Probabilistic version of Trotter's exponential product formula in Banach algebras

By A. LENARD in Bloomington (Indiana, U.S.A.)

1. Introduction and results

It is an elementary fact that the exponential function may be defined by the equivalent formulae

$$\exp(x) = \lim_{n\to\infty}\left(1+\frac{1}{n}x\right)^n = \sum_{n=0}^{\infty}\frac{1}{n!}x^n$$

not only when x is a real or complex number but also when it is a matrix with real or complex entries or a bounded operator acting on a Hilbert space or a Banach space or, even when it is an element of an abstract Banach algebra $\mathfrak{B}$ with identity I (for a definition of a Banach algebra see for instance [1]). If $\mathfrak{B}$ is not commutative then in general $\exp(x)\exp(y) \neq \exp(x+y)$. There is, however, a formula which replaces the addition law of the exponential function, namely

$$\lim_{n\to\infty}\left(\exp\left(\frac{x}{n}\right)\exp\left(\frac{y}{n}\right)\right)^n = \exp(x+y) \tag{1}$$

and this holds regardless whether x and y commute or not. Formula (1) is capable of further generalization; see TROTTER [2]. Specifically, x and y may be unbounded operators of a certain type, namely generators of continuous one-parameter operator semi-groups. In the present paper we are not concerned with Trotter's generalization, but we shall still refer to (1) as the Trotter product formula. The symbols $x, y, \ldots, a, b, \ldots$ shall generally denote elements of the Banach algebra $\mathfrak{B}$. The norm of $x\in\mathfrak{B}$ is written $\|x\|$.

Let $\mathbf{X}=(x_1, x_2, \ldots, x_m$ be any finite sequence of elements of $\mathfrak{B}$. With any such sequence we associate the product

$$T(\mathbf{X}) = \exp\left(\frac{x_1}{m}\right)\exp\left(\frac{x_2}{m}\right)\ldots\exp\left(\frac{x_m}{m}\right)$$

which will be called its Trotter product. Note that it depends essentially on the order of the factors, i.e. on $\mathbf{X}$ as a sequence, not merely as a set. We also write, for the mean of the elements of $\mathbf{X}$, $M(\mathbf{X}) = \frac{1}{m}(x_1 + x_2 + \cdots + x_m)$. Using this notation we can express the Trotter product formula as follows: If $\mathbf{X}_k$ (for $k = 1, 2, \ldots$) is the sequence of length $2k$ whose elements are alternatingly x and y, then

$$\lim_{k\to\infty} T(\mathbf{X}_k) = \exp\left(\lim_{k\to\infty} M(\mathbf{X}_k)\right). \tag{2}$$

We now raise the following question. Under what natural conditions on a sequence $\mathbf{X}_1, \mathbf{X}_2, \ldots$ (subject to the requirement that their lengths tend to infinity) will (2) hold?

We shall prove two theorems relevant to this question. The first theorem gives a rather general sufficient condition for an infinite sequence $\mathbf{X}_1, \mathbf{X}_2, \ldots$ to satisfy (2). The original Trotter formula is an obvious consequence of this condition. But our theorem also shows that to a certain extent the order of the factors in the Trotter product may be made subject to considerable rearrangement without destroying the validity of (2).

For any $\mathbf{X} = (x_1, x_2, \ldots, x_m)$ let us write $\varrho = \varrho(\mathbf{X}) = \mathrm{Max}_{1\le j\le m} \|x_j\|$. Let π denote a partition of the sequence $(1, 2, \ldots, m)$ into successive subsequences $(1, 2, \ldots, m_1)$, $(m_1+1, m_1+2, \ldots, m_1+m_2)$, $\ldots$, $(m-m_s+1, m-m_s+2, \ldots, m)$, and let $\mathbf{Y}_1, \mathbf{Y}_2, \ldots, \mathbf{Y}_s$ be the corresponding subsequences of $\mathbf{X}$. For any element $g \in \mathfrak{B}$ we introduce a quantity $\delta = \delta(\mathbf{X}, \pi, g)$ which measures the closeness to which g uniformly approximates the "partial means" of $\mathbf{X}$ induced by the partition π

$$\delta(\mathbf{X}, \pi, g) = \mathrm{Max}_{1\le j\le s} \|M(\mathbf{Y}_j) - g\|. \tag{3}$$

We also define a quantity

$$\eta = \eta(\pi) = \sum_{j=1}^{s} \left(\frac{m_j}{m}\right)^2. \tag{4}$$

If $\eta_0 = \max_{1\le j\le s} \left(\frac{m_j}{m}\right)$, we have clearly

$$\eta_0^2 \le \eta \le \eta_0, \tag{5}$$

so that η is a kind of a measure for the relative fineness of π.

Theorem 1. *Let $\mathbf{X}_1, \mathbf{X}_2, \ldots$ be an infinite sequence of finite sequences of elements of $\mathfrak{B}$. Suppose that $\varrho(\mathbf{X}_k)$ is bounded. Suppose that $g \in \mathfrak{B}$ exists and a sequence of partitions π_k of $\mathbf{X}_k$ into successive subsequences exists such that $\eta(\pi_k) \to 0$ and $\delta(\mathbf{X}_k, \pi_k, g) \to 0$ as $k \to \infty$. Then $M(\mathbf{X}_k) \to g$ and $T(\mathbf{X}_k) \to \exp(g)$ as $k \to \infty$.*

Note that the quantity $\delta(\mathbf{X}, \pi, g)$ is independent of the order of the elements x_i *within* each of the subsequences $\mathbf{Y}_j$ $(j = 1, 2, \ldots, s)$. Thus our theorem formulates mathematically the intuitive notion that the order of the factors in the Trotter product $T(\mathbf{X})$ is irrelevant "locally" and only essential "in the large".

Our second Theorem deals with the following problem. Suppose we consider infinite sequences $\mathbf{Z}=(x_1, x_2, \ldots)$ whose elements are all taken from a fixed finite subset $\{a_1, a_2, \ldots, a_\sigma\}$ of $\mathfrak{B}$. With any such infinite sequence we may consider the sequence of its finite sections $\mathbf{X}_1=(x_1)$, $\mathbf{X}_2=(x_1, x_2)$, ... Is it possible to characterize the extent of the set of those infinite sequences $\mathbf{Z}$ for which the generalized Trotter product formula (2) holds?

This question leads to measure-theoretic considerations, and can be naturally formulated in probabilistic terminology.

Theorem 2. *Let* $a_1, a_2, \ldots, a_\sigma \in \mathfrak{B}$, *and let* $p_1, p_2, \ldots, p_\sigma \geqq 0$, $\Sigma_\nu p_\nu = 1$. *Suppose an infinite (random) sequence* $\mathbf{Z}=(x_1, x_2, \ldots)$ *is formed by choosing, independently for each* $j \geqq 1$, $x_j = a_\sigma$ *with probability* p_σ. *Let* $g = p_1 a_1 + p_2 a_2 + \cdots + p_\sigma a_\sigma$. *Then the probability is unity that* $M(\mathbf{X}_k) \to g$ *and* $T(\mathbf{X}_k) \to \exp(g)$ *as* $k \to \infty$.

Thus in the sense of the given probability measure defined on the set of sequences $\mathbf{Z}$, for almost every sequence the generalized Trotter product formula holds.

2. An auxiliary inequality

Both theorems derive from an elementary estimate formulated as follows:

Lemma. *Let* $\mathbf{X}$ *be a finite sequence of elements of* $\mathfrak{B}$, π *any partition of it into successive subsequences, and* $g \in \mathfrak{B}$. *Let* $\varrho=\varrho(\mathbf{X})$, $\eta=\eta(\pi)$ *and* $\delta=\delta(\mathbf{X}, \pi, g)$ *be defined as above. Then*

$$\|T(\mathbf{X}) - \exp(g)\| \leqq e^{\|g\|}\left(e^{\delta+\frac{1}{2}\eta\varrho^2 e^{\varrho}} - 2e^{-\frac{1}{2}\eta\|g\|^2} + 1\right).$$

In the proof of the Lemma we shall make use of the following facts:

(A) If $P(x_1, x_2, \ldots, x_q)$ is a polynomial or a power series with non-negative real coefficients then

$$\|P(x_1, x_2, \ldots, x_q)\| \leqq P(\|x_1\|, \|x_2\|, \ldots, \|x_q\|)$$

for all $x_j \in \mathfrak{B}$ such that the right hand side is finite.

(B) For any $t \geqq 0$, $e^t - 1 - t \leqq \frac{1}{2} t^2 e^t$.

(C) For any $t \geqq 0$, $e^{t-\frac{1}{2}t^2} \leqq 1 + t \leqq e^t$.

Let $\mathbf{Y}_j$ $(j=1, 2, \ldots, s)$ be the subsequences of $\mathbf{X}$ produced by the partition π, and let m_j be their respective length, $m=\Sigma_j m_j$. Let $T(\mathbf{X})=y_1 y_2 \ldots y_s$ where y_j is the product of those factors in the product, taken in their proper order, which involve the elements $x_k \in \mathbf{Y}_j$. Write $y_j = 1+\frac{m_j}{m} g+r_j$, thus defining r_j. For notational convenience we now consider $j=1$. We have

$$r_1 = \left[y_1-1-\frac{m_1}{m} M(\mathbf{Y}_1)\right]+\frac{m_1}{m}[M(\mathbf{Y}_1)-g].$$

The norm of the second term is bounded by $\frac{m_1}{m}\,\delta$. To obtain a bound on the norm of the first term we note that if

$$y_1-1-\frac{m_1}{m} M(\mathbf{Y}_1) = \exp\left(\frac{x_1}{m}\right)\exp\left(\frac{x_2}{m}\right)\dots\exp\left(\frac{x_{m_1}}{m}\right)-1-\frac{1}{m}(x_1+x_2+\dots+x_{m_1})$$

is regarded as a power series in the x_j ($j=1, 2, \dots, m_1$) it has non-negative coefficients (the negative terms cancel!). Thus by principle (A) above we may replace x_j by $\|x_j\|$, and then taking (B) and the definition of ϱ into account we get

$$\left\|y_1-1-\frac{m_1}{m} M(\mathbf{Y}_1)\right\| \leqq \frac{1}{2}\left(\frac{m_1}{m}\right)^2 \varrho^2 e^{\varrho}.$$

Thus we have for $j=1, 2, \dots, s$

$$\|r_j\| \leqq \frac{1}{2}\left(\frac{m_j}{m}\right)^2 \varrho^2 e^{\varrho}+\frac{m_j}{m}\,\delta. \tag{6}$$

Next, let $z_j = 1+\frac{m_j}{m} g$, and consider the difference $T(\mathbf{X})-z_1 z_2 \dots z_s = y_1 y_2 \dots y_s - z_1 z_2 \dots z_s$. As a polynomial in g and the r_j, it has again non-negative coefficients, so we apply principle (A). The norms of r_j are majorized by (6), and so by the inequality (C)

$$1+\frac{m_j}{m}\|g\|+\|r_j\| \leqq e^{\frac{m_j}{m}\|g\|+\frac{1}{2}\left(\frac{m_j}{m}\right)^2 \varrho^2 e^{\varrho}+\frac{m_j}{m}\delta},$$

and

$$-\left(1+\frac{m_j}{m}\|g\|\right) \leqq -e^{\frac{m_j}{m}\|g\|-\frac{1}{2}\left(\frac{m_j}{m}\right)^2\|g\|^2},$$

where in the last step we used $\frac{m_j}{m}\|g\| \leqq \sqrt{\eta}\,\|g\| \leqq 1$. Thus we see that

$$\|y_1 y_2 \dots y_s - z_1 z_2 \dots z_s\| \leqq e^{\|g\|}\left(e^{\delta+\frac{1}{2}\eta\varrho^2 e^{\varrho}}-e^{-\frac{1}{2}\eta\|g\|^2}\right).$$

Arguing analogously, we have also

$$\|\exp(g)-z_1 z_2 \dots z_s\| \leqq e^{\|g\|}\left(1-e^{-\frac{1}{2}\eta\|g\|^2}\right).$$

The last two inequalities together imply the conclusion of the lemma.

We note that Theorem 1 is an immediate consequence of the lemma, since the estimate for $\|T(\mathbf{X}_k)-\exp(g)\|$ supplied by the lemma tends to zero as $k\to\infty$ if the hypotheses of the theorem are fulfilled.

3. Proof of Theorem 2

The idea of the proof of Theorem 2 is to find an appropriate sequence of partitions π_k $(k=1, 2, 3, \ldots)$ such that if we let $\delta_k=\delta(\mathbf{X}_k, \pi_k, g)$ then

(7) $$\mathbf{P}\{\lim_{k\to\infty}\delta_k = 0\} = 1,$$

and at the same time such that

(8) $$\lim_{k\to\infty}\eta(\pi_k) = 0.$$

Indeed, if (7) and (8) are fulfilled then the conclusion of Theorem 2 follows from Theorem 1.

Let $C_1 < C_2$ and $\beta<1$ be three positive constants. We define the partition π_k of $(1, 2, 3, \ldots, k)$ into successive subsequences of lengths $m_j=m_j(k)$ $(j=1, 2, \ldots, s=s(k))$ in such a manner that for all j and k

(9) $$C_1 k^\beta < m_j < C_2 k^\beta.$$

Since $\Sigma_j m_j=k$, it follows that

(10) $$s=s(k)=O(k^{1-\beta}),$$

and therefore

(11) $$\eta_k = \eta(\pi_k) = \sum_{j=1}^{s(k)}\left(\frac{m_j}{k}\right)^2 = O(k^{\beta-1}),$$

so that (8) holds. Note that in our probabilistic set-up the partitions π_k are not random variables (i.e. π_k is constant over the whole probability space).

Next, we remark that, by virtue of the Borel—Cantelli lemma, in order to prove (7) it is sufficient to prove that

(12) $$\sum_{k=1}^{\infty}\mathbf{P}\{\delta_k \geqq \varepsilon\} < \infty$$

for any positive ε.

Let $\mathbf{Y}_1, \mathbf{Y}_2, \ldots, \mathbf{Y}_s$ be the successive subsequences of $\mathbf{X}_k$ produced by the partition π_k. Let $N_{j\nu}$ $(j=1, 2, \ldots, s;\ \nu=1, 2, \ldots, \sigma)$ be the number of occurrences of a_ν among the elements of the subsequence $\mathbf{Y}_j$. The $N_{j\nu}$ are random variables subject to the multinomial distribution determined by the probabilities p_ν, and for different j they are independent. We have

$$\delta_k = \operatorname{Max}_{1\leqq j\leqq s}\left\|\sum_{\nu=1}^{\sigma}\left(\frac{N_{j\nu}}{m_j}-p_\nu\right)a_\nu\right\| \leqq AM_k,$$

where

$$A = \mathrm{Max}_\nu \|a_\nu\| \quad \text{and} \quad M_k = \mathrm{Max}_{1 \leqq j \leqq s} \sum_{\nu=1}^{\sigma} \left| \frac{N_{j\nu}}{m_j} - p_\nu \right|.$$

Thus we need to show

$$\sum_{k=1}^{\infty} \mathbf{P}\{M_k \geqq \varepsilon\} < \infty. \tag{13}$$

Since the following inclusion (implication) of events holds

$$\{M_k \geqq \varepsilon\} = \bigcup_{j=1}^{s(k)} \left\{ \sum_{\nu=1}^{\sigma} \left| \frac{N_{j\nu}}{m_j} - p_\nu \right| \geqq \varepsilon \right\} \subseteqq \bigcup_{j=1}^{s(k)} \bigcup_{\nu=1}^{\sigma} \left\{ \left| \frac{N_{j\nu}}{m_j} - p_\nu \right| \geqq \frac{\varepsilon}{\sigma} \right\},$$

we obtain for the probabilities of the complementary events

$$\mathbf{P}\{M_k < \varepsilon\} \geqq \mathbf{P} \bigcap_{j=1}^{s(k)} \bigcap_{\nu=1}^{\sigma} \left\{ \left| \frac{N_{j\nu}}{m_j} - p_\nu \right| < \frac{\varepsilon}{\sigma} \right\} = \tag{14}$$

$$= \prod_{j=1}^{s(k)} \mathbf{P} \bigcap_{\nu=1}^{\sigma} \left\{ \left| \frac{N_{j\nu}}{m_j} - p_\nu \right| < \frac{\varepsilon}{\sigma} \geqq \prod_{j=1}^{s(k)} \left[1 - \sum_{\nu=1}^{\sigma} \mathbf{P} \left\{ \left| \frac{N_{j\nu}}{m_j} - p_j \right| \geqq \frac{\varepsilon}{\sigma} \right\} \right].$$

The equality in (14) is due to the fact that we are dealing with the intersection (conjunction) of independent events.

Suppose N is the number of "successes" in a sequence of m Bernoulli trials with probability p for success. We have then the following fact [3]: given any $\alpha > 1$, for all sufficiently large m (depending only on α and p)

$$\mathbf{P} \left\{ \frac{|N - mp|}{[mp(1-p)]^{1/2}} \geqq (2\alpha \log m)^{1/2} \right\} < \frac{1}{m^\alpha}.$$

It follows that for any $\varepsilon > 0$

$$\mathbf{P} \left\{ \left| \frac{N}{m} - p \right| \geqq \varepsilon \right\} < \frac{1}{m^\alpha} \tag{15}$$

for all sufficiently large m (depending only on ε, α and p). If the inequality (15) is used for the probabilities on the right hand side of (14) we obtain

$$\mathbf{P}\{M_k < \varepsilon\} \geqq \prod_{j=1}^{s(k)} \left[1 - \frac{\sigma}{m_j^\alpha} \right]$$

which is valid for all sufficiently large k, since (9) implies that then all the m_j are large enough. But (9) and (10) show that for suitable constants C and C'

$$\mathbf{P}\{M_k < \varepsilon\} \geqq (1 - Ck^{-\alpha\beta})^{C' k^{1-\beta}} = 1 - O(k^{-\gamma}) \tag{16}$$

where $\gamma = \alpha\beta + \beta - 1$. Since $\alpha > 1$ was arbitrary we may suppose $\gamma > 1$, so that $\mathbf{P}\{M_k \geqq \varepsilon\} = O(k^{-\gamma})$ and (13) follows. This concludes the proof of Theorem 2.

References

[1] K. YOSIDA, *Functional Analysis* (Heidelberg—New York, 1968), Chapter XI.

[2] H. TROTTER, On the product of semi-groups of operators, *Proc. Amer. Math. Soc.*, **10** (1959), 545—551.

[3] W. FELLER, *Introduction to Probability Theory and its Applications.* I (New York, 1957). See Section VIII. 4, especially equ. (4. 5).

DEPARTMENT OF MATHEMATICS
INDIANA UNIVERSITY
BLOOMINGTON, INDIANA 47401

(Received April 22, 1970)