

COMPUTING THE DECAY OF A SIMPLE REVERSIBLE SUB-MARKOV SEMIGROUP

E. B. DAVIES

Abstract

Two different numerical methods for solving a non-self-adjoint evolution equation are compared in this paper. If the initial function lies in the domain of the operator, a recently proposed method that combines pseudospectral ideas and semigroup theory is shown to be considerably more accurate than a standard discretization method. One example is worked out in detail, but the methods used are of much wider applicability.

1. Introduction

We consider the evolution equation

$$f'_t = Af_t, \tag{1}$$

where

$$(Af)(x) = b^{-1}f''(x) - f'(x)$$

subject to Dirichlet boundary conditions, and $x \in (0, a)$. We start by assuming that the equation acts in the space $C_0(0, a)$ of continuous functions on $[0, a]$ that vanish at $0, a$, with the uniform (that is, supremum) norm $\| \cdot \|_\infty$. The operator A has eigenvalues

$$\lambda_n = -\frac{b}{4} - \frac{\pi^2 n^2}{a^2 b}, \quad \text{for } n = 1, 2, \dots,$$

with corresponding eigenfunctions

$$\phi_n(x) = e^{b(x-a)/2} \sin(\pi nx/a).$$

The solution of the evolution equation is given by $f_t = T_t f_0$, where $T_t = e^{At}$ is a strongly continuous, reversible, sub-Markov semigroup on $C_0(0, a)$.

It is known that there is a non-negative continuous kernel $K(t, x, y)$ on $(0, \infty) \times [0, a] \times [0, a]$ such that

$$(T_t f)(x) = \int_0^a K(t, x, y) f(y) dy$$

for all $t > 0$ and $f \in C_0(0, a)$. Moreover, $K(t, x, y) = 0$ if $x = 0, a$ or $y = 0, a$, because of the Dirichlet boundary conditions. This formula may be used to extend the semigroup to a variety of other spaces, including $L^2((0, a), dx)$. If we use the formula to extend T_t to $C[0, a]$, then the semigroup is not strongly continuous at $t = 0$ because the Dirichlet boundary conditions imply that $T_t f \in C_0(0, a)$ for all $t > 0$.

If one considers the semigroup as acting in $L^2((0, a), e^{-bx} dx)$, then $A = A^*$ and

$$\|T_t\|_2 = e^{\lambda_1 t} \quad \text{for all } t \geq 0.$$

Received 14 October 2003 , revised 2 January 2004; published 26 February 2004.

2000 Mathematics Subject Classification 65Mxx, 60J22, 47D06, 34L05

© 2004, E. B. Davies

The situation in $C_0(0, a)$ is very different. The operator T_t is a positivity-preserving contraction for all $t \geq 0$. The function

$$n(t) := \|T_t\|_\infty = \sup\{(T_t 1)(x) : 0 < x < 1\} \quad (2)$$

satisfies $0 \leq n(t) \leq 1$ for all $t \geq 0$, and is a monotonically decreasing function. It is also submultiplicative in the sense that

$$n(s + t) \leq n(s)n(t)$$

for all $s, t \geq 0$. Moreover,

$$n(t) \sim ce^{\lambda_1 t} \quad (3)$$

as $t \rightarrow \infty$. In probabilistic terms, T_t^* , considered as acting in $L^1((0, a), dx)$, describes a drift to the left with speed 1 combined with a diffusion with parameter b and killing at $0, a$. If b is large, one should expect $n(t) \sim 1$ for $0 \leq t < a$ and $n(t) \sim 0$ for $t > a$. Indeed, for the pure drift operator ($b = +\infty$), one has exactly such behaviour.

One goal of this paper is to compare different numerical procedures for evaluating $n(t)$. This semigroup was already considered in [4], where we explained the possible advantages of using a pseudospectral method for numerical computations. In this paper, we investigate the example in much more detail. If we put $a = 10$ and $b = 20$, then it is sufficiently singular to be interesting, but also exactly soluble, so that one can compare the different methods with the correct values. Our conclusions are explained further in Section 11, but may be summarized as follows. If the initial function f does not lie in the domain of the operator, then all the methods considered for computing $e^{At}f$ are slow or unreliable, or depend upon using very high-precision arithmetic. This is not because of the failure of standard algorithms described in [9, p. 384] and [11] for computing the exponentials of large matrices, but because of the difficulty of choosing an appropriate basis. For more regular f , however, the pseudospectral method is more accurate than that based upon the most common type of discretization. We accept that more-refined discretizations do better, but we expect that our new pseudospectral method will also be improved once its potential has been established.

2. The spectral method

The calculations in this section use the spectral expansion

$$(T_t 1)(x) = \sum_{n=1}^{\infty} \frac{e^{\lambda_n t} \langle 1, \phi_n^* \rangle}{\langle \phi_n, \phi_n^* \rangle} \phi_n(x), \quad (4)$$

where

$$\phi_n^*(x) = e^{-bx/2} \sin(\pi nx/a).$$

This method is not generally to be recommended. However, because all the relevant integrals can be written down in closed form, one can easily evaluate (4), *provided that* one is prepared to use very high-precision arithmetic. As a or b increases, the number of digits needed to carry out the computation increases rapidly. The computations in this section are used to provide a basis for assessing the other methods to be presented. Direct calculations show that

$$\langle 1, \phi_n^* \rangle = a\pi n \frac{1 - (-1)^n e^{-ba/2}}{b^2 a^2 / 4 + \pi^2 n^2}$$

and

$$\langle \phi_n, \phi_n^* \rangle = (a/2)e^{-ba/2}.$$

We summed the first 100 terms of the series (4) using Maple. Precise values of $n(t)$, obtained using the above spectral expansion, are given in Table 1.

The function $n(t)$ was evaluated with Digits := 20 and then re-evaluated with Digits := 25, 30, 35, . . . until it stabilized; the smallest stable value was recorded. We found that 100 terms were enough to make the sum stable for all the values of t considered except $t = 1$, when 150 terms were used. The column of Table 1 labelled x provides the value of x at which the supremum in (2) occurs. For $t \leq 4$, the value of x is not accurate: the function concerned is almost equal to 1 over a substantial proportion of the interval $(0, a)$.

We see, as expected, that there is little decay before $t = 10$, after which the norm decreases rapidly. We adapt one of the standard definitions of the mixing time τ of a Markov chain to our sub-Markov context, by taking it to be the solution of $n(\tau) = 1/4$; see [1, Chapter 4]. Our computations then show that $10 < \tau < 11$.

We emphasize that even if $t > \tau$, the decay is *not* well approximated by the formula (3) within the range of t considered. Table 2 presents values of

$$r(t) = \log(n(t)) - \log(n(t - 1)) \tag{5}$$

for various values of t .

Table 1: Values of $n(t)$ computed using the spectral method; see Appendix A.1.

t	$n(t)$	x	Digits
1	1.00000	8	45
2	1.00000	8	40
4	1.00000	8	35
6	0.99999	9.3	35
8	0.96593	9.722	35
10	0.39454	9.835	25
15	0.897×10^{-5}	9.88	20
20	0.188×10^{-12}	9.89	20
30	0.446×10^{-31}	9.90	20
40	0.259×10^{-51}	9.90	20

Table 2: Values of $r(t)$ as defined by (5); see Appendix A.2.

t	$r(t)$
20	-3.84938
30	-4.52004
40	-4.74733
50	-4.84919

According to the spectral expansion,

$$\lim_{t \rightarrow \infty} r(t) = \lambda_1 \sim -5.00493,$$

but we see that the convergence to this value is very slow. This implies that up to the values of t considered, many terms in the series are contributing towards the value of $n(t)$. One enters the asymptotic region of exponential decay only once $e^{\lambda_2 t} \ll e^{\lambda_1 t}$. This requires $3\pi^2 t > a^2 b$ or, given our choices of the constants, $t > 70$. We conclude that the formula (3) is useless in any practical sense; the size of $n(t)$ is unmeasurably small by the time the formula becomes relevant.

We also computed the values of $m(t) = \|e^{At} f\|_\infty$, for $f(x) = \sin(\pi x/a)$, by the same method. This function does not have the same intrinsic interest as $n(t)$, but will be used to study the extent to which the accuracy of the methods depends upon whether f is or is not in the domain of A . The results are presented in Table 3, in which x is the value at which $(e^{At} f)(x)$ takes its maximum. Once again, the numerical evidence indicates that the values are correct to the accuracy presented.

Table 3: Values of $m(t)$ computed using the spectral method; see [Appendix A.3](#).

t	$m(t)$	x
1	0.99508	6.0000
2	0.99018	7.0000
4	0.98045	9.0000
6	0.89058	9.7460
8	0.50619	9.8060
10	0.09263	9.8486

3. The discretization method

We compare the values of $n(t)$ and of $m(t)$, obtained above, with those obtained by a discretization of the differential operator A . The main issue is deciding how to replace A by a matrix. There is a huge literature on such matters, describing higher-order finite differences, higher-order finite elements, non-uniform grids, and so forth, but the difficulties that we will encounter apply to *all* of them. For this reason, we treat only the simplest discretization in detail. We also emphasize that one should not expect that a method known to be efficient when used for L^2 approximations would also be efficient for L^∞ approximations; this is one source of the difficulties that arise for this problem.

We divide the interval $(0, a)$ into N subintervals of equal length h , so that $Nh = a$. We use the centred difference method to replace the differential operator A by the $(N-1) \times (N-1)$ matrix

$$M_{r,s} = \begin{cases} -2/bh^2, & \text{if } r = s; \\ 1/bh^2 + 1/2h, & \text{if } r = s + 1; \\ 1/bh^2 - 1/2h, & \text{if } r = s - 1; \\ -0, & \text{otherwise.} \end{cases} \quad (6)$$

If $h \leq 2/b$, the off-diagonal entries are all non-negative. In addition, the row sums all vanish, except for those for $r = 1, N - 1$, which are negative. It follows by standard arguments (see [2, Chapter 7] and [8]) that e^{Mt} are positivity-preserving contractions on $l^\infty\{1, (N - 1)\}$ for all $t \geq 0$, provided that $h \leq 2/b$. We define

$$\tilde{n}_N(t) = \max\{(e^{Mt}1)_r : 1 \leq r \leq N - 1\},$$

and we compute the exponential using Matlab's `expm` algorithm; similar results would no doubt be obtained with other standard exponential algorithms. We found no evidence of the failure of `expm`: we are primarily interested in the best choice of the basis of trial functions, not in the linear algebra performed after making this choice. However, we point out that if the test function space has very high dimension (say 10^5 or more), the computation of the matrix e^{At} is not a routine matter, because it is full.

Table 4 gives the values of $\tilde{n}_N(t)$ for various t and N . One sees that the values obtained are accurate for large enough N . If one repeats the computation for $N < 100$, one observes a dramatic loss in accuracy: one obtains $\tilde{n}_{99}(t) = 1.0050$ for $t = 1, 2, 4$ and $\tilde{n}_{90}(t) = 1.0526$ for $t = 1, 2, 4, 6$. The reason is that for such N , all of the entries $M_{r,r+1}$ are negative. There is therefore no built-in reason why $\|e^{Mt}\|_\infty$ should be less than or equal to 1, and there is a consequent lack of stability in the computations.

This well-known phenomenon confirms the need to be very cautious when discretizing an evolution equation that contains two terms operating on very different time-scales: in this case, a large hyperbolic term and a much smaller parabolic term. One can avoid the problem in this case by using a backward difference approximation for the first-order term instead of a centered difference approximation. However, we will see in the next section that *every* discretization runs up against a fundamental difficulty for short times.

The case $N = 100$ is borderline: the matrix M is then lower triangular and bidiagonal. The values of $\|e^{Mt}\|_\infty$ computed using `expm` coincide with the exact values of the same expression to nine decimal places for $1 \leq t \leq 10$. This does not, of course, mean that they coincide with the norms of e^{At} , which act in an infinite-dimensional space. We put $N = 110$ instead of $N = 100$ in Table 4 because this provides considerably better accuracy with very little extra cost.

If one increases the value of N , the results converge slowly to the true value of $n(t)$, at the cost of increasing the run-time of the computation; the memory requirements also grow rapidly, because e^{At} is a full matrix. We obtained $\tilde{n}_{500}(10) = 0.39565$ and $\tilde{n}_{1000}(10) = 0.39476$.

Table 4: Values of $\tilde{n}_N(t)$ computed using the discretization method; see Appendix A.4.

t	$\tilde{n}_{110}(t)$	$\tilde{n}_{200}(t)$	$\tilde{n}_{300}(t)$	$n(t)$
1	1.00000	1.00000	1.00000	1.00000
2	1.00000	1.00000	1.00000	1.00000
4	1.00000	1.00000	1.00000	1.00000
6	0.99999	0.99999	0.99999	0.99999
8	0.97139	0.96684	0.96632	0.96593
10	0.42446	0.40210	0.39763	0.39454

Table 5 compares the true values of $m(t)$ with the values of $\tilde{m}_N(t)$ obtained using Matlab's `expm` algorithm with the same discretization of the interval $(0, a)$, as before. We see that the accuracy of the results is better than it was for $n(t)$, but not very much so.

Table 5: Values of $\tilde{m}_N(t)$ computed using the discretization method; see Appendix A.5.

t	$\tilde{m}_{110}(t)$	$\tilde{m}_{200}(t)$	$\tilde{m}_{300}(t)$	$m(t)$
1	0.99508	0.99508	0.99508	0.99508
2	0.99018	0.99018	0.99018	0.99018
4	0.98046	0.98045	0.98045	0.98045
6	0.90264	0.89307	0.89151	0.89058
8	0.52120	0.50999	0.50785	0.50619
10	0.10159	0.09448	0.09319	0.09263

4. Short-time asymptotics

In this section and the next, we show that the above approximation procedure cannot be as accurate as Table 4 suggests. We start by pointing out that there is an unavoidable problem when seeking approximations for short enough times. These arise from the following argument.

LEMMA 1. *The semigroup $T_t = e^{At}$ on $L^\infty(0, a)$ satisfies*

$$\|T_t 1 - 1\|_\infty = 1$$

for all $t > 0$. If one regards T_t as acting in $C_0(0, a)$ provided with the uniform norm, then one has

$$\|T_t - I\| \geq 1$$

for all $t > 0$.

Proof. Since each T_t is a positivity-preserving contraction on L^∞ , we have $0 \leq T_t 1 \leq 1$ for all $t > 0$. The boundary conditions on A imply that $T_t 1 \in C_0(0, a)$ for all $t > 0$, and it follows that

$$\|T_t 1 - 1\|_\infty = 1.$$

If $\|T_t - I\| < 1$ as an operator on $C_0(0, a)$, the same holds for $(T_t - I)^*$ acting on $L^1(0, a)$, regarded as a Banach subspace of $C_0(0, a)^*$. By taking duals again, we deduce that $\|T_t - I\| < 1$ when considered as an operator acting on $L^\infty(0, a)$. This contradicts $\|T_t 1 - 1\|_\infty = 1$. \square

Now suppose that some method of approximating T_t by a semigroup e^{Mt} acting in C^N has been devised. Suppose also that when estimating all vectors and operators with respect to the L^∞ norm in C^N , we have $\|e^{Mt}\| \leq 2$ for $0 \leq t \leq 1$, and that the function $1 \in L^\infty(0, a)$ is represented by the sequence $1 \in \mathbb{C}^N$.

LEMMA 2. We have

$$\|e^{Mt}1 - 1\|_\infty < 1/2, \quad (7)$$

provided that

$$0 < t \leq \frac{1}{4\|M\|}.$$

In the particular case of the discrete approximation of Section 3 the bound (7) holds, provided that

$$0 < t \leq \frac{bh^2}{16}.$$

Proof. The first statement follows from

$$\|e^{Mt}1 - 1\|_\infty = \left\| \int_0^t e^{Ms} M1 \, ds \right\|_\infty \leq 2t\|M\|_\infty.$$

The second uses the extra fact that

$$\|M\|_\infty = \frac{4}{bh^2}. \quad \square$$

We can compare the above estimates with the results obtained by computing

$$\rho(t) = \|e^{Mt}1 - 1\|_\infty \quad (8)$$

for small $t > 0$, using the method of the last section. When computing Table 6, we put $a = 10$, $b = 20$ and $h = 1/20$, so that $N = 200$. According to Lemma 2, $\rho(t) < 1/2$ for $0 \leq t \leq 1/320 \sim 0.003$, but Table 6 shows that this also holds for $0 \leq t \leq 0.02$. Neither result is compatible with the behaviour of T_t as proved in Lemma 1.

The results demonstrate that the discretization does not have acceptable behaviour with respect to the L^∞ norm for small $t > 0$. This being the case, one has to be very cautious about accepting results obtained using it for larger t .

Table 6: Values of $\rho(t)$ as defined by (8); see Appendix A.6.

t	$\rho(t)$
0	0
0.01	0.248
0.02	0.420
0.03	0.541
0.04	0.630
0.05	0.697
0.1	0.866
0.2	0.960
0.5	0.997
1.0	0.99992

The above difficulty cannot be avoided by using a different type of discretization. The two lemmas above show that the same problem *must* arise for small $t > 0$. If any numerical procedure generates a substantial error over a very short time, there is no obvious basis for believing it for longer times, unless one can find a reason why the error should die out over longer times. In the next section, we show that this problem is not purely theoretical.

5. *Bounds on the actual errors*

Table 4 gives the impression that for $a = 10$, $b = 20$ and $h = 1/20$, the discretization method gives fairly accurate results for $1 \leq t \leq 10$. On the other hand, the results of Section 4 show that it is not reliable for $t < 1$. In this section we show that the discretization method is not as accurate as Table 4 suggests for *any* $t > 0$. We show this in Table 7, which presents the values of

$$\sigma_N(t) = \|e^{M_N t} \mathbf{1} - P_N e^{M_{2N} t} \mathbf{1}\|_\infty \tag{9}$$

for $N = 200$, $N = 300$ and various t , where M_N is defined by (6).

The projection P_N is needed in the above equation because the two matrices act in spaces of different dimensions. In order to make the comparison, one should regard C^{2N-1} as obtained by sampling a function in $C_0(0, a)$ at the points $x = ra/2N$ for $1 \leq r \leq 2N-1$; the projection P_N from C^{2N-1} onto C^{N-1} is then obtained by discarding the odd sampling points.

Table 7: Values of $\sigma_N(t)$ as defined by (9); see [Appendix A.7](#).

t	$\sigma_{200}(t)$	$\sigma_{300}(t)$
0.1	0.0251	0.0100
0.5	0.0265	0.0102
1	0.0267	0.0103
2	0.0267	0.0103
4	0.0267	0.0103
6	0.0267	0.0103
8	0.0259	0.0100
10	0.0142	0.0056

Table 7 shows that the error is present even for small t , and then persists as t increases. By plotting the differences $e^{M_N t} \mathbf{1} - P_N e^{M_{2N} t} \mathbf{1}$, one finds that for every t in the range, the error is concentrated near the ends of the interval $(0, a)$. This is as expected, since the difficulty is associated with the fact that the initial function $\mathbf{1}$ does not lie in the domain of the operator A : it does not satisfy the boundary conditions at $0, a$.

6. *A trigonometric approximation method*

One can apply `expm` to a matrix obtained from a trigonometric basis. The fact that a substantial proportion of the off-diagonal entries of the matrix M defined in (10) are

negative suggests that the use of M will not give good results unless the dimension of the test function space is very large. There is also a second reason to doubt the likely accuracy of this method.

Let

$$e_n(x) = \sqrt{2/a} \sin(\pi n x/a),$$

and put

$$M(m, n) = \langle A e_n, e_m \rangle \tag{10}$$

for $1 \leq m, n \leq N$. Then one has to compute

$$n_N^\dagger(t) := \max \left\{ \sum_{m=1}^N \beta_m e_n(x) : 0 \leq x \leq a \right\}$$

where $\beta = \expm(t * M) * \alpha \in \mathbb{C}^N$ and $\alpha \in \mathbb{C}^N$ is given by

$$\alpha_n = \langle 1, e_n \rangle = \frac{\sqrt{2a}}{\pi n} (1 - (-1)^n).$$

These coefficients decrease slowly, and so one needs a large value of N to get a good initial approximation.

Table 8 shows the results of using this method to compute the norm $n_N^\dagger(t)$ for various N and t . The results confirm the above expectations.

Table 8: Values of $n_N^\dagger(t)$ computed using the trigonometric method; see [Appendix A.8](#).

t	$n_{100}^\dagger(t)$	$n_{200}^\dagger(t)$	$n_{300}^\dagger(t)$	$n(t)$
1	1.02322	1.00254	1.00067	1.00000
2	1.02320	1.00253	1.00067	1.00000
4	1.02320	1.00253	1.00067	1.00000
6	1.02320	1.00253	1.00066	0.99999
8	0.99069	0.96959	0.96669	0.96593
10	0.39513	0.39680	0.39364	0.39454

7. The pseudospectral method

In this section we consider the same problem, using the recently introduced pseudospectral expansion of [4]. The advantage of this method is that it uses a test function space adapted to the operator in question, leading to the hope that it may be of much smaller dimension than the use of finite element methods would need for comparable accuracy. For background material on pseudospectra, we refer to [3, 6, 7, 12, 13, 14]. The basic idea is to approximate the initial function by a linear combination of approximate (that is, pseudo-) eigenfunctions e_σ . These are required to satisfy $\|e_\sigma\| = 1$ and

$$\inf\{\|e_\sigma - w\| + \|Aw - z_\sigma w\| : w \in \text{Dom}(A)\} < \varepsilon$$

for some pre-assigned ‘acceptable’ error $\varepsilon > 0$.

As ε decreases, the e_σ might become more singular, with the result that using true eigenfunctions (that is, putting $\varepsilon = 0$) is numerically infeasible. The computation of suitable pseudo-eigenfunctions is a part of the theory of pseudospectra; given $\varepsilon > 0$, they are typically determined by one or more continuous parameters, and one includes a finite number of true eigenfunctions in the family if the norms of the corresponding spectral projections are of order 1.

In the present example, suitable unnormalised pseudo-eigenfunctions were obtained in [4], and are of the form

$$e_\sigma(x) = e^{(-b/2+b\delta+i\sigma)(a-x)} - e^{(-b/2-b\delta-i\sigma)(a-x)},$$

where $N = 2M + 1$, $\delta = 1/2 - c/(ab)$ and $\sigma \in \mathbb{R}$. Note that the e_σ satisfy the boundary conditions exactly at $x = a$, but only approximately at $x = 0$ (we assume a choice such as $c = 20$). This is an overcomplete family, and one has to choose some discrete subfamily before starting computations. An analogy with the theory of Fourier series suggests taking $\sigma = 2\pi m/a$, where $m \in \mathbb{Z}$. We will show later that the use of overcomplete families of e_σ yields approximations of comparable accuracy, but in this section we make the above choice.

As in [4], we approximate $f_t = T_t f$ for $f = 1$ by

$$g_t = E(\phi_t) \in L^2(0, a). \quad (11)$$

In this equation, $E : \mathbb{C}^N \rightarrow L^2(0, a)$ is obtained by concatenating the set of pseudo-eigenfunctions

$$e_m(x) = e^{(-b/2+b\delta+2\pi im/a)(a-x)} - e^{(-b/2-b\delta-2\pi im/a)(a-x)},$$

where $N = 2M + 1$, $\delta = 1/2 - c/(ab)$ and $-M \leq m \leq M$. Specifically,

$$(E\psi)(x) = \sum_{m=-M}^M \psi_m e_m(x),$$

for all $\psi \in \mathbb{C}^N$. Finally,

$$(\phi_t)_m = e^{\mu_m t} \phi_m \in \mathbb{C}^N,$$

where

$$\phi = E \setminus f. \quad (12)$$

We have changed the differential operator A slightly from what is considered in [4], and so the e_m are concentrated around $x = a$ rather than around $x = 0$. The pseudo-eigenvalues μ_n are given by

$$\mu_m = b^{-1} (-b/2 + b\delta + 2\pi im/a)^2 + (-b/2 + b\delta + 2\pi im/a).$$

The test functions depend upon the parameter c , which must satisfy $0 < c < ab/2$. They are of order 1 near $x = a$, and of order e^{-c} near $x = 0$. The computation becomes more accurate as c increases, at the cost of being more ill-conditioned. For $c = 35$ it breaks down, because each of the test functions used is then of order 10^{-16} near $x = 0$, and hence appears to be equal to 0 when using double precision arithmetic; it is therefore impossible to obtain a good initial approximation to $f = 1$ at $t = 0$.

We approximated $n(t)$ for $a = 10$ and $b = 20$ by putting $f_0 = 1$ and evaluating

$$\bar{n}_N(t) := \|g_t\|_\infty.$$

We put $c = 20$, although other similar values lead to similar results. We replaced $(0, a)$ by 10^4 equally spaced points, and used the trapezium rule for evaluating integrals when computing E ; this was a significant source of error, and no doubt a more sophisticated procedure could be devised. We computed the values of $\bar{n}_N(t)$ for $N = 101$ and t from 1 to 10. Table 9 presents the computed values.

The same values were obtained with considerably smaller N for $0 \leq t \leq 6$. For $t \geq 8$, the results were not stable to the accuracy shown, even when N was increased to 400. Note that there is no structural reason why this numerical method should yield values in the range $[0, 1]$, so the fact that this is the case itself provides evidence that the approximation method is working; for the discretization method, this was a structural feature of the approximation, and therefore proves nothing about its accuracy. Almost all of the run-time is spent computing ϕ using (12). On comparing Tables 4 and 9, we conclude that for the initial function $f = 1$ the discretization method is to be preferred, although neither is very accurate.

Table 9: Values of $\bar{n}_N(t)$ computed using the pseudospectral method; see Appendix A.9.

t	$\bar{n}_{101}(t)$	$n(t)$
1	1.00000	1.00000
2	1.00000	1.00000
4	1.00000	1.00000
6	0.99999	0.99999
8	0.970	0.96593
10	0.422	0.39454

Table 10: Values of $\bar{m}_N(t)$ computed using the pseudospectral method; see Appendix A.10.

t	$\bar{m}_{81}(t)$	$m(t)$
1	0.99508	0.99508
2	0.99018	0.99018
4	0.98045	0.98045
6	0.89058	0.89058
8	0.50629	0.50619
10	0.09346	0.09263

The situation is very different for the initial function $f(x) = \sin(\pi x/a)$. Table 10 presents the values of $\bar{m}_N(t)$ computed using the pseudospectral method for $a = 10$, $b = 20$, $c = 20$ and $N = 81$. For $t \leq 6$, similar accuracy could be obtained for much smaller values of N . It is clear that the pseudospectral method is much more accurate in this case than it was when computing $n(t)$.

8. Overcomplete families

If $\{e_m\}_{m=1}^\infty$ is a sequence of vectors in an abstract Hilbert space \mathcal{H} , one can try to expand a general vector $f \in \mathcal{H}$ as a norm-convergent series $f = \sum_{m=1}^\infty \phi_m e_m$. This cannot succeed for all f unless the sequence is complete – that is, unless its linear span is dense in \mathcal{H} . If the sequence is overcomplete (that is, it remains complete after the removal of some member), then the expansion is not unique, so there cannot be a canonical formula for determining the coefficients ϕ_m . Historically, most effort has therefore been devoted to minimal complete sequences of vectors, and in particular to Riesz bases.

In this section we show that expansions in terms of overcomplete sequences are feasible numerically, and that they are sometimes more accurate than traditional expansions. We explain this first by reference to ordinary Fourier series, and apply it to our differential operator only in the next section.

Given $u > 0$ and $m \in \mathbb{Z}$, let $e_{m,u} \in L^2(0, 1)$ be the function

$$e_{m,u}(x) = e^{2\pi i m x / u}. \quad (13)$$

If $u = 1$, this is the familiar complete orthonormal Fourier basis. If $0 < u < 1$, it is not complete, and if $u > 1$, it is overcomplete. Given a smooth function on $[0, 1]$, the standard Fourier expansion converges slowly unless the function is periodic. However, if $u > 1$, such a function can be extended to a smooth periodic function on $[0, u]$, and the standard Fourier expansion of this extended function in $L^2(0, u)$ converges rapidly. Restricting back to $L^2(0, 1)$, we see that there exists a rapidly converging expansion of f in terms of $e_{n,u}$ for any $u > 1$. The nonuniqueness of the extension of f from $[0, 1]$ to $[0, u]$ might seem to be a barrier to the application of this method, but it can be resolved by the following regularized least-squares approximation method, with parameter δ ; see [10].

Given any finite sequence of vectors $\{e_m\}_{m=1}^M$ with $\|e_m\| \leq 1$ in a Hilbert space \mathcal{H} , $f \in \mathcal{H}$, a positive weight w on $\{1, \dots, M\}$ and $\delta > 0$, we define

$$\mathcal{E}_M(f) = \min \left\{ \left\| f - \sum_{m=1}^M \phi_m e_m \right\|^2 + \delta \langle w\phi, \phi \rangle : \phi \in \mathbb{C}^M \right\}, \quad (14)$$

where we henceforth also use w to denote the diagonal matrix whose diagonal entries are w_1, \dots, w_M .

LEMMA 3. $\mathcal{E}_M(f)$ is minimized by a unique $\phi \in \mathbb{C}^M$ which depends linearly on f . It may be written in the (Matlab) form $\phi = \tilde{E} \setminus (f \oplus 0)$, where the linear operator \tilde{E} is defined by (16). The condition number of \tilde{E} (the ratio of its largest and smallest singular values) satisfies

$$\text{cond}(\tilde{E}) \leq \left(\frac{\delta w_2 + M}{\delta w_1} \right)^{1/2}, \quad \text{where } w_1 = \min\{w\} \text{ and } w_2 = \max\{w\}. \quad (15)$$

Proof. If $E : \mathbb{C}^M \rightarrow \mathcal{H}$ is defined by

$$E\psi = \sum_{m=1}^M \psi_m e_m,$$

then the variation of parameters method shows that the minimum on the right-hand side of (14) is achieved for

$$\phi = (E^* E + \delta w)^{-1} E^* f,$$

which depends linearly upon f .

One can alternatively characterize ϕ as the vector in \mathbb{C}^M that minimizes $\|\tilde{E}\phi - (f \oplus 0)\|$, where $\tilde{E} : \mathbb{C}^M \rightarrow \mathcal{H} \oplus \mathbb{C}^M$ is defined by

$$\tilde{E}\phi = (E\phi) \oplus (\delta^{1/2}w^{1/2}\phi). \quad (16)$$

Every entry of the matrix E^*E has magnitude at most 1, so $\|E^*E\| \leq M$. It then follows from (16) that

$$\delta w_1 \|\phi\|^2 \leq \|\tilde{E}\phi\|^2 \leq (\delta w_2 + M)\|\phi\|^2$$

for all ϕ , so the condition number of \tilde{E} satisfies (15). \square

The advantage of using \tilde{E} instead of E is that for an overcomplete family, E is very badly conditioned. For small positive $\delta > 0$, one obtains almost as good a result using \tilde{E} , which is a better conditioned matrix. In examples it appears that $\delta = 10^{-10}$ often yields an acceptable result.

There are two advantages of choosing a weight w other than $w = 1$. It allows one to discriminate progressively against undesirable choices of e_m : for example, highly oscillatory test functions. Secondly, in our applications we are interested in $\|\phi\|_1$, not in the quantity $\|w^{1/2}\phi\|_2$, which arises in (16). Using the obvious inequality

$$\|\phi\|_1 \leq \|w^{-1/2}\|_2 \|w^{1/2}\phi\|_2,$$

we see that the two norms may be of the same order of magnitude if

$$\sum_{m=1}^{\infty} w_m^{-1} < \infty.$$

We illustrate the power of these ideas by reference to the Fourier-type example at the start of this section. We chose $f \in L^2(0, 1)$ to be $f(x) = \sqrt{3}x$, so that $\|f\|_2 = 1$. We put $\delta = 10^{-10}$ and $w = 1$, and we let m range from -10 to 10 . We replaced $(0, 1)$ by 10^4 equally spaced points, and we used the trapezium rule for evaluating integrals as before. Table 11 shows how the error $\|f - E\phi\|_2$ and $\|\phi\|_2$ depend upon the choice of u . We see that for $0 < u < 1$ the error is large, corresponding to the fact that the sequence (13) is then

Table 11: Approximation errors for the overcomplete Fourier expansion method; see Appendix A.11.

u	$\ f - E\phi\ _2$	$\ \phi\ _2$
0.8	0.4443	0.9019
0.9	0.3574	0.9407
1.0	0.1201	0.9928
1.1	8.196×10^{-3}	1.0014
1.2	7.291×10^{-4}	1.0150
1.5	2.473×10^{-6}	1.0688
2.0	3.225×10^{-9}	1.1814
2.5	2.591×10^{-11}	1.3099

incomplete. For $u > 1$, the error is much smaller than it is for the complete orthonormal set corresponding to $u = 1$.

If $f \in C^\infty[0, 1]$ and $u > 1$, then f may be extended (in several ways) to a smooth function \tilde{f} on \mathbb{R} that is periodic with period u . Neglecting a scale factor, the sequence $\{e_m\}_{m=-\infty}^\infty$ may be extended to the standard Fourier sequence $\{\tilde{e}_m\}_{m=-\infty}^\infty$ for $L^2(0, u)$. The Fourier coefficients of \tilde{f} in $L^2(0, u)$ with respect to the standard Fourier basis decrease at a super-polynomial rate, and one may then apply the following lemma.

LEMMA 4. Let \mathcal{H} be contained in $\tilde{\mathcal{H}}$, and let $\{e_m\}_{m=1}^\infty$ be the orthogonal projections onto \mathcal{H} of a complete orthonormal sequence $\{\tilde{e}_m\}_{m=1}^\infty$ in $\tilde{\mathcal{H}}$. Let f be the orthogonal projection of some $\tilde{f} \in \tilde{\mathcal{H}}$ onto \mathcal{H} , and let $\phi \in \mathbb{C}^M$ be the sequence that minimizes $\mathcal{E}_M(f)$. Then

$$\mathcal{E}_M(f) \leq \sum_{m=M+1}^\infty |\langle \tilde{f}, \tilde{e}_m \rangle|^2 + \delta \sum_{m=1}^M w_m |\langle \tilde{f}, \tilde{e}_m \rangle|^2.$$

Proof. This is a routine computation, once one replaces the minimizing ϕ in $\mathcal{E}_M(f)$ by the sequence $\phi_m = \langle \tilde{f}, \tilde{e}_m \rangle$, $1 \leq m \leq M$. \square

9. More on semigroup approximation

In the Fourier example just discussed, the value $u = 1$ at which the sequence passed from being incomplete to being overcomplete was known a priori. When applying the pseudospectral approximation method to a general one-parameter semigroup, the sequence of vectors e_n is generated numerically, and it is not possible to ensure that they form a Riesz basis. One can try to take enough vectors to ensure overcompleteness, and then confirm success a posteriori by observing the small size of $\|E\phi - f\|$.

The following modification of [4, Theorem 5] shows how to approximate a semigroup using an overcomplete sequence of pseudoeigenvectors. In [4], we assumed that $\phi = E \setminus f$, but this fact was not used. For the right-hand side of (18) to be small, we need that $\|f - E\phi\|$ and $\varepsilon \|\phi\|_1$ should both be small. If the vectors e_1, \dots, e_N are part of an infinite overcomplete sequence, this may be achieved by putting $\phi = \tilde{E} \setminus (f \oplus 0)$, as described in the last section.

THEOREM 5. Let T_t be a one-parameter semigroup on \mathcal{B} satisfying $\|T_t\| \leq M e^{\gamma t}$ for all $t \geq 0$. Let $\varepsilon > 0$, and let $\{e_n\}_{n=1}^N$ be a finite sequence of normalized approximate eigenvectors of A in \mathcal{B} , in the sense that $\|e_n\| = 1$ and

$$\inf\{\|e_n - w\| + \|Aw - \lambda_n w\| : w \in \text{Dom}(A)\} < \varepsilon \quad (17)$$

for all n , where $\text{Re}(\lambda_n) \leq \gamma$ for all n . Let $E : \mathbb{C}^N \rightarrow \mathcal{B}$ be defined by $E\psi = \sum_{n=1}^N \psi_n e_n$. Given $\phi \in \mathbb{C}^N$, let $(\phi_t)_n = e^{\lambda_n t} \phi_n$. Then

$$\|T_t f - E\phi_t\| < \|f - E\phi\| M e^{\gamma t} + \varepsilon \|\phi\|_1 (1 + M + Mt) e^{\gamma t}. \quad (18)$$

Proof. Using (17), let $w_n \in \text{Dom}(A)$ satisfy

$$\|e_n - w_n\| + \|Aw_n - \lambda_n w_n\| < \varepsilon$$

for each n .

Then for any $t > 0$, we have

$$\begin{aligned}
 \|T_t e_n - e^{\lambda_n t} e_n\| &\leq \|e^{\lambda_n t} (e_n - w_n)\| + \|T_t (e_n - w_n)\| + \|T_t w_n - e^{\lambda_n t} w_n\| \\
 &< \varepsilon(1 + M)e^{\gamma t} + \left\| \int_0^t T_{t-s} e^{\lambda_n s} (A w_n - \lambda_n w_n) ds \right\| \\
 &\leq \varepsilon(1 + M)e^{\gamma t} + \int_0^t M e^{\gamma(t-s)} e^{\gamma s} \|A w_n - \lambda_n w_n\| ds \\
 &\leq \varepsilon(1 + M + Mt)e^{\gamma t}.
 \end{aligned}$$

Hence

$$\begin{aligned}
 \|T_t f - E\phi_t\| &\leq \|T_t f - T_t E\phi\| + \|T_t E\phi - E\phi_t\| \\
 &\leq \|f - E\phi\| M e^{\gamma t} + \sum_{n=1}^N |\phi_n| \|T_t e_n - e^{\lambda_n t} e_n\| \\
 &\leq \|f - E\phi\| M e^{\gamma t} + \varepsilon \|\phi\|_1 (1 + M + Mt) e^{\gamma t}. \quad \square
 \end{aligned}$$

We repeated the calculations of Section 7, but with the new definition $\phi = \tilde{E} \setminus (f \oplus 0)$ instead of (12). Referring back to (16), we put $w(s) = 1 + s^2$ and chose the constant δ to be 10^{-10} . Putting $N = 81$, we denoted the result by $\check{m}_u(t)$, where u denotes the value of an overcompleteness parameter u . We computed $\check{m}_u(t)$ for $t = 1, \dots, 10$, and for various values of u . The (unnormalized) pseudoeigenfunctions are now given by

$$e_m(x) = e^{(-b/2 + b\delta + 2\pi im/ua)(a-x)} - e^{(-b/2 - b\delta - 2\pi im/ua)(a-x)},$$

so the case $u = 1$ can be compared directly with the results of Table 8, which uses E instead of \tilde{E} .

We see from Table 12 that the results are not useful for $u = 0.5$. If $u \geq 1$, they are not heavily dependent upon the value of u . If one puts $\delta = 10^{-15}$, Matlab considers the matrix \tilde{E} to be ill-conditioned, and the graph of the sequence ϕ that it obtains is very irregular.

Table 12: Values of $\check{m}_u(t)$ for various choices of the overcompleteness constant u ; see Appendix A.10.

t	$\check{m}_{0.5}(t)$	$\check{m}_1(t)$	$\check{m}_{1.5}(t)$	$\check{m}_2(t)$	$m(t)$
1	0.96672	0.99508	0.99508	0.99508	0.99508
2	0.93801	0.99018	0.99018	0.99018	0.99018
4	0.85509	0.98045	0.98045	0.98045	0.98045
6	0.07462	0.89058	0.89058	0.89058	0.89058
8	0.00028	0.50623	0.50626	0.50635	0.50619
10	0.00002	0.09319	0.09196	0.09335	0.09263

10. Fractional powers of A

In this section we show that the pseudospectral method can be used to approximate a variety of other functions $f(A)$, provided that one has a prior estimate of $\|f(A)\|$. We will investigate one particular case.

If e^{-At} is a one-parameter contraction semigroup on \mathcal{B} , then $e^{-A^\alpha t}$ is also a one-parameter contraction semigroup for all $0 < \alpha < 1$. The definition of the fractional power is a little involved, except when A has purely discrete spectrum, in which case one has the property that if $Af = \lambda f$, then $\operatorname{Re}(\lambda) \geq 0$ and $A^\alpha f = \lambda^\alpha f$. In the general case, one has the formula

$$e^{-A^\alpha t} f = \int_0^\infty \rho(s, t, \alpha) e^{-As} f \, ds,$$

where ρ is a positive density with

$$\int_0^\infty \rho(s, t, \alpha) \, ds = 1$$

for all $t > 0$ and $0 < \alpha < 1$; see [15, p. 259]. This formula implies immediately that if e^{-At} are positivity-preserving contractions on $C_0(X)$ or on $L^p(X)$ for any $1 \leq p \leq \infty$, then the same applies to $e^{-A^\alpha t}$. If also $e^{-At} 1 \in C_0(X)$ for all $t > 0$, then the same applies to $e^{-A^\alpha t}$, from which we deduce that

$$\|e^{-A^\alpha t} 1 - 1\|_\infty = 1$$

for all $t > 0$. In other words, the short-time approximation problems for $e^{-A^\alpha t}$ also occur for e^{-At} . In other respects, however, the semigroup $e^{-A^\alpha t}$ has quite different properties. Its generator is not a local operator, and the corresponding Markov process has discontinuous sample paths. The transition from approximately Markov behaviour to rapid decay is less sharp, the smaller the value of α .

One can apply the methods of Section 3 to the new semigroup simply by replacing (11) by

$$g_t = E \left(\sum_{m=-M}^M e^{-\mu_m^\alpha t} \phi_m \right)$$

where E, ϕ_m and μ_m are defined as before, except for a change of sign in μ_m . The justification for doing this depends upon the following lemma. We expect that the proof can be extended to other values of α .

LEMMA 6. *Let e^{-At} be a one-parameter contraction semigroup acting on \mathcal{B} . If $\|u\| = 1$, $\operatorname{Re}(\mu) \geq 0$ and*

$$\|u - w\| + \|Aw - \mu w\| < \varepsilon < 1/2,$$

then

$$\|e^{-A^{1/2}t} u - e^{-\mu^{1/2}t} u\| < 2\varepsilon + 2\varepsilon^{1/2}t, \tag{19}$$

for all $t > 0$.

Proof. We first observe that the left-hand side of (19) is bounded above by $2\varepsilon + \beta$, where

$$\beta := \|e^{-A^{1/2}t} w - e^{-\mu^{1/2}t} w\|.$$

The formula

$$e^{-As}w - e^{-\mu s}w = \int_0^s e^{-Au - \mu(s-u)}(\mu w - Aw) du,$$

together with the crude bound $\|w\| < 1 + \varepsilon$, implies that

$$\|e^{-As}w - e^{-\mu s}w\| < \min\{\varepsilon s, 2(1 + \varepsilon)\}$$

for all $s > 0$. It follows that

$$\begin{aligned} \beta &= \left\| \int_0^\infty \rho(s, t, 1/2) (e^{-As}w - e^{-\mu s}w) ds \right\| \\ &\leq \int_0^\infty \rho(s, t, 1/2) \|e^{-As}w - e^{-\mu s}w\| ds \\ &\leq \int_0^\infty \rho(s, t, 1/2) \min\{\varepsilon s, 2(1 + \varepsilon)\} ds. \end{aligned}$$

We estimate this with the aid of the formula

$$\rho(s, t, 1/2) = \frac{t}{2\sqrt{\pi s^3}} e^{-t^2/4s}.$$

See [2, Section 2.4] or [6, Section 3.8]. Unfortunately, no explicit formula of this type is available for other values of α . We obtain

$$\begin{aligned} \beta &\leq \int_0^\infty \frac{t}{2\sqrt{\pi s^3}} \min\{\varepsilon s, 2(1 + \varepsilon)\} ds \\ &= \int_0^\infty \frac{1}{2\sqrt{\pi u^3}} \min\{\varepsilon ut^2, 2(1 + \varepsilon)\} du \\ &= \sqrt{\frac{8(1 + \varepsilon)}{\pi}} \varepsilon^{1/2} t. \end{aligned} \quad \square$$

11. Conclusions

We showed in Section 4 that the numerical computation of $e^{At}1$ is an intrinsically hard problem for small $t > 0$, and we confirmed that the difficulty is genuine in Tables 6 and 7. The discretization method is better than the pseudospectral method for the initial function $f = 1$, but both are rather inaccurate for trial function spaces of a few hundred dimensions. The ultimate reason for these problems is the lack of continuity of the semigroup e^{At} at $t = 0$, when acting in the space $C[0, a]$.

The situation changes entirely when one applies the semigroup to a function lying in the domain of its generator. Table 13 compares the errors $\delta_N(t)$ and $\delta_{110}(t)$ in the computations of the function $m(t) = \|e^{At}f\|_\infty$, where $f(x) = \sin(\pi x/a)$, using the pseudospectral and discretization methods respectively. Once again, we recorded all the relevant data to nine decimal places. In this case, the advantages of the pseudospectral method over the discretization method are clear.

The potential advantage of the pseudospectral method is that it uses a test function space that is adapted to the operator, and which may therefore be of relatively low dimension. We have indicated briefly how the sequence of pseudo-eigenvectors might be generated in more general situations, but further investigations are needed to validate that method.

A semi-classical procedure for generating pseudo-eigenvectors is described in [5], and is currently being developed further.

Table 13: Comparison between pseudospectral and discretization methods.

t	N	$\bar{\delta}_N(t)$	$\tilde{\delta}_{110}(t)$
1	51	7×10^{-9}	3.3×10^{-7}
2	41	$< 10^{-9}$	6.6×10^{-7}
4	41	$< 10^{-9}$	1.3×10^{-6}
6	41	5.1×10^{-7}	1.2×10^{-2}
8	81	1.1×10^{-4}	1.5×10^{-2}
10	81	8.4×10^{-4}	9×10^{-3}

Acknowledgements. We would like to thank N. J. Higham and the referees for helpful comments on an early version of this paper. We also acknowledge financial support under the EPSRC grant GR/R81756/01.

Appendix A. *List of programs used*

The following programs are available to subscribers to the journal at

<http://www.lms.ac.uk/jcm/7/lms2003-028/appendix-a>.

Appendix A.1.

The program for computing $n(t)$ using the spectral method is available at

<http://www.lms.ac.uk/jcm/7/lms2003-028/appendix-a/mixing3.mws>

(see Table 1).

Appendix A.2.

The program for computing $r(t) = \log(n(t)) - \log(n(t - 1))$ is available at

<http://www.lms.ac.uk/jcm/7/lms2003-028/appendix-a/mixing4.mws>

(see Table 2).

Appendix A.3.

The program for computing $m(t)$ using the spectral method is available at

<http://www.lms.ac.uk/jcm/7/lms2003-028/appendix-a/mixing5.mws>

(see Table 3).

Appendix A.4.

The program for computing values of $\tilde{n}_N(t)$ using the discretization method is available at

<http://www.lms.ac.uk/jcm/7/lms2003-028/appendix-a/expm7.m.mws>

(see Table 4).

Appendix A.5.

The program for computing values of $\tilde{m}_N(t)$ using the discretization method is available at

<http://www.lms.ac.uk/jcm/7/lms2003-028/appendix-a/expm8.m.mws>

(see Table 5).

Appendix A.6.

The program for computing $\rho(t) = \|e^{Mt} \mathbf{1} - \mathbf{1}\|_\infty$ is available at

<http://www.lms.ac.uk/jcm/7/lms2003-028/appendix-a/expm9.m>

(see Table 6).

Appendix A.7.

The program for computing $\sigma_N(t) = \|e^{M_N t} \mathbf{1} - P_N e^{M_{2N} t} \mathbf{1}\|_\infty$ is available at

<http://www.lms.ac.uk/jcm/7/lms2003-028/appendix-a/expm10.m>

(see Table 7).

Appendix A.8.

The program for computing $n_N^\dagger(t)$ using the trigonometric method is available at

<http://www.lms.ac.uk/jcm/7/lms2003-028/appendix-a/expm4.m>

(see Table 8).

Appendix A.9.

The program for computing $\bar{n}_N(t)$ using the pseudospectral method is available at

<http://www.lms.ac.uk/jcm/7/lms2003-028/appendix-a/pseudo9.m>

(see Table 9).

Appendix A.10.

The program for computing $\bar{m}_N(t)$ using the pseudospectral method is available at

<http://www.lms.ac.uk/jcm/7/lms2003-028/appendix-a/pseudo12.m>

(see Tables 10 and 12).

Appendix A.11.

The program for computing approximation errors for the overcomplete Fourier expansion method is available at

<http://www.lms.ac.uk/jcm/7/lms2003-028/appendix-a/overcomplete1.m>

(see Table 11).

References

1. D. ALDOUS and J. FILL, ‘Reversible Markov chains and random walks on graphs’, <http://stat-www.berkeley.edu/users/aldous/RWG/book.html>; monograph in preparation. 3
2. E. B. DAVIES, *One-parameter semigroups* (Academic Press, London, 1980). 5, 17
3. E. B. DAVIES, ‘Pseudospectra of differential operators’, *J. Operator Theory* 43 (2000) 243–262. 9
4. E. B. DAVIES, ‘Approximating semigroups by using pseudospectra’, Preprint, 2003. 2, 9, 10, 14
5. E. B. DAVIES, ‘Semi-classical states for non-self-adjoint Schrödinger operators’, *Comm. Math. Phys.* 200 (1999) 35–41. 18
6. E. B. DAVIES, ‘Spectral theory’, 2003, Chapters 1-3; <http://www.mth.kcl.ac.uk/MAO/Semigroups/DaviesEBrianSemigroupsBook.html>. 9, 17
7. M. EMBREE and L. N. TREFETHEN, ‘Pseudospectra gateway’; <http://www.comlab.ox.ac.uk/pseudospectra>. 9
8. K. J. ENGEL and R. NAGEL, *One-parameter semigroups for linear evolution equations*, Grad. Texts in Math. 194 (Springer, 1999). 5
9. G. H. GOLUB and C. F. VAN LOAN, *Matrix computations* (Johns Hopkins Univ. Press, 1983). 2
10. P. C. HANSEN, *Rank-deficient and discrete ill-posed problems: numerical aspects of linear inversion*, SIAM Monogr. Math. Model. Comput. 4 (SIAM, Philadelphia, PA, 1997). 12
11. C. B. MOLER and C. F. VAN LOAN, ‘Nineteen dubious ways to compute the exponential of a matrix, twenty-five years later’, *SIAM Rev.* 45 (2003) 3–49. 2
12. S. C. REDDY and L. N. TREFETHEN, ‘Pseudospectra of the convection-diffusion operator’, *SIAM J. Appl. Math.* 54 (1994) 1634–1649. 9
13. L. N. TREFETHEN, ‘Pseudospectra of linear operators’, *SIAM Review* 39 (1997) 383–406. 9
14. L. N. TREFETHEN, ‘Computation of pseudospectra’, *Acta Numerica* 8 (1999) 247–295. 9
15. K. YOSIDA, *Functional analysis* (Springer, Berlin, 1965). 16

E. B. Davies E.Brian.Davies@kcl.ac.uk
http://www.mth.kcl.ac.uk/staff/eb_davies.html

Department of Mathematics
King’s College
Strand
London
WC2R 2LS