OPERATOR-LIMIT DISTRIBUTIONS IN PROBABILITY THEORY

ZBIGNIEW J. JUREK · J. DAVID MASON

Wiley Series in Probability and Mathematical Statistics:
Probability and Mathematical Statistics Section—

Operator-Limit Distributions in Probability Theory

ZBIGNIEW J. JUREK

Institute of Mathematics
University of Wrocław
Wrocław, Poland

J. DAVID MASON

Department of Mathematics
The University of Utah
Salt Lake City, Utah

A Wiley-Interscience Publication
JOHN WILEY & SONS, INC.
New York • Chichester • Brisbane • Toronto • Singapore

This text is printed on acid-free paper.

Library of Congress Cataloging in Publication Data:

Jurek, Zbigniew J.
Operator-limit distributions in probability theory/Zbigniew J. Jurek and J. David Mason.
p. cm.—(Wiley series in probability and mathematical statistics. Probability and mathematical statistics)
"A Wiley-Interscience publication."
Includes bibliographical references and index.
ISBN 0-471-58595-5 (acid-free)
1. Central limit theorem. I. Mason, J. David (Jesse David), 1935- . II. Title. III. Series.
QA273.67.J87 1993
519.2—dc20 93-7158

Printed in the United States of America

10 9 7 6 5 4 3 2 1

Symbols and Notations

	Defined on Page		Defined on Page
$\langle \cdot , \cdot \rangle$	3	$\mathrm{ID}(X)$	28
$\Vert \cdot \Vert_{\mu}$	223	$\mathrm{ID}_{\log}(X)$	36
$\Vert \cdot \Vert_{\mu, B}$	223	$\mathbf{Inv}(\mu)$	50
$\Vert \cdot \Vert_{Q}$	92	$\ker(A)$	6
$\mathbf{A}(\mu)$	53	$\mathscr{K}_Q$	110
$\langle A; a \rangle$	24	$\mathrm{lin}(A)$	105
$\mathscr{A}_I$	60	$\mathrm{lin}_Q(A)$	105
$\mathbf{A}_J$	77	$\mathscr{L}(\xi)$	22
$[a, R, M]$	33	L_M	94
$\mathbf{Aut}(X)$	4	$L_0(Q)$	108
$\mathscr{A}(X)$	24	$L(X, Y)$	4
$\mathscr{B}(S)$	20	$\hat{\mu}$	25
$C_b(S)$	20	μ	26
$C_0(\mathbb{R}^d)$	4	μ^0	26
$\mathscr{D}(\Gamma)$	10	$\mu_n \Rightarrow \mu$	20
$\mathbf{D}(\mu)$	53	$\mu * \nu$	23
$\mathbf{D}_J(\mu)$	77	$\mathscr{O}$	18
$\mathrm{DONA}(\mu)$	240	$\mathrm{OL}(\mathbb{R}^d)$	72
$D(S; [0, \infty))$	42	Π_{x^*}	32
$D(S; [a, b])$	40	$\mathscr{P}(S)$	20
$\mathbf{End}(X)$	4	$\mathscr{Q}$	18
$e(m)$	30	$\mathrm{sem}\{a\}$	2
$\mathscr{E}_u(\mu)$	85	$\mathrm{sem}\ A$	2
$\mathscr{E}_{cu}(\mu)$	88	$\mathrm{sem}\{\mathbf{F}\}$	75
$\exp(A)$	15	S_Q	93
$[F; G]$	196	$\mathscr{T}(\mathbb{H})$	15
$\mathscr{F}(\mathbb{R}^d)$	47	$\mathscr{T}_Q$	127
$\mathrm{GDOA}(\mu)$	239	$\mathscr{T}_B$	198

Preface

The theory of limit distributions occupies a central place in probability and mathematical statistics. It describes limit phenomena of triangular series or sequences of independent random observations. Typically, one looks at statistics (functions) constructed from given random variables. Often these functions are assumed to be linear. Therefore, this leads to consideration of limit distributions of the sequences

$$a_n(\xi_1 + \cdots + \xi_n) + x_n, \qquad n \geq 1, \tag{0.1.1}$$

where $\xi_1, \xi_2, \ldots$ are independent $\mathbb{R}$-valued random variables and a_n and x_n are real constants. But often when the observations $\xi_1, \xi_2, \ldots$ are random vectors ($\mathbb{R}^d$ or Banach space valued), the normalization in (0.1.1) is still done by *scalars*. In such a setting, we have coordinate-wise one-dimensional problems again. One should allow the interaction between coordinate normalization in (0.1.1) to be consistent with the structure of $\mathbb{R}^d$, that is, one should allow normalization by arbitrary *linear operators*. Thus, the main aims or principles for this book are:

1. Present a theory of limit distributions of sequences

$$A_n(\xi_1 + \cdots + \xi_n) + x_n, \qquad n \geq 1, \tag{0.1.2}$$

   where $\xi_1, \xi_2, \ldots$ are $\mathbb{R}^d$-valued random vectors and $A_1, A_2, \ldots$ are matrices (operators) on $\mathbb{R}^d$. This explains the title, *Operator-Limit Distributions in Probability Theory*.

2. Present proofs which do not appeal to the one-dimensional results of (0.1.1). In other words, this exposition is much more "functional" or "coordinate-free."

3. Present complete exposition for $\mathbb{R}^d$ and indicate the essential differences for infinite-dimensional Banach space valued random variables.

It seems natural that the normalization of the random vectors be consistent with the algebraic structure of the space in which they take their values. Hence, for the real line one uses scalars, for a linear space one uses linear operators, and for a group one uses group endomorphisms. (As far as we know, the group case is completely open. For probabilities on groups, convolution powers are usually investigated and often the normalized Haar measure is the only limit measure.)

Limits of (0.1.1) are called *stable* distributions if $\xi_1, \xi_2, \ldots$ are independent and identically distributed. If ξ_n's are independent and the triangular array $\{a_n \xi_j: 1 \le j \le n\}$ is infinitesimal, then limits of (0.1.1) are called *selfdecomposable* distributions (or Lévy class L_0 distributions). Stable measures have been extensively investigated for the last 60 years or so [cf. Gnedenko and Kolmogorov (1954), Zolotarev (1986), and Linde (1986)]. The class L_0 of selfdecomposable measures contains the class of stable laws. It is related to autoregressive sequences, that is, a sequence $\{X_n\}$ such that $X_{n+1} \stackrel{d}{=} cX_n + \varepsilon_n$, whose $0 < c < 1$, X_n's are independent of ε_n's. Furthermore, L_0 distributions arise in limits of Ising models for ferromagnetism in statistical physics [cf. deConinck (1984)]. Elements from L_0 can be viewed as limits as $t \to \infty$ of some stochastic processes which are given by random integrals and are similar to Ornstein–Uhlenbeck processes (cf. Section 3.6). Finally, there is also the notion of self-similar processes. These processes satisfy $\{X(at): t \ge 0\}$ and $\{a^H X(t): t \ge 0\}$ have the same finite-dimensional distributions, where H is a scalar (or a matrix). In the case when the process has stationary independent increments, the distribution of $X(1)$ is (operator) stable, and in the case that the increments are only independent, the distribution of $X(1)$ is (operator) selfdecomposable [cf. Lamperti (1962, 1972), Hudson and Mason (1982), and Sato (1991)].

Limit distributions of sequences of the form (0.1.2) will be called *operator-stable* and *operator-selfdecomposable* when $\xi_1, \xi_2, \ldots$ are i.i.d. or only independent and infinitesimally small, respectively. Sakovic (1961, 1965), Ph.D. dissertation under B. V. Gnedenko, was the first to investigate operator-stable laws on $\mathbb{R}^d$. On the other hand, Fisz (1954) proved a theorem on the convergence of operator-types (normalization by matrices). Both, Fisz and Sakovic used a coordinate-wise approach which led to some computational difficulties and which perhaps explains why this direction was not continued later (cf. the above aim 2). Independently of Sakovic, Sharpe (1969), Ph.D. dissertation under S. Kakutani, gave a complete characterization of operator-stable measures on $\mathbb{R}^d$ and his proof used some algebraic methods. Similarly, Urbanik (1972) gave a "functional" proof for his description of operator-selfdecomposable measures. This was the beginning of a period of extensive study of operator-limit distributions. This functional approach also prompted the development of new purely algebraic methods (decomposability semigroups) in probability. This book attempts to summarize that period of investigation.

Chapter 1 is of an auxiliary character. It compiles well-known facts from the theory of limit distributions (infinitely divisible measures, weak convergence, Skorohod spaces) without proofs. On the other hand, algebraic facts (Numakura theorem, Lie groups, one-parameter semigroups of operators) are given with proofs. Bibliographic comments will always be given at the end of each chapter.

To be able to study the limits of (0.1.1) or (0.1.2), one needs theorems on the relationship between the limit distribution of the pair of sequences

$$\{\xi_n\} \quad \text{and} \quad \{A_n\xi_n + x_n\}. \tag{0.1.3}$$

We refer to these as convergence of *operator types* and these are proved in Chapter 2. Besides that, Chapter 2 contains theorems on decomposability and symmetry semigroups associated with probability measures. These semigroups will be the fundamental tools in Chapters 3 and 4, but they are of interest in themselves and therefore are presented in separate sections. The point is that some topological and algebraic properties of these semigroups are equivalent to measure theoretical properties. Examples in Section 2.5 indicate the "trouble" points when one deals with random variables having values in infinite-dimensional linear spaces. The last two sections specialize some results to one-dimensional and infinite-dimensional spaces. Bibliographic comments are given in Section 2.8.

Chapter 3 deals with operator-selfdecomposable measures, that is, limits of (0.1.2) with independent ξ_n's and infinitesimal triangular array $\{A_n\xi_j: 1 \le j \le n, n \ge 1\}$. First, properties of norming sequences $\{A_n\}$ are discussed, in particular, semigroups of operators generated by them. Then the main Urbanik decomposability theorem for operator-selfdecomposable measures, Theorem 3.3.5, is proved. Their characteristic functions are described in Section 3.4. Section 3.6 establishes random integral representations of operator-selfdecomposable probabilities, thus showing the connection with Ornstein–Uhlenbeck type processes. Subsequently, infinitesimal generators for such processes are found. Section 3.8 proves absolute continuity of full $\exp(-tQ)$-decomposable measures on $\mathbb{R}^d$, whereas Section 3.9 deals with selfdecomposable measures on arbitrary Banach spaces.

In Chapter 4, we investigate operator-stable distributions, that is, limits of (0.1.2) with i.i.d. sequences $\xi_1, \xi_2, \ldots$. The main characterization due to Sharpe (partially due to Sakovic) is proved in Theorem 4.2.12. Structural characterization of operator-stable distributions and operator exponents of such measures are proved in Section 4.6. Then commuting exponents and elliptically symmetric measures are discussed. Section 4.11 gives descriptions of the domain of normal attraction, that is, normalization is by operators of the form n^{-B} for some operator exponent B. In the case of the *generalized* domain of attraction, that is, using arbitrary norming operators A_n, we decided to omit the results. This general case is presently solved in the finite-dimensional setting, but it is dealt with by appealing to knowledge of one-dimensional results and applying them uniformly in every direction [cf. Hahn and Klass (1981, 1985) and Griffin (1986)]. This is not in the spirit of our aim 2 and therefore is not included here. The existence of moments of operator-stable laws are presented in Section 4.12, whereas the existence of a complete set of independent univariate marginals is given in Section 4.13. The special case of the usual multivariate stable laws is presented in Section 4.14. Chapter 4 closes by considering some special cases and with bibliographic comments.

The book ends with an epilogue in which we briefly discuss some areas of operator-limit distributions theory which are omitted, in particular, operator-semistability, $\mathscr{G}$-stability for a group $\mathscr{G}$ of bounded linear operators, and normalization by a one-parameter semigroup of operators.

This book is designed to be used by graduate students as a textbook either in a classroom or for directed studies. For that purpose, we added Chapter 1. Definitely, we do not suggest one starts with Chapter 1 since it is for reference purposes only and it should be consulted as needed. Because of aim 2, it is not necessary that readers be familiar with classical stability or decomposability (on $\mathbb{R}^1$ or a Banach space). As a by-product, those are discussed at the end of each chapter. We do not have sections with exercises, but many of the corollaries and lemmas can be assigned as homework. Because of aim 3, students can persue their own research in the area of operator-limit distributions on infinite-dimensional linear spaces and on topological groups. Those interested only in operator-stability can skip Chapter 3 since it is quite independent of Chapter 4.

We also see this book as the main reference for further research. It is the first monograph covering such a selection of limit laws. In some sense it can be viewed as a complement to the books by Araujó and Giné (1980), Linde (1986), and Zolotarev (1986).

It took a long time to complete this book due in large part to the "long-distance" communication between Poland and the United States. Thus, we divided the work and the responsibility. Of course, we share the responsibility for any errors and oversights. Much of the results are due to others and we point this out in the bibliographic comments at the end of each chapter.

We benefited from discussions with many of those working in the areas of operator-limit theorems. We thank them all for their encouraging, discouraging, or neutral comments. Those whom we particularly thank include (in alphabetical order): M. G. Hahn (Tufts University), W. Hazod (University of Dortmund), W. N. Hudson (Auburn University), R. Jajte (University of Łódz), M. Klass (University of California, Berkeley), W. Krakowiak, J. Kucharczak, B. Mincer, and T. Rajba (all from University of Wrocław), K. Sato (University of Nagoya), H. Tucker (University of California, Irvine), K. Urbanik (University of Wrocław), J. A. Veeh (Auburn University), and M. Yamazato (Nagoya Institute of Technology).

Last, but not least, we would like to thank our closest loved ones who have suffered through these years. Their constant reminding question: "When will you be ready?" was always motivating and encouraging on a road up to this line.

ZBIGNIEW J. JUREK
J. DAVID MASON

Contents

CHAPTER 1

Preliminaries

The purpose of this introductory chapter is to collect all the needed facts for future references. The theorems and tools from outside of probability theory are presented with complete proofs and are in the generality that covers our needs. However, facts from probability theory, mainly concerned with weak limit theorems, are presented mostly without proofs. References for them are given in the last section of this chapter.

1.1 NUMAKURA THEOREM

Let S be a Hausdorff topological space. If in S there is defined a single-valued product ab which is associative and continuous, then S is called a *topological semigroup*. By a *subsemigroup* of S we mean a nonempty subset A such that $A^2 \subset A$, that is, $ab \in A$ for all $a, b \in A$. Also, A is called a *subgroup* of S if $xA = Ax = A$ for all $x \in A$. By a *left* (*right*) *ideal* of S we mean a subset M such that $SM \subset M$ ($MS \subset M$). When M is both a left and a right ideal of S, M is called an *ideal* of S. Finally, an element $a \in S$ is called an *idempotent* of S provided $a^2 = a$. Obviously, the zero ($a \cdot 0 = 0 \cdot a = 0$ for all $a \in S$) and the identity ($e \cdot a = a \cdot e = a$ for all $a \in S$) elements of S are idempotent, whenever they exist.

Theorem 1.1.1. *If S is a compact semigroup, then S contains a compact subgroup, and hence at least one idempotent.*

Proof. Fix $a \in S$ and let $K(a)$ denote the set of all limit points of the sequence $\{a^n\}_{n \geq 1}$, that is,

$$K(a) := \bigcap_{n=1}^{\infty} \overline{\{a^i : i \geq n\}},$$

where A^- denotes the closure of $A \subset S$. We assert that $K(a)$ is a compact, commutative subsemigroup of S. The commutativity of $K(a)$ is given by: let $b_1 := \lim_{n\to\infty} a^{m_n}$, $b_2 := \lim_{n\to\infty} a^{k_n}$ with $\{m_n\}$ and $\{k_n\}$ strictly increasing. By continuity of products, $b_1 b_2 = \lim_{n\to\infty} a^{m_n+k_n} = b_2 b_1$. The subsemigroup property follows from a similar argument and the compactness is obvious.

To complete the proof, we show that $xK(a) = K(a)$ for all $x \in K(a)$. Clearly, we have $xK(a) \subset K(a)$. Suppose there is an $x \in K(a)$ for which $xK(a)$ is a proper subset of $K(a)$. Then there is a $z \in K(a)$ such that $z \notin xK(a)$. By continuity of products, there are open neighborhoods V, W, and U of x, $K(a)$, and z, respectively, such that $VW \cap U = \varnothing$. Since $x, z \in K(a)$, there are an integer m and a sequence of integers $\{n_i\}_{i\geq 1}$ such that $n_{i+1} > n_i > m$, $a^m \in V$, and $a^{n_i} \in U$ for all $i \geq 1$. Let b be a limit point of $\{a^{n_i-m}\}_{i\geq 1}$. Then $b \in K(a) \subset W$, so there is an integer j such that $a^{n_i-m} \in W$ for all $i \geq j$. Hence, $a^{n_i} = a^m a^{n_i-m} \in VW$ for $i \geq j$. But, $a^{n_i} \in U$, which contradicts $VW \cap U = \varnothing$. Thus, $xK(a) = K(a)x = K(a)$, so $K(a)$ is a subgroup of S. Obviously, the identify of $K(a)$ is an idempotent of S.

Q.E.D.

For $A \subset S$, let sem A denote the smallest closed subsemigroup in S containing A. In case $A = \{a\}$, sem$\{a\}$ is called a *monothetic* semigroup.

Theorem 1.1.2. *If the monothetic semigroup* sem$\{a\}$ *is compact, then the set of all limit points of* $\{a^n\}_{n\geq 1}$, $K(a)$, *is the minimal ideal in* sem$\{a\}$ *and the unit element,* e, *of the subgroup* $K(a)$ *is the only idempotent in* sem$\{a\}$.

Proof. Since sem$\{a\} = \{a^n: n \geq 1\} \cup K(a)$ and $K(a)$ is a commutative subgroup of sem$\{a\}$, we see that $K(a)$ is an ideal in sem$\{a\}$. Let b be an idempotent in sem$\{a\}$. When $b \in K(a)$, $b = e$ since $K(a)$ is a group. When $b \in \{a^n: n \geq 1\}$, then $b = a^m$ for some m, so for every $k \geq 1$, $b = b^k = a^{km}$, which implies $b \in K(a)$. This shows that the identity of $K(a)$ is the only idempotent of sem$\{a\}$.

It remains to show $K(a)$ is a minimal ideal. First, note that if H is a minimal ideal in sem$\{a\}$, then $xH = Hx = H$ for all $x \in$ sem$\{a\}$, because otherwise xH or Hx is a proper subideal of H in sem$\{a\}$. Hence, H is a subgroup with unit element $e \in K(a)$, since sem$\{a\}$ has only one idempotent. Second, note that $H = G(e)$, where H is a minimal ideal in sem$\{a\}$ and $G(e)$ is a maximal subgroup of sem$\{a\}$ containing e. We know that $G(e)$ exists by the Zorn lemma. To see this, we obviously have $H \subset G(e)$. Let $z \in G(e)$. Then $zH = H$, so there is $z^* \in H$ such that $zz^* = z^*z = e$. Hence, $z \in H$, so $G(e) \subset H$. Third, we show that $K(a) = G(e)$. Since $K(a)$ is a subgroup with identity e, we have $K(a) \subset G(e)$. Suppose there is $b \in G(e)$ and $b \notin K(a)$. Since $b \in \text{sem}\{a\}$, $b = a^m$ for some integer m. Since $a^m e = a^m$ and product is continuous, for every neighborhood W of b, there is a neighborhood V of e such that $bV \subset W$. Since $e \in K(a)$, $V \supset \{a^{n_k}: k \geq 1\}$ for some sequence $n_k < n_{k+1}$. Hence, $a^m a^{n_k} = a^{m+n_k} \in W$, so b must be in $K(a)$. This contradiction shows that $K(a) = G(e)$. These three steps show that $K(a)$ is the minimal ideal in $\text{sem}\{a\}$. Q.E.D.

Corollary 1.1.3 (Numakura Theorem). *Let S be a compact semigroup. For each $a \in S$, the monothetic semigroup* $\text{sem}\{a\}$ *is compact and the set of limit points of $\{a^n\}_{n \geq 1}$, $K(a)$, is a subgroup. Moreover, $K(a)$ is the minimal ideal of* $\text{sem}\{a\}$ *[hence, for x in* $\text{sem}\{a\}$, *$xK(a) = K(a)$] and the identity element, e, of the group $K(a)$ is the only idempotent in* $\text{sem}\{a\}$.

For future reference, we have the following corollary.

Corollary 1.1.4. *If the monothetic semigroup* $\text{sem}\{a\}$ *is compact, then there is $b \in K(a)$ such that $ab = ba = e$.*

Proof. Since $K(a)$ is commutative and $aK(a) = K(a)$, such $b \in K(a)$ exists. Q.E.D.

1.2 LINEAR SPACES

Let X be a real *Banach* space, that is, X is a real linear, normed, complete space, with norm $\|\cdot\|$. By X^* we denote its *topological dual Banach* space, that is, $x^* \in X^*$ are continuous linear functionals on X, and $\langle \cdot, \cdot \rangle$ is the dual pair between X^* and X. When the norm in X is given by a scalar product, X is called a *Hilbert* space. In that case, X^* is isomorphic to X and the dual pair is simply the scalar product. Furthermore, all real separable Hilbert spaces are isomorphic to l_2, the space of all real square-summable sequences with

$$\langle x, y \rangle := \sum_i x_i y_i, \qquad \|x\| := (\langle x, x \rangle)^{1/2}$$

for $x = \{x_i\}_{i \geq 1}$ and $y = \{y_i\}_{i \geq 1}$. Besides this example, we will deal with the following ones.

(a) $X = C_0(\mathbb{R}^d)$ is the set of all real-valued continuous functions on $\mathbb{R}^d$, d-dimensional Euclidean space, which vanish at infinity. The point ∞ can be used in the one-point compactification of $\mathbb{R}^d$. By the *Riesz representation theorem*, each $x^* \in [C_0(\mathbb{R}^d)]^*$ is uniquely determined by a finite Borel measure m on $\mathbb{R}^d$, not necessarily positive, such that, for $f \in C_0(\mathbb{R}^d)$,

$$\langle x^*, f \rangle = \int f(x) m(dx), \qquad \|x^*\| = m(\mathbb{R}^d).$$

(b) $X = \mathbb{R}^d$. Then $(\mathbb{R}^d)^* = \mathbb{R}^d$ and

$$\langle x, y \rangle = \sum_{i=1}^{d} x_i y_i.$$

Also, it is easy to see that in all finite-dimensional inner product spaces all norms are *equivalent*, that is, if $\|\cdot\|_1$ and $\|\cdot\|_2$ are two norms on X, then there are positive constants c_1 and c_2 such that for all $x \in X$, $c_1\|x\|_1 \leq \|x\|_2 \leq c_2\|x\|_1$.

Let X and Y be Banach spaces. By a *bounded linear operator A* from X into Y, we mean a function A: $X \to Y$ such that (1) A is linear, that is, $A(\alpha_1 x_1 + \alpha_2 x_2) = \alpha_1 A x_1 + \alpha_2 A x_2$ for all $x_1, x_2 \in X$ and all $\alpha_1, \alpha_2 \in \mathbb{R}^1$, and (2) there is a constant C such that $\|Ax\| \leq C\|x\|$ for all $x \in X$; the first norm is in Y and the second norm, possibly different, is in X. The infimum of all C in (2) is denoted by $\|A\|$, and is called the *norm of the operator A*. The assumption that A is bounded and linear is equivalent to A being continuous and linear from X to Y, where the topologies are given by the norms. The collection $L(X, Y)$ of all bounded linear operators from X into Y, using the operator norm, is also a Banach space. When $X = Y$, $L(X, Y)$ is denoted by $\mathrm{End}(X)$; in which case, we also have that the product of two operators in $\mathrm{End}(X)$ is a continuous linear operator: if $A, B \in \mathrm{End}(X)$, then AB: $X \to X$ is given by $(AB)x = A(Bx)$ for $x \in X$. Moreover, $\|AB\| \leq \|A\|\,\|B\|$ for all $A, B \in \mathrm{End}(X)$. With this multiplication of operators, $\mathrm{End}(X)$ becomes a topological semigroup. By $\mathrm{Aut}(X)$, we denote the set of all invertible operators in $\mathrm{End}(X)$. These inverses are also continuous and linear, so $\mathrm{Aut}(X)$ is a topological group.

Let $\mathscr{D}(A)$ be a linear subspace of a Banach space X, and let A be a linear operator from $\mathscr{D}(A)$ into a Banach space Y. By the *graph* of A is meant the set

$$\text{graph } A := \{(x, Ax): x \in \mathscr{D}(A)\} \subset X \times Y.$$

The product space $X \times Y$ can be treated as a Banach space; for example, $\|(x, y)\| := \|x\|_1 + \|y\|_2$, where $\|\cdot\|_1$ and $\|\cdot\|_2$ are the norms in X and Y, respectively. We say that the operator A is *closed* if its graph is a closed subset of $X \times Y$, that is, if $x_n \in \mathscr{D}(A)$, $x_n \to x_0$ in X, $Ax_n \to y_0$ in Y implies that $x_0 \in \mathscr{D}(A)$ and $y_0 = Ax_0$. An operator B defined on $\mathscr{D}(B) \subset X$ is called an *extension* of A with its domain $\mathscr{D}(A)$ whenever $\mathscr{D}(A) \subset \mathscr{D}(B)$ and $Ax = Bx$ for all $x \in \mathscr{D}(A)$.

The following is a criterion which determines when an operator A with domain $\mathscr{D}(A)$ has a closed extension.

Theorem 1.2.1. *Let A be a linear operator with domain $\mathscr{D}(A) \subset X$. Then A has a closed extension if and only if there is no $y \neq 0$ such that $(0, y)$ belongs to* (graph $A)^-$, *the closure of* graph A. *In this case,* (graph $A)^-$ *is the* graph A^-, *the smallest closed extension of A.*

Proof. Since graph A is a linear subspace of $X \times Y$, so is (graph $A)^-$. Since $(x, y_1), (x, y_2) \in$ (graph $A)^-$ implies that $y_1 = y_2$ $[(x, y_1) - (x, y_2) = (0, y_1 - y_2) \in$ (graph $A)^-]$, the relation (graph $A)^- \subset X \times Y$ determines a function, A^-. We see that the linearity and closeness of (graph $A)^-$ implies that A^- is linear and closed. Obviously, graph $A^- =$ (graph $A)^-$. It is also obvious that (graph $A)^-$ is the graph of the smallest closed extension of A. The converse is trivial. Q.E.D.

We use A^- to denote the smallest closed extension of the linear operator A.

Corollary 1.2.2. *If a linear operator A has an extension which is a closed linear operator B, then A^- exists.*

Proof. Since (graph $A)^- \subset$ graph B, there is no $y \neq 0$ such that $(0, y) \in$ (graph $A)^-$. Now, apply Theorem 1.2.1. Q.E.D.

Obviously, all bounded linear operators are closed. Other examples of closed operators are the infinitesimal generators of one-parameter semigroups [cf. Proposition 4.4.1(c)]. Also, note that the sum of a closed linear operator and a bounded linear operator is closed.

Now, we wish to define the trace of a linear operator on $\mathbb{R}^d$. When A is a $d \times d$ matrix, we define trace(A) to be the sum of the numbers on its main diagonal. Since similar matrices have the same trace, we may define the trace(A) for A a linear operator on $\mathbb{R}^d$ to be the trace of any matrix which represents A in an ordered basis.

For $A \in \mathrm{End}(\mathbb{R}^d)$, we may define e^A by the series $e^A := \sum_{n=0}^{\infty}(n!)^{-1}A^n$. Since $\sum_{n=0}^{\infty}(n!)^{-1}\|A\|^n$ converges, $e^A \in \mathrm{End}(\mathbb{R}^d)$. Actually, $e^A \in \mathrm{Aut}(\mathbb{R}^d)$ since the inverse of e^A is given by e^{-A}.

It is well known how the determinant of a $d \times d$ matrix A is defined. For $A \in \mathrm{End}(\mathbb{R}^d)$, $\det(A)$ denotes the determinant of any matrix which represents A in a basis. Then $\det(e^A) = e^{\mathrm{trace}(A)}$.

The final topic of this section is the primary decomposition theorem of $\mathbb{R}^d$. Let $A \in \mathrm{End}(\mathbb{R}^d)$. The *minimal polynomial for A* is the unique polynomial $g(\cdot)$ over the reals with the properties that (1) the coefficient of its highest term is one, (2) $g(A) = 0$, and (3) if $h(A) = 0$, where $h(\cdot)$ is a polynomial over the reals, then the degree of $h(\cdot)$ is greater than or equal to the degree of $g(\cdot)$. A polynomial over the reals is said to be *irreducible* if it is not possible to factor it into the product of two polynomials, each having degree greater than or equal to one. Let $\ker(A)$ denote the set $\{x \in \mathbb{R}^d\colon Ax = 0\}$, and call it the *null space of A*. A subspace $W \subset \mathbb{R}^d$ is said to be *A-invariant* if $A(W) \subset W$. We write $\mathbb{R}^d = W_1 \oplus \cdots \oplus W_k$ if each W_i is a subspace of $\mathbb{R}^d$ and each $v \in \mathbb{R}^d$ has a unique representation of the form $v = v_1 + \cdots + v_k$ with $v_i \in W_i$ for all i. We also say that $\mathbb{R}^d$ is the *direct sum* of the subspaces $W_1, \ldots, W_k$. A nonzero polynomial $a_n x^n + a_{n-1}x^{n-1} + \cdots$ is called a *monic* polynomial if $a_n = 1$.

Theorem 1.2.3 (Primary Decomposition Theorem). *Let $A \in \mathrm{End}(\mathbb{R}^d)$, and let $g(\cdot)$ be its minimal polynomial. Assume $g = g_1^{r_1} \cdots g_k^{r_k}$, where the g_i are distinct irreducible monic polynomials over $\mathbb{R}^1$ and the r_i are positive integers. Let $W_i := \ker(g_i(A)^{r_i})$ for $1 \le i \le k$. Then*

- **(i)** $\mathbb{R}^d = W_1 \oplus \cdots \oplus W_k$;
- **(ii)** *each W_i is A-invariant;*
- **(iii)** *if A_i is the restriction of A to W_i, then the minimal polynomial of A_i is $g_i^{r_i}$.*