accounted for in a simple way as in (5.2).

6. An Application

As a simple demonstration of the applicability of formula (5.6), we consider the problem of finding an approximation, for fixed $t>0$, of the solution u of the following Cauchy problem involving the one-dimensional heat operator in an infinite medium

$$\frac{\partial u}{\partial t}(x, t)=\frac{\partial^2 u}{\partial x^2}(x, t), \qquad x\in\mathbb{R},\ t>0 \tag{6.1}$$

$$u(x, 0)=u_0(x), \qquad x\in\mathbb{R}. \tag{6.2}$$

In (6.2), u_0 is restricted to the class of functions

$$X=\{f: f, f' \text{ are absolutely continuous and } f, f', f''\in L^2(\mathbb{R})\}. \tag{6.3}$$

The closed-form solution of (6.1) and (6.2) is well known and can be expressed as

$$\begin{aligned} u(x, t)&=\tfrac{1}{2}(\pi t)^{-1/2}\int_{-\infty}^{\infty} e^{-(x-y)^2/4t}u_0(y)\,dy, && t>0,\\ &=(G_t * u_0)(x), && t>0, \end{aligned}$$

where G_t, defined by

$$G_t(x)=\tfrac{1}{2}(\pi t)^{-1/2}\,e^{-x^2/4t}, \qquad t>0,\ x\in\mathbb{R},$$

is the Gauss–Weierstrass kernel.

If we define a family of operators $\{T(t; A); t\geqslant 0\}$ on $L^2(\mathbb{R})$ by

$$T(t; A)f=\begin{cases} G_t * f, & t>0\\ f, & t=0, \end{cases}$$

then $\{T(t; A); t\geqslant 0\}$ is a (C_0) semigroup of bounded operators on $L^2(\mathbb{R})$ with index 0 and infinitesimal generator A defined by

$$Af=f'', \qquad f\in D(A)=X,$$

where X is given by (6.3). Since the resolvent operator $R(\mu; A)$ associated with A is defined on $L^2(\mathbb{R})$ by

$$R(\mu; A) f = R_\mu * f \qquad (\mu > 0),$$

where

$$R_\mu(x) = \tfrac{1}{2}\mu^{-1/2} e^{-\mu^{1/2}x \operatorname{sgn} x},$$

we are led, via (5.6), to the following approximation $\tilde{u}$ for u at fixed t:

$$\tilde{u}(x, t) = -0.102 u_0(x) + \frac{1}{t}(R_{1/t} * u_0)(x).$$

The error involved in this approximation is

$$\varepsilon = \|\tilde{u}(\cdot, t) - u(\cdot, t)\|_2 \leqslant 0.102 \|u_0\|_2,$$

where $\| \ \|_2$ denotes the usual norm on $L^2(\mathbb{R})$.

It is worth remarking that, in this particular example, formula (5.6) can be deduced using more familiar methods. To see this, we note initially that the Fourier transform F and its inverse F^{-1} are isometric isomorphisms of $L^2(\mathbb{R})$ onto $L^2(\mathbb{R})$ and, moreover,

$$T(t; A) f = F^{-1} \hat{G}_t F f, \qquad \hat{G}_t = (2\pi)^{-1/2} e^{-\omega^2 t},$$
$$R(\mu; A) f = F^{-1} \hat{R}_\mu F f, \qquad \hat{R}_\mu = (2\pi)^{-1/2} (\mu + \omega^2)^{-1}.$$

Consequently,

$$\begin{aligned} &\| T(t; A) - (2\pi)^{-1/2} a_0 I - a_1 R(\mu; A) \| \\ &\quad = \sup\{ \| F^{-1}(\hat{G}_t - (2\pi)^{-1/2} a_0 - a_1 \hat{R}_\mu) F f \|_2 ; f \in L^2(\mathbb{R}), \| f \|_2 = 1 \} \\ &\quad = \operatorname{ess\,sup}\{ |\hat{G}_t(\omega) - (2\pi)^{-1/2} a_0 - a_1 \hat{R}_\mu(\omega)| ; \omega \in \mathbb{R} \} \\ &\quad = (2\pi)^{-1/2} \sup\{ |e^{\lambda t} - a_0 - a_1(\mu - \lambda)^{-1}| ; \lambda \leqslant 0 \} \quad (\lambda = -\omega^2). \end{aligned} \tag{6.4}$$

The problem of minimising (6.4) is clearly identical to that of minimising $\varepsilon^{(1)}$ given by (5.3) and therefore, proceeding as in Section 5, we are led once more to formula (5.5).

Although the above analysis involved only routine application of the Fourier transform and led to the same minimisation problem as that obtained using spectral theory, it should be emphasised that the Fourier transform is unsuitable for dealing with equations with non-constant coefficients. The method described in the earlier sections, however, can, in theory, cater for both constant and non-constant coefficient cases.

REFERENCES

1. S. HEIKKILA AND G. F. ROACH, On equivalent norms and the contraction mapping principle, *J. Nonlinear Anal. TMA* **8** No. 10 (1984), 1241–1252.
2. R. E. KLEINMAN AND G. F. ROACH, On modified Green's functions in exterior problems for the Helmholtz equation, *Proc. Roy. Soc. London Ser. A* **383** (1982), 313–332.
3. R. E. KLEINMAN AND G. F. ROACH, Operators of minimal norm via modified Green's functions, *Proc. Roy. Soc. Edinburgh, Ser. A* **94** (1983), 163–178.
4. R. KRESS AND G. F. ROACH, On the convergence of successive approximations for an integral equation in a Green's function approach to the Dirichlet problem, *J. Math. Anal. Appl.* **55** (1976), 102–111.
5. J. MIKA AND D. C. PACK, Approximations to inverses of normal operators, *Proc. Roy. Soc. Edinburgh, Ser. A* **103** (1986), 335–345.
6. G. F. ROACH, On the approximate solution of elliptic self-adjoint boundary value problems, *Arch. Ration. Mech. Anal.* **27** No. 3 (1967), 243–254.
7. G. F. ROACH, Approximate Green's functions and the solution of related integral equations, *Arch. Ration. Mech. Anal.* **36** No. 1 (1970), 79–88.
8. G. F. ROACH, "Weighted Norms and the Contraction Mapping Principle," pp. 1–11, Univ. Gottingen NAM-Bericht No. 10, 1974.
9. W. RUDIN, "Functional Analysis," McGraw–Hill, New York, 1973.
10. J. TODD (Ed.), "Survey of Numerical Analysis," McGraw–Hill, New York, 1962.