Letters in Mathematical Physics **68:** 131–138, 2004.

Optimal Error Estimates in Operator-Norm Approximations of Semigroups

V. BENTKUS[1] and V. PAULAUSKAS[2]
[1]*Department of Mathematics, Vilnius Pedagogical University, Studentu 39, LT2004, Vilnius, Lithuania. e-mail: bentkus@ktl.mii.lt*
[2]*Department of Mathematics and Informatics, Vilnius University, Naugarduko 24, Vilnius, 03225, Lithuania and Institute of Mathematics and Informatics.*
e-mail: vygantas.paulauskas@maf.vu.lt

(Received: 30 December 2003; revised version: 8 April 2004)

Abstract. We demonstrate that an idea related to the Central Limit Theorem and approximations by accompanying laws in probability theory is useful to get optimal convergence rates in some approximation formulas for operators. As examples we provide a bound for Euler approximations of bounded holomorphic semigroups; a bound for error in approximation of a power of operators by accompanying exponents, which is a useful tool in analysis of the Trotter–Kato formula, and can be considered as an extended version of Chernoff's '$\sqrt{n}$ lemma'.

Mathematics Subject Classifications (2000). 47D03, 41A80.

Key words. Chernoff theory, Euler approximations, Trotter–Kato formula, holomorphic semigroups, Ritt condition.

1. Introduction and Results

In a recent paper [15] results and methods of probability theory were applied to obtain bounds for errors for approximations of some semigroups of operators. At the same time in [1] a new approach was introduced for analysis of errors in the Central Limit Theorem and in approximations by accompanying laws. An idea of this approach is to use 'multiplicative' representations of differences. In this Letter we show that such representations are useful tools in problems related to operator convergence, and they lead to optimal bounds for errors. Our aim in this Letter is not to provide most general results, but rather on examples to explain the method, which we believe will be useful in operator theory.

Let us recall the well-known Chernoff '$\sqrt{n}$-lemma'. Let A be a linear contraction of a Banach space X. Then $\mathrm{e}^{t(A-I)}$, $t \geqslant 0$, is a contraction semigroup, and

$$\|A^n x - \mathrm{e}^{n(A-I)} x\| \leqslant n^{1/2} \|Ax - Ix\|, \quad \text{for all } x \in X,$$

([8], Lemma 2). In the next theorem we provide an operator norm bound for error in approximation of A^n by the accompanying exponent $\mathrm{e}^{n(A-I)}$.

Write $B=I-A$, for an operator A is a Banach space X, and consider the condition

$$(n+1)\|tB(I-tB)^n\| \leqslant K, \quad \text{for all } 0 \leqslant t \leqslant 1, \;\; n=0,1,2,\dots \tag{1.1}$$

with some constant K independent of t and n.

THEOREM 1.1. *Assume that A is a contraction of a Banach space X and satisfies condition* (1.1). *Then we have*

$$\Delta_{n,s} \overset{\text{def}}{=} \|(I-sB/n)^n - \mathrm{e}^{-sB}\| \leqslant 4K^2/n, \tag{1.2}$$

for all $0 \leqslant s \leqslant n$ and $n=1,2,\dots$. In particular, with $s=n$, we have

$$\Delta_n \overset{\text{def}}{=} \|A^n - \mathrm{e}^{n(A-I)}\| \leqslant 4K^2/n, \quad \textit{for all } n=1,2,\dots. \tag{1.3}$$

We provide proofs in Section 2.

To see that (1.3) is optimal, consider the function $f(x)=|x^n-\exp(n(x-1))|$ of real variable $0 \leqslant x \leqslant 1$. It is not difficult to show that $\max_{0 \leqslant x \leqslant 1} f(x) > c/n$, with some positive constant $c>0$.

The constant 4 in Theorem 1.1 can be easily made smaller.

In the case $t=1$ condition (1.1) specifies to

$$(n+1)\|A^n - A^{n+1}\| \leqslant K. \tag{1.4}$$

For an operator A consider the condition

$$\|(A-\lambda I)^{-1}\| \leqslant c|\lambda-1|^{-1}, \quad \text{for } |\lambda|>1, \;\; \lambda \in \mathbb{C}, \tag{1.5}$$

where c is a constant. Condition (1.5) can be traced back to [17]. It is known (see [13]), that Ritt's condition is equivalent to (1.4) combined with the power boundedness condition $\sup_n \|A^n\| < \infty$.

We note that the conditions of Theorem 1.1 are equivalent to the Ritt condition (recall, that in Theorem 1.1 we assume that A is contraction). Since contractions satisfy the power boundedness condition $\sup_n \|A^n\| \leqslant 1$, it suffices to check (1.1) $\Longleftrightarrow$ (1.5). To prove (1.1) $\Longrightarrow$ (1.5) we use (1.1) $\Longrightarrow$ (1.4) and note that (1.4) $\Longrightarrow$ (1.5) by the Nagy and Zemánek result. To prove (1.5) $\Longrightarrow$ (1.1) we note that all operators $I-tB$, $0 \leqslant t \leqslant 1$, satisfy (1.5) with the same constant, if A satisfies (1.5). Therefore Nagy and Zemánek's result applied to $I-tB$ yields (1.1).

Theorem 1.1 extends to bounded operators. Write

$$K_0 = \sup_{s \geqslant 0} \|\mathrm{e}^{s(A-I)}\|.$$

A sufficient condition for $K_0 < \infty$ is the power boundedness of A (hence, the Ritt condition as well). This is clear by an expansion of e^{sA} into the Taylor series.

THEOREM 1.2. *Assume that an operator A satisfies* (1.1) *and that $K_0<\infty$. Then the bounds of Theorem* 1.1 *still hold provided that we replace the constant K^2 by K_0K^2.*

It is clear that Ritt's condition implies the conditions of Theorem 1.2. Whether the conditions of Theorem 1.2 imply Ritt's condition, is an open question.

We conjecture that one of the inequalities

$$\sup_n n\Delta_n<\infty, \quad \sup_n \sup_{0\leqslant s\leqslant n} n\Delta_{n,s}<\infty$$

is equivalent to the Ritt condition.

For quasi-sectorial contractions with semi-angle $\alpha\in[0,\pi/2)$ in Hilbert spaces, Cachia and Zagrebnov ([4]) established a bound $\Delta_n=\mathrm{O}(n^{-1/3})$. Naturally, quasi-sectorial contractions satisfy the Ritt condition. By an application of probabilistic type arguments, the bound was improved to $\Delta_n=\mathrm{O}(n^{-1/2}\ln^{1/2}n)$ in [15]. One can refine the probabilistic arguments and prove $\Delta_n=\mathrm{O}(n^{-1/2})$, see [2]. We do not know how to refine the probabilistic arguments further in order to improve the bound. The new approach in this paper is based on multiplicative representations of differences introduced in [1], and consists of:

(1) a choice of a curve, say $\gamma(\tau)$, $0\leqslant\tau\leqslant 1$, connecting two close objects, say a and b, such that $\gamma(0)=a$ and $\gamma(1)=b$. In the proof of (1.3) of Th. 1.1 we take $a=A^n$ and $b=\mathrm{e}^{n(I-A)}$ and $\gamma(\tau)=(I-\tau B)^n\mathrm{e}^{-n(1-\tau)B}$, where $B=I-A$;
(2) an application of the mean value theorem (or Newton–Leibnitz formula) along the curve, that is,

$$b-a=\textstyle\int_0^1\gamma'(\tau)\,\mathrm{d}\tau. \tag{1.6}$$

The main difficulty using this approach lies into the choice of a 'right' curve γ. Once a right γ is found, all subsequent proofs become convenient, easy and short. In general, we have no formal rules to describe a way to choose γ. In probability (see [1]) the choice of γ is quite a complicated problem since one wants to preserve along the curve numerous properties of a and b. In the setting of the present article the choices of γ are much simpler.

We note that our approach can be used to describe other situations, like cases where a natural convergence rate is $\mathrm{O}(n^{-\alpha})$, $0<\alpha<1$, instead of $\mathrm{O}(n^{-1})$.

Theorem 1.1 can be used to analyze the Trotter–Kato formula. One of the central results in approximation of semigroups is the following Chernoff-type theorem, stating that under some conditions on a family $F(t)$, $t\geqslant 0$, of contractions in a Banach space one has

$$\delta_n\stackrel{\text{def}}{=}\|F^n(t/n)^n-\mathrm{e}^{tC}\|\to 0, \quad \text{as } n\to\infty, \tag{1.7}$$

where C is a generator of a semigroup, defined as the closure of the strong derivative $F'(0)$. Important examples of the families F are

$$e^{tA}e^{tB} \quad \text{and} \quad e^{tB/2}e^{tA}e^{tB/2}, \qquad t \geqslant 0,$$

with unbounded and noncommuting operators A and B. In such cases one speaks about the so-called Trotter–Kato product formula, and there is a rich literature related to the convergence $\delta_n \to 0$ and upper bounds for δ_n, see [4–7, 10, 11, 14] and references therein. For self-adjoint operators in Hilbert spaces the optimal bound $O(n^{-1})$ for the error in the Trotter–Kato formula is obtained by Ichinose, *et al.* [11], using a very nice reduction to analysis of resolvents from [10]. The rate is the same as in the famous result of Sophus Lie for matrices A and B in finite dimensional space. In the infinite dimensional setting optimal error bounds so far are obtained only for self-adjoint operators in a Hilbert space for the following reason. Estimation of δ_n usually starts with the identity (see, for example, relation (2.3) in [10]

$$F^n(t/n) - e^{tC} = F^n(t/n) - e^{n(F(t/n)-I)} + e^{n(F(t/n)-I)} - e^{tC}.$$

It follows that

$$\delta_n \leqslant \Delta_n^* + \left\| e^{n(F(t/n)-I)} - e^{tC} \right\|, \quad \Delta_n^* \stackrel{\text{def}}{=} \| F^n(t/n) - e^{n(F(t/n)-I)} \|, \tag{1.8}$$

with δ_n introduced in (1.7). We note that in Theorems 1.1 and 1.2 the quantity Δ_n is a particular case of Δ_n^* with $F(t/n) = A$. For self-adjoint operators $\Delta_n = O(n^{-1})$ by an application of the spectral theorem. As far as we know, the bound $\Delta_n = O(n^{-1})$ of Theorem 1.1 presents the first optimal result for nonself-adjoint operators. We note, that Δ_n has the same form as in approximation of distributions of sums of independent identically distributed random variables by accompanying laws in probability.

The second term in (1.8) we intend to estimate in the nearest future, to appear elsewhere.

We conclude the introduction by providing an optimal error bound for Euler approximations of semigroups of operators in Banach spaces.

THEOREM 1.3. *Let* e^{-tA}, $t \geqslant 0$, *be a semigroup of operators in a Banach space. Assume that there exists a constant* K *independent of* n *and* t *such that*

$$n \| tA(I + tA)^{-n} \| \leqslant K, \tag{1.9}$$

and

$$\| e^{-tA} \| \leqslant K, \quad \| tAe^{-tA} \| \leqslant K, \tag{1.10}$$

for all $n = 1, 2, \ldots$ *and* $t \geqslant 0$. *Then we have*

$$\| (I + tA/n)^{-n} - e^{-tA} \| \leqslant 4K^3 n^{-1}.$$

Bounded holomorphic semigroups of operators satisfy conditions (1.9) and (1.10). Indeed, the inequality $\|(I+tA)^{-n}\| \leqslant K$ holds by Theorem 5.2 in [16]. Using integral representations provided in Remark 5.4 of Pazy's book, it is not difficult to show that this inequality implies (1.9). The first inequality in (1.10) holds according to the definition. The second one holds by Lemma 2.38 in [9].

The bound of Theorem 1.3 is optimal, which can be seen considering the special case where operators are real numbers.

Theorem 1.3 extends and refines a bit a result in [15], which was proved for m-sectorial operators in Hilbert spaces with a semi-angle $0<\alpha<\pi/2$. A proof in [15] has a probabilistic background related to induction type proofs in the Central Limit Theorem in probability theory. [15] improved to $\mathrm{O}(n^{-1})$ a bound $\mathrm{O}(n^{-1}\ln n)$ of [4]. For bounded holomorphic semigroups [3] obtained a bound $\mathrm{O}(n^{-1}\ln n)$. We are grateful to V. Cachia for early communication of his work. In the special case of m-sectorial operators, a comparison of the proof of Theorem 1.3 with those in [3,4,15] shows, that less than one page long proof in the present paper is indeed short. The proof can be even shortened if instead of the constant 4 we would take a generic absolute constant. Concluding, we note that our proof is extendible to analysis of other semigroups.

2. Proofs

We would like to thank a referee who proposed the next lemma and its proof.

LEMMA 2.1. *Assume that a bounded operator satisfies* $(n+1)\|A^n - A^{n+1}\| \leqslant K$, *for all* $n=0,1,2,\ldots$ *Then the operator* $B=I-A$ *satisfies*

$$\|tB\mathrm{e}^{-tB}\| \leqslant K, \; for \; t \geqslant 0. \tag{2.1}$$

Proof. Expanding e^{tA} into the Taylor series, we get

$$tB\mathrm{e}^{-tB} = \mathrm{e}^{-t}\sum_{n=0}^{\infty}\frac{t^{n+1}}{n!}(A^n - A^{n+1}).$$

Applying $\|A^n - A^{n+1}\| \leqslant K/(n+1)$ and summing the series, we derive (2.1). □

Proof of Theorem 1.1. We prove only (1.3) since the proof of (1.2) differs only in technical details. Write

$$\gamma(\tau) = u^n(\tau), \quad u(\tau) = (I-\tau B)\mathrm{e}^{-(1-\tau)B}.$$

Then we have $\gamma(0)=\mathrm{e}^{n(A-I)}$ and $\gamma(1)=A^n$. An application of the Newton–Leibnitz formula (1.6) yields

$$\Delta_n \leqslant \theta_1 + \theta_2 \quad \text{with} \quad \theta_1 \stackrel{\text{def}}{=} \int_0^{1/2} \|\gamma'(\tau)\|\,\mathrm{d}\tau, \qquad \theta_2 \stackrel{\text{def}}{=} \int_{1/2}^{1} \|\gamma'(\tau)\|\,\mathrm{d}\tau. \tag{2.2}$$

In view of (2.2), to prove the theorem, it suffices to show that (note that $5\ln 2 \leqslant 4$)

$$\theta_1 \leqslant \frac{K^2 \ln 2}{n}, \qquad \theta_2 \leqslant \frac{4K^2 \ln 2}{n}.$$

We have

$$\gamma'(\tau) = n u^{n-1}(\tau) u'(\tau), \quad u'(\tau) = -\tau B^2 \mathrm{e}^{-(1-\tau)B}$$

and

$$\gamma'(\tau) = -n(I - \tau B)^{n-1} \tau B^2 \mathrm{e}^{-(1-\tau)nB}. \tag{2.3}$$

Let us estimate θ_1. Regrouping the factors in (2.3), we get

$$\|\gamma'(\tau)\| \leqslant \varrho_1 \varrho_2 \quad \text{with } \varrho_1 = n\|\tau B(I-\tau B)^{n-1}\|, \quad \varrho_2 = \|B \mathrm{e}^{-(1-\tau)nB}\|.$$

Applying condition (1.1) we have $\varrho_1 \leqslant K$. The first inequality in (2.1) implies $\varrho_2 \leqslant K/(n(1-\tau))$. Hence $\varrho_1\varrho_2 \leqslant K^2/(n(1-\tau))$. Integrating over [0, 1/2], we derive $\theta_1 \leqslant (K^2 \ln 2)/n$.

Let us estimate θ_2. We consider here only the case of an even $n = 2m$ since the case of odd $n = 2m+1$ can be examined similarly splitting $n - 1 = m + m$ (in a similar proof of Theorem 1.3 we consider the case of odd n). Regrouping the factors in (2.3) and estimating $\|\mathrm{e}^{-sB}\| \leqslant 1$, for $s \geqslant 0$, we get

$$\|\gamma'(\tau)\| \leqslant \varrho_3 \varrho_4 \quad \text{with } \varrho_3 = n\|\tau B(I-\tau B)^{m}\|, \quad \varrho_4 = \|B(I-\tau B)^{m-1}\|.$$

Condition (1.1) yields

$$\varrho_3 \leqslant \frac{Kn}{m+1} \quad \text{and} \quad \varrho_4 \leqslant \frac{K}{\tau m}.$$

We note that $n/(m(m+1)) \leqslant 4$. Hence, we have $\varrho_3\varrho_4 \leqslant 4K^2/n\tau$. Integrating over the interval [1/2, 1], we get $\theta_2 \leqslant (4K^2 \ln 2)/n$. □

Proof of Theorem 1.2. The proof repeats that of Theorem 1.1. The only one difference is that now we use $\|\mathrm{e}^{-sB}\| \leqslant K_0$ instead of $\|\mathrm{e}^{-sB}\| \leqslant 1$ in estimation of θ_2. This leads to the estimate $\theta_2 \leqslant (4K_0K^2 \ln 2)/n$. Noting that $K_0 \geqslant 1$, we see that as a final bound we can choose $4K_0K^2/n$. □

Proof of Theorem 1.3. The design of the proof is very similar to that of Theorem 1.1. Write $C = tA/n$ and

$$\gamma(\tau) = u^n(\tau), \quad u(\tau) = (I + \tau C)^{-1} \mathrm{e}^{-(1-\tau)C}.$$

Then we have $\gamma(0) = \mathrm{e}^{-tA}$ and $\gamma(1) = (I + tA/n)^{-n}$. An application of the Newton–Leibnitz formula (1.6) yields

$$\Delta_n \leqslant \theta_1 + \theta_2 \quad \text{with } \theta_1 \stackrel{\text{def}}{=} \int_0^{1/2} \|\gamma'(\tau)\|\,\mathrm{d}\tau, \quad \theta_2 \stackrel{\text{def}}{=} \int_{1/2}^{1} \|\gamma'(\tau)\|\,\mathrm{d}\tau. \tag{2.4}$$

In view of (2.4), to prove the theorem, it suffices to show that (note that $K \geqslant 1$ and $5\ln 2 \leqslant 4$)

$$\theta_1 \leqslant \frac{K^2 \ln 2}{n}, \qquad \theta_2 \leqslant \frac{4K^3 \ln 2}{n}.$$

We have

$$\gamma'(\tau) = nu^{n-1}(\tau)u'(\tau), \qquad u'(\tau) = \tau C^2 (I + \tau C)^{-2} \mathrm{e}^{-(1-\tau)C}$$

and

$$\gamma'(\tau) = n\tau C^2 (I + \tau C)^{-n-1} \mathrm{e}^{-(1-\tau)tA}. \tag{2.5}$$

Let us estimate θ_1. Regrouping the factors in (2.5), we get

$$\|\gamma'(\tau)\| \leqslant \varrho_1 \varrho_2 \quad \text{with} \quad \varrho_1 = n\|\tau C(I + \tau C)^{-n-1}\|, \quad \varrho_2 = \|C\mathrm{e}^{-(1-\tau)tA}\|.$$

Applying condition (1.9) we have $\varrho_1 \leqslant K$. An application of condition (1.10) shows that $\varrho_1 \leqslant K/(n(1-\tau))$. Hence $\varrho_1\varrho_2 \leqslant K^2/(n(1-\tau))$. Integrating over the interval $[0, 1/2]$, we derive $\theta_1 \leqslant (K^2 \ln 2)/n$.

Let us estimate θ_2. We consider only the case of an odd $n = 2m + 1$. Regrouping the factors in (2.5) and using $\|\mathrm{e}^{-sA}\| \leqslant K$ (see (1.10)), we get

$$\|\gamma'(\tau)\| \leqslant \varrho_3 \varrho_4 \quad \text{with} \quad \varrho_3 = \|\tau C(I + \tau C)^{-m-1}\|, \quad \varrho_4 = Kn\varrho_3/\tau.$$

Condition (1.9) yields $\varrho_3 \leqslant K/(m+1)$. We note that $n/(m+1)^2 \leqslant 4$. Hence, we have $\varrho_3\varrho_4 \leqslant 4K^3/(n\tau)$. Integrating $[1/2, 1]$, we get $\theta_2 \leqslant (4K^3 \ln 2)/n$. □

Acknowledgement

The authors are grateful to the anonymous referees whose remarks and suggestions improved the presentation of the results.

References

1. Bentkus, V.: A new method for approximations in probability and operators theories, (in Russian, an English translation in *Lithuanian Math. J.*), *Liet. matem. rink.* **43**(4) (2003), 444–470.
2. Bentkus, V. and Paulauskas, V.: Optimal error estimates in operator-norm approximations of some semigroups, http:/www.mathematik.uni-bielefeld.de/fgweb/preserv.html, a preprint of Bielefeld University 03–053 (2003).
3. Cachia, V.: Euler's exponential formula for non C_0 semigroups, *Semigroup Forum* **68** (2004) 1–24.
4. Cachia, V. and Zagrebnov V. A.: Operator-norm approximation of semigroups by quasi-sectorial contractions, *J. Funct. Anal.* **180** (2001), 176–194.
5. Cachia, V. and Zagrebnov, V. A.: Operator-norm convergence of the Trotter product formula for holomorphic semigroups, *J. Operator Theory* **46**(1) (2001), 199–213.

6. Cachia, V., Neidhardt, H. and Zagrebnov, V.: Comments on the Trotter product formula error-bound estimates for nonself-adjoint semigroups, *Integral Equations Operator Theory* **42** (2002), 425–448.
7. Cachia, V., Neidhardt, H. and Zagrebnov, V.: Accretive perturbations and error estimates for the Trotter product formula, *Integral Equations Operator Theory* **39**(4) (2001), 396–412.
8. Chernoff, P. R.: Note on product formulas for operator semigroups, *J. Funct. Anal.* **2** (1968), 238–242.
9. Davies, E. B.: *One-Parameter Semigroups*, Academic Press, London, 1980, pp. 1–230.
10. Ichinose, T. and Tamura, Hideo: The norm convergence of the Trotter–Kato product formula with error bound, *Comm. Math. Phys.* **217** (2001), 489–502.
11. Ichinose, T., Tamura, Hideo, Tamura, Hiroshi and Zagrebnov, V. A.: Note on the Paper 'The norm convergence of the Trotter–Kato product formula with error bound' by Ichinose and Tamura, *Comm. Math. Phys.* **221** (2001), 499–510.
12. Kato, T.: *Perturbation Theory for Linear Operators*, Springer-Verlag, Berlin, 1966.
13. Nagy B. and Zemánek, J.: A resolvent condition implying power boundedness, *Studia Math.* **134**(2) (1999), 143–151.
14. Neidhardt, H. and Zagrebnov, V.: Trotter–Kato product formula and operator–norm convergence, *Comm. Math. Phys.* **205** (1999), 129–159.
15. Paulauskas, V.: On operator-norm approximation of some semigroups by quasi-sectorial operators, *J. Funct. Anal.* **207**(1) (2003), 58–67.
16. Pazy, A.: *Semigroups of Linear Operators and Applications to Partial Differential Equations*, Appl. Math. Sci. 44, Springer-Verlag, New York, 1983.
17. Ritt, R. K.: A condition that $\lim_{n\to\infty} T^n = 0$, *Proc. Amer. Math. Soc.* **4** (1953), 898–899.