Springer INdAM Series 30

Dario Andrea Bini
Fabio Di Benedetto
Eugene Tyrtyshnikov
Marc Van Barel *Editors*

Structured Matrices in Numerical Linear Algebra

Analysis, Algorithms and Applications

Springer

Springer INdAM Series

Volume 30

More information about this series at http://www.springer.com/series/10283

Dario Andrea Bini • Fabio Di Benedetto •
Eugene Tyrtyshnikov • Marc Van Barel
Editors

Structured Matrices in Numerical Linear Algebra

Analysis, Algorithms and Applications

Springer

Editors
Dario Andrea Bini
Department of Mathematics
University of Pisa
Pisa, Italy

Fabio Di Benedetto
Department of Mathematics
University of Genoa
Genoa, Italy

Eugene Tyrtyshnikov
Institute of Numerical Mathematics
Russian Academy of Sciences
Moscow, Russia

Marc Van Barel
Department of Computer Science
KU Leuven
Heverlee, Belgium

ISSN 2281-518X ISSN 2281-5198 (electronic)
Springer INdAM Series
ISBN 978-3-030-04087-1 ISBN 978-3-030-04088-8 (eBook)
https://doi.org/10.1007/978-3-030-04088-8

Library of Congress Control Number: 2019936552

This Springer imprint is published by the registered company Springer Nature Switzerland AG.
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

Solving a mathematical model, by means of computational techniques in general or more specifically by means of numerical analysis, is ultimately reduced to solving problems in linear algebra. In fact, very often models are linear by their nature, while sometimes they are nonlinear but their solution is achieved by means of linearization. This to some extent explains why numerical linear algebra and matrix analysis have undergone such extensive development in recent decades.

In this process, one typically encounters problems like solving linear systems, or solving standard or generalized eigenvalue problems, as well as matrix polynomial equations or polynomial eigenvalue problems where the size of the matrices involved in the model is extremely large or in some cases even infinite. For such problems standard (general-purpose) algorithms cannot work due to their extreme complexity; therefore, one has to exploit the specific properties which originate from the peculiar features of the model. In the language of linear algebra, these properties are translated in terms of *structures* that the matrices involved in the model share; often, structured matrices reveal themselves in a clear form and appear to show all their properties immediately. Sometimes, however, structures are hidden and difficult to discover, and their properties seem hardly exploitable. Their analysis and exploitation is not just a challenge but also a mandatory step which is necessary to design highly effective ad hoc algorithms for the solution of large-scale problems from applications. In fact, general-purpose algorithms, say Gaussian elimination for solving linear systems, cannot be used to solve problems of large size while a smart exploitation of the available structures enables one to design effective solution algorithms even for problems of a huge size.

The importance of matrix structures has grown over the years. Analyzing structures from the theoretical point of view, turning them into effective solution algorithms, constructing software which implements the algorithms, and verifying its effectiveness by direct computation is one of the most exciting challenges that covers abstract theory, design and analysis of algorithms, software implementation, and applications.

This volume presents a selected number of peer-reviewed papers concerning structured matrix analysis and its applications. The topics discussed concern theory,

algorithms, and applications in which structured matrices are involved. The subjects range from abstract topics such as the theory of generalized locally (block) Toeplitz matrices and the analysis of matrix subspaces and quadratic kernels to more numerical issues such as error analysis of algorithms for tensor manipulation and analysis of the derivative of matrix geometric means. Moreover, other structured oriented topics are developed, e.g., analysis of companion pencil and block Fiedler companion matrices, together with analysis of the tridiagonal symmetric eigenvalue problem, computation of bivariate matrix functions, and solution of the saddle point problem. Among the applications are analysis of the stability of gyroscopic systems, numerical solution of 2D hard scattering problems of damped waves, fractional reaction-diffusion equations, and the problem of multi-frame super-resolution reconstruction from video clips.

All the papers correspond to talks presented at the INdAM meeting *Structured Matrices in Numerical Linear Algebra: Analysis, Algorithms and Applications* held in Cortona, Italy, on September 4–8, 2017.

This workshop aimed to continue in both form and spirit the series of conferences on *Structured Matrices* and their applications held in Cortona, Italy, every 4 years between 1996 and 2008 and continued in Leuven, Belgium, in September 2012 and in Kalamata, Greece, in September 2014.

The book will be of interest to graduate students in mathematics and researchers in numerical linear algebra and scientific computing, as well as engineers and applied mathematicians.

Pisa, Italy — Dario Andrea Bini
Genoa, Italy — Fabio Di Benedetto
Moscow, Russia — Eugene Tyrtyshnikov
Heverlee, Belgium — Marc Van Barel
October 2018

Contents

About the Editors

Dario Andrea Bini, a Full Professor of Numerical Analysis since 1986, has held a permanent position at the University of Pisa since 1989. His research mainly focuses on numerical linear algebra problems, on structured matrix analysis and on the design and analysis of algorithms for polynomial and matrix computations. The author of three research books and more than 120 papers, he also serves on the editorial boards of three international journals.

Fabio Di Benedetto is an Associate Professor of Numerical Analysis at the Department of Mathematics of the University of Genoa, where he teaches courses on Numerical Analysis for undergraduate and graduate students, since 2000. His main research interests concern the solution of large-scale numerical linear algebra problems, with special attention to structured matrices analysis with applications to image processing. He is the author of more than 30 papers.

Eugene Tyrtyshnikov, Professor and Chairman at the Lomonosov Moscow State University, is a Full Member of the Russian Academy of Sciences and Director of the Institute of Numerical Mathematics of the Russian Academy of Sciences, Moscow. He completed his Ph.D. in Numerical Mathematics at Moscow State University, and his postdoctoral studies at the Siberian Branch of the Russian Academy of Sciences, Novosibirsk. His research interests concern numerical analysis, linear and multilinear algebra, approximation theory and related applications. He is the associate editor of many international journals and the author of more than 100 papers and 8 books.

Marc Van Barel received his Ph.D. in Computer Engineering (Numerical Analysis and Applied Mathematics) from the KU Leuven, where he is currently a Full Professor at the Department of Computer Science. His work mainly focuses on numerical (multi-)linear algebra, approximation theory, orthogonal functions and their applications in systems theory, signal processing, machine learning, etc. He is the author or co-author of more than 140 papers and 4 books. Currently, he serves on the editorial boards of three international journals.

Spectral Measures

Giovanni Barbarino

Abstract The theory of spectral symbols links sequences of matrices with measurable functions expressing their asymptotic eigenvalue distributions. Usually, a sequence admits several spectral symbols, and it is not clear if a canonical one exists. Here we present a way to connect the sequences with the space of probability measure, so that each sequence admits a uniquely determined measure. The methods used are similar to those employed in the theory of generalized locally Toeplitz (GLT) sequences: a goal of this present contribution is in fact that of explaining how the two concepts are connected.

Keywords Probability measures · Generalized locally Toeplitz sequences · Complete pseudo-metrics · Ergodic formula

1 Introduction

A *matrix sequence* is an ordered collection of complex valued matrices with increasing size, and is usually denoted as $\{A_n\}_n$, where $A_n \in \mathbb{C}^{n\times n}$. We will refer to the space of matrix sequences with the notation

$$\mathscr{E} := \{\{A_n\}_n : A_n \in \mathbb{C}^{n\times n}\}.$$

It is often observed in practice that matrix sequences, $\{A_n\}_n$, generated by discretization methods applied to linear differential equations possess a *spectral symbol*, that is a measurable function describing the asymptotic distribution of the

G. Barbarino (✉)
Scuola Normale Superiore, Pisa, Italy
e-mail: giovanni.barbarino@sns.it

D. A. Bini et al. (eds.), *Structured Matrices in Numerical Linear Algebra*,
Springer INdAM Series 30, https://doi.org/10.1007/978-3-030-04088-8_1

eigenvalues of A_n. We recall that a spectral symbol associated with a sequence $\{A_n\}_n$ is a measurable function $k : D \subseteq \mathbb{R}^n \to \mathbb{C}$ satisfying

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} F(\lambda_i(A_n)) = \frac{1}{l(D)} \int_D F(k(x))dx$$

for every continuous function $F : \mathbb{C} \to \mathbb{C}$ with compact support, where D is a measurable set with finite Lebesgue measure $l(D) > 0$ and $\lambda_i(A_n)$ are the eigenvalues of A_n. In this case we write

$$\{A_n\}_n \sim_\lambda k(x).$$

We can also consider the singular values of the matrices instead of the eigenvalues. In the same setting, if

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} F(\sigma_i(A_n)) = \frac{1}{l(D)} \int_D F(|k(x)|)dx$$

for every continuous function $F : \mathbb{R} \to \mathbb{C}$ with compact support, where $\sigma_i(A_n)$ are the singular values of A_n, then $\{A_n\}_n$ possesses a *singular value symbol*, and we write

$$\{A_n\}_n \sim_\sigma k(x).$$

The space of matrix sequences is a complete pseudometric space when endowed with a pseudometric inducing the *approximating classes of sequences* (acs) convergence, that we will redefine in the next section. One fundamental property of this metric is that it identifies sequences that differ by a sequence admitting zero as singular value symbol (called zero-distributed sequences). In particular, it has been shown that such sequences share the same singular value symbol, but the distance between two sequences with the same singular value symbol is not usually zero.

The main observation of this note is that for any measurable function $k(x)$, the operator

$$\phi(F) := \int_D F(k(x))dx \qquad \phi : C_c(\mathbb{C}) \to \mathbb{C}$$

is linear and continuous and can be represented by a unique probability measure μ. We call μ a *spectral measure*, and we associate it with any sequence $\{A_n\}_n$ that has $k(x)$ as spectral symbol. It turns out that if a sequence admits a spectral measure, then it is uniquely determined, differently from the spectral symbols. The space of probability spectral measures is moreover a complete metric space with the Lévy–Prokhorov distance π, and it corresponds to a pseudometric d' on matrices called

modified optimal matching distance. The main result is that d' identifies sequences admitting the same spectral symbol, differently from the acs distance.

Theorem 1 *If* $\{A_n\}_n \sim_\lambda f(x)$, *then*

$$\{B_n\}_n \sim_\lambda f(x) \iff d'(\{A_n\}_n, \{B_n\}_n) = 0.$$

A different approach to the uniqueness problem for the spectral symbol is embodied in the theory of GLT sequences. For specific sequences, called *generalized locally Toeplitz* (GLT) sequences, we can choose one of their symbols, and denote it as *GLT symbol* of the sequence

$$\{A_n\}_n \sim_{GLT} k(x, \theta).$$

In the case of diagonal matrix sequences, the choice of one symbol can be seen as a particular sorting of their eigenvalues, as expressed in the following theorem, proved in the last section, and which represents a generalization of the results in [3].

Theorem 2 *Given a diagonal sequence* $\{D_n\}_n$ *and one of its spectral symbols* $k : [0, 1] \to \mathbb{C}$, *then*

$$\{P_n D_n P_n^T\} \sim_{GLT} k(x) \otimes 1$$

for some P_n *permutation matrices.*

The paper is organized in the following way: In Sect. 2 we recall basic definitions such as the acs convergence, the optimal matching distance d, and the theory of GLT sequences. Moreover, we define the modified optimal matching distance d' since it is a slight variation of d, and we discuss how it is connected to d_{acs}. In Sect. 3 we introduce the spectral measures and we study their relationships with the spectral symbols. In particular, we notice how the vague convergence and the Lévy–Prokhorov distance π on the probability measures lead to a reformulation of the definition of spectral symbol/measure. In Sect. 4, we prove that the pseudometrics π and d' are actually equivalent, and we explain how this fact leads to the proofs of the above reported theorems.

2 Prerequisites

2.1 Complete Pseudometrics

The space of matrix sequences that admit a spectral symbol on a fixed domain D has been shown to be closed with respect to a notion of convergence called the approximating classes of sequences (acs) convergence. This notion and this result are due to Serra-Capizzano [11], but were actually inspired by Tilli's pioneering

paper on LT sequences [12]. Given a sequence of matrix sequences $\{B_{n,m}\}_{n,m}$, it is said to be acs convergent to $\{A_n\}_n$ if there exists a sequence $\{N_{n,m}\}_{n,m}$ of "small norm" matrices and a sequence $\{R_{n,m}\}_{n,m}$ of "small rank" matrices such that for every m there exists n_m with

$$A_n = B_{n,m} + N_{n,m} + R_{n,m}, \qquad \|N_{n,m}\| \leq \omega(m), \qquad \mathrm{rk}(R_{n,m}) \leq nc(m)$$

for every $n > n_m$, and

$$\omega(m) \xrightarrow{m\to\infty} 0, \qquad c(m) \xrightarrow{m\to\infty} 0.$$

In this case, we will use the notation $\{B_{n,m}\}_{n,m} \xrightarrow{acs} \{A_n\}_n$.

This notion of convergence has been shown to be metrizable on the whole space $\mathscr{E}$. Given a matrix $A \in \mathbb{C}^{n\times n}$, we can define the function

$$p(A) := \min_{i=1,\dots,n+1} \left\{ \frac{i-1}{n} + \sigma_i(A) \right\},$$

where $\sigma_1(A) \geq \sigma_2(A) \geq \cdots \geq \sigma_n(A)$ are the singular values of A, and by convention $\sigma_{n+1}(A) = 0$. The function $p(A)$ is subadditive, so we can introduce the pseudometric d_{acs} on the space of matrix sequences

$$d_{acs}(\{A_n\}_n, \{B_n\}_n) = \limsup_{n\to\infty} p(A_n - B_n).$$

It has been proved [6, 8] that this distance induces the acs convergence already introduced. In other words,

$$d_{acs}\left(\{A_n\}_n, \{B_{n,m}\}_{n,m}\right) \xrightarrow{m\to\infty} 0 \iff \{B_{n,m}\}_{n,m} \xrightarrow{acs} \{A_n\}_n.$$

One fundamental property of this metric is that it identifies sequences whose difference admits zero as singular value symbol (called zero-distributed sequence). In particular, it has been shown that such sequences share the same singular value symbol, in case one of them admits singular value symbol.

Lemma 1 *Let $\{A_n\}_n, \{B_n\}_n \in \mathscr{E}$. We have*

$$\{A_n - B_n\}_n \sim_\sigma 0 \iff d_{acs}(\{A_n\}_n, \{B_n\}_n) = 0.$$

In this case, if $k : D \subseteq \mathbb{R}^n \to \mathbb{C}$ where D is a measurable set with finite Lebesgue measure $l(D) > 0$, then

$$\{A_n\}_n \sim_\sigma k(x) \iff \{B_n\}_n \sim_\sigma k(x).$$

In [2], it has been first proved that the pseudometric d_{acs} on the space of matrix sequences is complete. In Theorem 2.2 of [4], we find sufficient conditions for a pseudometric on $\mathscr{E}$ to be complete. Here we need a different result, but the proof is almost identical.

Lemma 2 *Let d_n be pseudometrics on the space of matrices $\mathbb{C}^{n\times n}$ bounded by the same constant $L > 0$ for every n. Then the function*

$$d(\{A_n\}_n, \{B_n\}_n) := \limsup_{n\to\infty} d_n(A_n, B_n)$$

is a complete pseudometric on the space of matrix sequences.

2.2 Optimal Matching Distance

Let $v, w \in \mathbb{C}^n$ be vectors with components

$$v = [v_1, v_2, \ldots, v_n], \quad w = [w_1, w_2, \ldots, w_n].$$

We recall the pseudometric on $\mathbb{C}^n$ called *optimal matching distance* defined in Bhatia's book [5].

Definition 1 Given $v, w \in \mathbb{C}^n$, the pseudometric of the optimal matching distance is defined as

$$d(v, w) := \min_{\sigma\in S_n} \max_{i=1,\ldots,n} |v_i - w_{\sigma(i)}|,$$

where S_n is the symmetric group of permutation of n objects.

Given $A \in \mathbb{C}^{n\times n}$, let $\Lambda(A) \in \mathbb{C}^n$ be the vector of the eigenvalues. We can extend the distance d to matrices in the following way.

Definition 2 Given $A, B \in \mathbb{C}^{n\times n}$, we define

$$d(A, B) := d(\Lambda(A), \Lambda(B)).$$

Notice that the order of the eigenvalues in $\Lambda(A)$ and $\Lambda(B)$ does not affect the quantity $d(A, B)$. It is easy to see that d is still a pseudometric on $\mathbb{C}^{n\times n}$. This is still not enough for our purposes, since we want a distance that sees two matrices differing for few eigenvalues as very similar. For this reason, we modify the previous metric, and we introduce a new function d' called *modified optimal matching distance*.

Definition 3 Given $v, w \in \mathbb{C}^n$, the modified optimal matching distance is defined as

$$d'(v, w) := \min_{\sigma \in S_n} \min_{i=1,\dots,n+1} \left\{ \frac{i-1}{n} + |v - w_\sigma|_i^{\downarrow} \right\},$$

where

$$|v - w_\sigma| = [|v_1 - w_{\sigma(1)}|, |v_2 - w_{\sigma(2)}|, \dots, |v_n - w_{\sigma(n)}|]$$

and $|v - w_\sigma|_i^{\downarrow}$ is the i-th greatest element in $|v - w_\sigma|$, with the convention $|v - w_\sigma|_{n+1}^{\downarrow} := 0$.

Given $A, B \in \mathbb{C}^{n \times n}$, we define

$$d'(A, B) := d'(\Lambda(A), \Lambda(B))$$

and if $\{A_n\}_n, \{B_n\}_n \in \mathscr{E}$, we can also define

$$d'(\{A_n\}_n, \{B_n\}_n) := \limsup_{n \to \infty} d'(A_n, B_n).$$

Notice that $d'(v, w) \leq 1$ for every $v, w \in \mathbb{C}^n$, so $d'(A, B) \leq 1$ for every pair of matrices of the same size, and $d'(\{A_n\}_n, \{B_n\}_n) \leq 1$ for every pair of sequences $\{A_n\}_n, \{B_n\}_n \in \mathscr{E}$. We referred to d' as a distance, but we need to prove it.

Lemma 3 *The function d' is a complete pseudometric on $\mathscr{E}$.*

Proof Let us prove that d' is a pseudometric on $\mathbb{C}^n$. First, it is easy to see that $d'(v, w)$ is always a finite nonnegative real number, and it is symmetric since

$$\begin{aligned} d'(v, w) &= \min_{\sigma \in S_n} \min_{i=1,\dots,n+1} \left\{ \frac{i-1}{n} + |v - w_\sigma|_i^{\downarrow} \right\} \\ &= \min_{\sigma \in S_n} \min_{i=1,\dots,n+1} \left\{ \frac{i-1}{n} + |w - v_{\sigma^{-1}}|_i^{\downarrow} \right\} = d'(w, v). \end{aligned}$$

Moreover, given any $\tau \in S_n$, we have

$$\begin{aligned} d'(v, w) &= \min_{\sigma \in S_n} \min_{i=1,\dots,n+1} \left\{ \frac{i-1}{n} + |v - w_\sigma|_i^{\downarrow} \right\} \\ &= \min_{\sigma \in S_n} \min_{i=1,\dots,n+1} \left\{ \frac{i-1}{n} + |v_\tau - w_{\sigma\tau}|_i^{\downarrow} \right\} = d'(v_\tau, w), \end{aligned}$$

so we can permute the elements of the vectors as we like. Let $v, w, z \in \mathbb{C}^n$ and let us sort their elements in such a way that

$$d'(v,w) = \min_{i=1,\dots,n+1} \left\{ \frac{i-1}{n} + |v_i - w_i| \right\},$$
$$d'(w,z) = \min_{i=1,\dots,n+1} \left\{ \frac{i-1}{n} + |w - z|_i^{\downarrow} \right\},$$

meaning that the permutation realizing the minimum in both cases is the identity, and that $|v_i - w_i| \geq |v_j - w_j|$ whenever $i \leq j$. Moreover, let s, r, q be the greatest indices that satisfy

$$d'(v,w) = \frac{s-1}{n} + |v_s - w_s|, \qquad d'(w,z) = \frac{r-1}{n} + |w_q - z_q|.$$

Let I, J be two sets of indices defined as

$$I = \{1, 2, \dots, s-1\}, \qquad J = \{j : |w_j - z_j| > |w_q - z_q|\}.$$

Notice that $\#I = s-1$ and $\#J = r-1$. Let us consider two cases.

- Suppose $I \cup J = \{1, \dots, n\}$. We obtain that

  $$\#I + \#J = r + s - 2 \geq n$$

  and hence

  $$d'(v,z) \leq 1 \leq \frac{s-1}{n} + \frac{r-1}{n} \leq d'(v,w) + d'(w,z).$$

- Suppose $I \cup J \neq \{1, \dots, n\}$. Let k be the index not belonging to $I \cup J$ that maximizes $|v_i - z_i|$. If we consider the identity permutation, we deduce that

  $$d'(v,z) \leq \min_{i=1,\dots,n+1} \left\{ \frac{i-1}{n} + |v - z|_i^{\downarrow} \right\},$$

  but the number of indices such that $|v_i - z_i|$ is greater than $|v_k - z_k|$ is at most $\#I \cup J \leq r + s - 2$, and consequently

  $$d'(v,z) \leq \frac{r+s-2}{n} + |v_k - z_k|.$$

  The index k does not belong to I or to J, so

  $$|v_k - w_k| \leq |v_s - w_s|, \qquad |w_k - z_k| \leq |w_q - z_q|.$$

From the latter we infer that

$$\begin{aligned} d'(v,z) &\leq \frac{r+s-2}{n} + |v_k - z_k| \\ &\leq \frac{s-1}{n} + |v_k - w_k| + \frac{r-1}{n} + |w_k - z_k| \\ &\leq \frac{s-1}{n} + |v_s - w_s| + \frac{r-1}{n} + |w_q - z_q| \\ &= d'(v,w) + d'(w,z). \end{aligned}$$

This shows that d' is a pseudometric on $\mathbb{C}^n$ and consequently it is a pseudometric even on $\mathbb{C}^{n\times n}$. Thanks to Lemma 2, we can conclude that d' is a complete pseudometric on $\mathscr{E}$. □

In the general case, the two pseudometrics have no common features, but, when dealing with diagonal matrices, we can prove the following lemma.

Lemma 4 *Given $\{D_n\}_n, \{D'_n\}_n \in \mathscr{E}$ sequences of diagonal matrices, there exists a sequence $\{P_n\}_n$ of permutation matrices such that*

$$d'(\{D'_n\}_n, \{D_n\}_n) = d_{acs}(\{D'_n\}_n, \{P_n D_n P_n^T\}_n).$$

Proof Let v^n and v'^n be the vectors of the ordered diagonal entries of D_n and D'_n, so that

$$v^n_i := [D_n]_{i,i}, \qquad v'^n_i := [D'_n]_{i,i}.$$

Let $\tau_n \in S_n$ be the permutations satisfying

$$\begin{aligned} d'(D'_n, D_n) &= \min_{\sigma\in S_n} \min_{i=1,\dots,n+1} \left\{ \frac{i-1}{n} + |v'^n - v^n_\sigma|^\downarrow_i \right\} \\ &= \min_{i=1,\dots,n+1} \left\{ \frac{i-1}{n} + |v'^n - v^n_{\tau_n}|^\downarrow_i \right\}. \end{aligned}$$

Let also P_n be the permutation matrices associated with τ_n. We know that

$$\begin{aligned} p(D'_n - P_n D_n P_n^T) &= \min_{i=1,\dots,n+1} \left\{ \frac{i-1}{n} + \sigma_i(D'_n - P_n D_n P_n^T) \right\} \\ &= \min_{i=1,\dots,n+1} \left\{ \frac{i-1}{n} + |v'^n - v^n_{\tau_n}|^\downarrow_i \right\} \\ &= d'(D'_n, D_n). \end{aligned}$$

As a consequence

$$d_{acs}(\{D'_n\}_n, \{P_n D_n P_n^T\}_n) = \limsup_{n\to\infty} p(D'_n - P_n D_n P_n^T)$$
$$= \limsup_{n\to\infty} d'(D'_n, D_n) = d'(\{D'_n\}_n, \{D_n\}_n).$$

□

2.3 GLT Matrix Sequences

A matrix sequence $\{A_n\}_n$ may have several different singular value symbols, even on the same domain. For specific sequences, called *generalized locally Toeplitz* (GLT) sequences, we can choose one of their symbols, and denote it as *GLT symbol* of the sequence

$$\{A_n\}_n \sim_{GLT} k(x, \theta).$$

where the chosen symbols have all the same domain $D = [0, 1] \times [-\pi, \pi]$. If we denote with $\mathscr{M}_D$ the set of measurable functions on D, and with $\mathscr{G}$ the set of GLT sequences, then the choice of the symbol can be seen as a map

$$S : \mathscr{G} \to \mathscr{M}_D.$$

Both $\mathscr{G}$ and $\mathscr{M}_D$ are $\mathbb{C}$ algebras and pseudometric spaces with the distances d_{acs} and d_m, inducing respectively the acs convergence and the convergence in measure. In [9] and in [2] several properties of the map S are proved.

Theorem 3

1. *S is a homomorphism of $\mathbb{C}$ algebras. Given $\{A_n\}_n, \{B_n\}_n \in \mathscr{G}$, and $c \in \mathbb{C}$, we have that*

$$S(\{A_n + B_n\}_n) = S(\{A_n\}_n) + S(\{B_n\}_n),$$

$$S(\{A_n B_n\}_n) = S(\{A_n\}_n) \cdot S(\{B_n\}_n),$$

$$S(\{cA_n\}_n) = cS(\{A_n\}_n).$$

2. *The kernel of S are exactly the zero-distributed sequences.*
3. *S preserves the distances. Given $\{A_n\}_n, \{B_n\}_n \in \mathscr{G}$ we have*

$$d_{acs}\left(\{A_n\}_n, \{B_n\}_n\right) = d_m(S(\{A_n\}_n), S(\{B_n\}_n)).$$

4. *S is onto. All measurable functions are GLT symbols.*

5. *GLT symbols are singular value symbols:*

$$\{A_n\}_n \in \mathscr{G} \implies \{A_n\}_n \sim_\sigma S(\{A_n\}_n)$$

6. *The graph of S is closed in $\mathscr{G} \times \mathscr{M}_D$. If $\{B_{n,m}\}_{n,m}$ are sequences in $\mathscr{G}$ that converge acs to $\{A_n\}_n$, and their symbols converge in measure to $k(x,\theta)$, then $S(\{A_n\}_n) = k(x,\theta)$.*

The diagonal sampling sequences are denoted as $\{D_n(a)\}_n$, where $a : [0,1] \to \mathbb{C}$ is a measurable function, and

$$D_n(a) = \underset{i=1,\dots,n}{\operatorname{diag}} a\left(\frac{i}{n}\right) = \begin{pmatrix} a\left(\frac{1}{n}\right) & & & \\ & a\left(\frac{2}{n}\right) & & \\ & & \ddots & \\ & & & a(1) \end{pmatrix}$$

It is easy to verify that when $a : [0,1] \to \mathbb{C}$ is an almost everywhere (a.e.) continuous function, we have $\{D_n(a)\}_n \sim_{\sigma,\lambda} a(x)$. Furthermore, if $a(x)$ is continuous, we know that these sequences have as GLT symbol

$$\{D_n(a)\}_n \sim_{GLT} a(x) \otimes 1,$$

where $a \otimes 1 : [0,1] \times [-\pi,\pi] \to \mathbb{C}$ is a function constant in the second variable. This is not true for every $a(x)$ measurable, so we resort to the following result.

Lemma 5 *Given any $a : [0,1] \to \mathbb{C}$ measurable function, and $a_m \in C([0,1])$ continuous functions that converge in measure to $a(x)$, there exists an increasing and unbounded map $m(n)$ such that*

$$\{D_n(a_{m(n)})\}_n \sim_{GLT} a(x) \otimes 1 \qquad \{D_n(a_{m(n)})\}_n \sim_\lambda a(x)$$

Proof Easy corollary of Lemma 3.4 and Theorem 3.1 in [3]. □

3 Spectral Measures

3.1 Radon Measures

Let $\{A_n\}_n \in \mathscr{E}$ be a sequence with a spectral symbol $k(x)$ with domain D. By definition, we have

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} G(\lambda_i(A_n)) = \frac{1}{l(D)} \int_D G(k(x))dx.$$

Let $\phi : C_c(\mathbb{C}) \to \mathbb{C}$ be the functional defined as

$$\phi(G) := \frac{1}{l(D)} \int_D G(k(x))dx.$$

The latter is a continuous and linear map, and if we restrict it to real valued compacted supported functions, it is also a positive operator, since

$$G(x) \geq 0 \quad \forall x \in \mathbb{C} \implies \phi(G) = \frac{1}{l(D)} \int_D G(k(x))dx \geq 0.$$

Let us now recall Riesz theorem [1].

Theorem 4 (Riesz) *Let $\phi : C_c(X) \to \mathbb{R}$ be a positive linear and continuous function, where X is a Hausdorff and locally compact space. There exists a uniquely determined Radon positive measure μ such that*

$$\phi(F) = \int_X F d\mu \qquad \forall F \in C_c(X).$$

If $G \in C_c(\mathbb{C})$ is a complex valued map, we can always decompose it into $G = G_1 + iG_2$ where G_1 and G_2 are real valued and supported on a compact. Since ϕ is linear, we get

$$\phi(G) = \phi(G_1) + i\phi(G_2) = \int_{\mathbb{C}} G_1 d\mu + i \int_{\mathbb{C}} G_2 d\mu = \int_{\mathbb{C}} G d\mu$$

so ϕ induces a unique measure μ. We can thus define a *spectral measure*.

Definition 4 Given $\{A_n\}_n \in \mathscr{E}$, we say that it has a spectral measure μ if

$$\lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} G(\lambda_i(A_n)) = \int_{\mathbb{C}} G d\mu$$

for every $G \in C_c(\mathbb{C})$.

Let $G_m \in C_c(\mathbb{C})$ be a sequence of nonnegative real valued maps such that $\|G_m\|_\infty \leq 1$ and

$$G_m(x) = 1 \qquad \forall\ |x| \leq m.$$

We find that

$$\int_{\mathbb{C}} G_m d\mu = \lim_{n\to\infty} \frac{1}{n} \sum_{i=1}^{n} G_m(\lambda_i(A_n)) \leq 1$$

and hence

$$\mu(\mathbb{C}) = \lim_{m\to\infty} \mu\left(\{x : |x| \le m\}\right) \le \limsup_{m\to\infty} \int_{\mathbb{C}} G_m d\mu \le 1.$$

This proves that all the measures we consider are finite. Since all the finite measures over the Borelian set are Radon, we will now simply say "measure" instead of "Radon measure." We showed that any measurable function induces a finite measure, but we can actually prove that it induces a probability measure, and also that any probability measure is induced by a function.

Lemma 6 *Let $D \subseteq \mathbb{R}^n$ be a measurable set with finite nonzero measure. Then, for any $k \in \mathscr{M}_D$ there exists a probability measure μ such that*

$$\frac{1}{l(D)} \int_D G(k(x))dx = \int_{\mathbb{C}} G d\mu \qquad \forall\, G \in C_c(\mathbb{C}).$$

Let J be the real interval $[0, 1]$. *Then for every probability measure μ there exists a measurable function $k \in \mathscr{M}_J$ such that*

$$\int_0^1 G(k(x))dx = \int_{\mathbb{C}} G d\mu \qquad \forall\, G \in C_c(\mathbb{C}).$$

Proof Given $k \in \mathscr{M}_D$, we already showed that Riesz theorem identifies a unique finite measure μ such that

$$\frac{1}{l(D)} \int_D G(k(x))dx = \int_{\mathbb{C}} G d\mu \qquad \forall\, G \in C_c(\mathbb{C}).$$

Let us consider $M > 0$ and denote

$$\chi_M(x) = \begin{cases} 1 & |x| \le M, \\ 0 & |x| > M. \end{cases}$$

Moreover, let us fix $\varepsilon > 0$, so that for every $M > 0$ we can find $G_M \in C_c(\mathbb{C})$ such that

$$\chi_M(x) \le G_M(x) \le \chi_{M+\varepsilon}(x) \qquad \forall x \in \mathbb{C}.$$

We infer

$$\int_{\mathbb{C}} \chi_{M-\varepsilon} d\mu \le \int_{\mathbb{C}} G_{M-\varepsilon} d\mu = \frac{1}{l(D)} \int_D G_{M-\varepsilon}(k(x))dx \le \frac{1}{l(D)} \int_D \chi_M(k(x))dx,$$

$$\frac{1}{l(D)} \int_D \chi_M(k(x))dx \le \frac{1}{l(D)} \int_D G_M(k(x))dx = \int_{\mathbb{C}} G_M d\mu \le \int_{\mathbb{C}} \chi_{M+\varepsilon} d\mu$$

so that

$$\int_{\mathbb{C}} \chi_{M-\varepsilon} d\mu \leq \frac{1}{l(D)} \int_D \chi_M(k(x)) dx \leq \int_{\mathbb{C}} \chi_{M+\varepsilon} d\mu.$$

When we let ε go to zero, we obtain that the integrals coincide on the indicator functions of closed intervals

$$\int_{\mathbb{C}} \chi_M d\mu = \frac{1}{l(D)} \int_D \chi_M(k(x)) dx.$$

The symbol $k(x)$ is a measurable function, so it is *sparsely unbounded*, meaning that

$$\lim_{M\to\infty} l(\{x : |k(x)| > M\}) = \lim_{M\to\infty} \int_D \chi_{|x|>M}(k(x)) dx = 0.$$

With the latter, we can conclude that μ is a probability measure

$$\mu(\mathbb{C}) = \lim_{M\to+\infty} \int_{\mathbb{C}} \chi_{|x|\leq M} d\mu = \lim_{M\to\infty} \frac{1}{l(D)} \int_D \chi_{|x|\leq M}(k(x)) dx = 1.$$

Given any probability measure μ, we know that the space $(\mathbb{C}, \mu)$ is a *standard probability space*, meaning that it is isomorphic to a space $X = I \sqcup E$, where I is a real finite interval with the Lebesgue measure, and $E = \{x_1, x_2, \dots\}$ is a discrete numerable set with an atomic measure ν. In particular, the isomorphism $\varphi : \mathbb{C} \to X$ satisfies

$$\mu(U) = l \oplus \nu(\varphi(U)) \qquad \forall\, U \in \mathscr{B}(\mathbb{C}).$$

and if the atomic measure is $\nu = \sum_{i=1}^{+\infty} c_i \delta_{x_i}$, then

$$1 = \mu(\mathbb{C}) = l \oplus \nu(X) = l(I) + \sum_{i=1}^{+\infty} c_i.$$

If we call $S = \nu(X) = \sum_{i=1}^{+\infty} c_i$, then we can take $I = [S, 1]$. Let $g : [0, 1] \to X$ be a map defined as

$$g(x) := \begin{cases} x_k & \sum_{i=1}^{k-1} c_i \leq x < \sum_{i=1}^{k} c_i, \\ x & x \geq S. \end{cases}$$

This has the same distribution as $l \oplus \nu$, since for every measurable map $G : X \to \mathbb{C}$ we obtain

$$\int_X G d(l \oplus \nu) = \sum_{i=1}^{+\infty} c_i G(x_i) + \int_S^1 G(x)dx = \int_0^1 G(g(x))dx.$$

Let now $k := \varphi^{-1} \circ g : [0, 1] \to \mathbb{C}$ be a measurable function, and $G \in C_c(\mathbb{C})$. We conclude that

$$\int_{\mathbb{C}} G d\mu = \int_X G \circ \varphi^{-1} d(l \oplus \nu) = \int_0^1 G(\varphi^{-1}(g(x)))dx = \int_0^1 G(k(x))dx.$$

□

A corollary of the latter result is that any sequence with a spectral symbol admits a probability spectral measure, and also the opposite holds. Moreover, if we call $\mathbb{P}$ the set of probability measures on $\mathbb{C}$, then we can also prove that any measure $\mu \in \mathbb{P}$ is a spectral measure.

Corollary 1 *All measures in $\mathbb{P}$ are spectral measures.*

Proof Let J be the real interval $[0, 1]$. Given any $k \in \mathscr{M}_J$, then there exists a sequence of continuous functions $k_m \in \mathscr{M}_J$ converging to k in measure. Using Lemma 5, we find that k is a spectral symbol, so every function in $\mathscr{M}_J$ is a spectral symbol.

Given now a measure $\mu \in \mathbb{P}$, Lemma 6 shows that it is induced by a measurable function in $\mathscr{M}_J$, so μ is also a spectral symbol. This implies that every measure in $\mathbb{P}$ is a spectral measure. □

3.2 Vague Convergence

We notice that every matrix A_n can be associated with an atomic probability measure μ_{A_n} with support on its eigenvalues

$$\mu_{A_n} := \frac{1}{n} \sum_{i=1}^{n} \delta_{\lambda_i(A_n)}.$$

Let us return again to the definition of spectral measure and notice that it can be rewritten as

$$\lim_{n\to\infty} \int_{\mathbb{C}} G d\mu_{A_n} = \int_{\mathbb{C}} G d\mu \qquad \forall G \in C_c(\mathbb{C}).$$

This is actually the definition of *vague convergence* for measures.

The space $\mathbb{P}$ endowed with the vague convergence is a complete metric space, using the *Lévy–Prokhorov metric* [10]

$$\pi(\mu, \nu) = \inf\{\, \varepsilon > 0 \mid \mu(A) \leq \nu(A^\varepsilon) + \varepsilon,\ \ \nu(A) \leq \mu(A^\varepsilon) + \varepsilon\ \forall A \in \mathscr{B}(\mathbb{C})\,\}$$

where

$$A^\varepsilon := \{\, x \in \mathbb{C} \mid dist(x, A) < \varepsilon\,\} = \{\, x + y \mid x \in A, |y| < \varepsilon\,\}.$$

Since every matrix is associated with an atomic probability measure, we can extend the definition of π to matrices and sequences.

Definition 5 Let $A, B \in \mathbb{C}^{n\times n}$ and let μ_A, μ_B be the probability atomic measures associated with their spectra, defined as

$$\mu_A := \frac{1}{n}\sum_{i=1}^{n} \delta_{\lambda_i(A)}, \qquad \mu_B := \frac{1}{n}\sum_{i=1}^{n} \delta_{\lambda_i(B)}.$$

The Lévy–Prokhorov metric on $\mathbb{C}^{n\times n}$ is defined as

$$\pi(A, B) := \pi(\mu_A, \mu_B).$$

The Lévy–Prokhorov metric on $\mathscr{E}$ is defined as

$$\pi(\{A_n\}_n, \{B_n\}_n) := \limsup_{n\to\infty} \pi(\mu_{A_n}, \mu_{B_n}).$$

Again, we need to prove that the latter is actually a pseudometric.

Lemma 7 *The Lévy–Prokhorov metric is a pseudometric on $\mathbb{C}^{n\times n}$ and a complete pseudometric on $\mathscr{E}$.*

Proof The Lévy–Prokhorov metric is an actual metric on the space of probability measures, so all the properties can be transferred to the space of matrices $\mathbb{C}^{n\times n}$, except for the identity of matrices with zero distance, since two different matrices may have the same eigenvalues. Thus it is a pseudometric on $\mathbb{C}^{n\times n}$, and by Lemma 2, it is a complete pseudometric on $\mathscr{E}$. □

Since every matrix is associated with an atomic probability measure, we can also use the same notation for mixed elements, like

$$\pi(A, \nu) := \pi(\mu_A, \nu).$$

The considered notation is useful since the definition of spectral measure is given by

$$\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}G(\lambda_i(A_n))=\int_{\mathbb{C}}G d\mu \qquad \forall G\in C_c(\mathbb{C})$$

and, when $\mu\in\mathbb{P}$, it can be rewritten as

$$\{A_n\}_n\sim_\lambda \mu \iff \pi(A_n,\mu)\xrightarrow{n\to+\infty}0.$$

The distance π on $\mathscr{E}$ is consistent with the distance between their spectral probability measures, as shown in the following result.

Lemma 8 *If $\{A_n\}_n\sim_\lambda\mu$ and $\{B_n\}_n\sim_\lambda\nu$, with $\{A_n\}_n,\{B_n\}_n\in\mathscr{E}$ and $\mu,\nu\in\mathbb{P}$, then*

$$\pi(\{A_n\}_n,\{B_n\}_n)=\pi(\mu,\nu)=\lim_{n\to\infty}\pi(A_n,B_n).$$

Proof Using the triangular property, we infer

$$\pi(\mu,\nu)\le\pi(\mu,A_n)+\pi(A_n,B_n)+\pi(B_n,\nu),$$

$$\pi(\mu,\nu)\ge-\pi(\mu,A_n)+\pi(A_n,B_n)-\pi(B_n,\nu).$$

Thus we obtain

$$\pi(\mu,\nu)\le\liminf_{n\to\infty}\pi(\mu,A_n)+\pi(A_n,B_n)+\pi(B_n,\nu)=\liminf_{n\to\infty}\pi(A_n,B_n),$$

$$\pi(\mu,\nu)\ge\limsup_{n\to\infty}-\pi(\mu,A_n)+\pi(A_n,B_n)-\pi(B_n,\nu)=\limsup_{n\to\infty}\pi(A_n,B_n).$$

By exploiting the latter relationships we conclude that

$$\begin{aligned}\pi(\{A_n\}_n,\{B_n\}_n)&=\limsup_{n\to\infty}\pi(A_n,B_n)\\&\le\pi(\mu,\nu)\le\\&\liminf_{n\to\infty}\pi(A_n,B_n)\le\pi(\{A_n\}_n,\{B_n\}_n).\end{aligned}$$

□

It is noteworthy to stress the importance of the probability condition on the measures. In fact, it is possible to find a sequence that admits a spectral measure but does not admit a spectral symbol, when the spectral measure is not a probability measure. Moreover, the Lévy–Prokhorov metric is defined only on probability

measures and if $\mu_n \in \mathbb{P}$ vaguely converge to a measure not in $\mathbb{P}$, then the sequence μ_n is not even a Cauchy sequence for π.

4 Main Results

4.1 Connection Between Measures

First of all, we prove that π and d' are equivalent pseudometrics on $\mathscr{E}$.

Lemma 9 *If $\{A_n\}_n, \{B_n\}_n \in \mathscr{E}$, then*

$$\pi(\{A_n\}_n, \{B_n\}_n) \leq d'(\{A_n\}_n, \{B_n\}_n) \leq 2\pi(\{A_n\}_n, \{B_n\}_n).$$

Proof Let us first prove that for any $A, B \in \mathbb{C}^{n\times n}$, we have

$$\pi(A, B) \leq d'(A, B) \leq 2\pi(A, B).$$

Let $\Lambda(A)$ and $\Lambda(B)$ be ordered so that

$$i < j \implies |\lambda_i(A) - \lambda_i(B)| \geq |\lambda_j(A) - \lambda_j(B)|$$

and

$$s := d'(A, B) = \frac{k-1}{n} + |\lambda_k(A) - \lambda_k(B)|.$$

In particular, we deduce that

$$|\lambda_i(A) - \lambda_i(B)| \leq s \qquad \forall i \geq k$$

and consequently, for any subset $U \subseteq \mathbb{C}$, we obtain the inequality

$$\#\{\lambda_i(A) \in U,\ i \geq k\} \leq \#\{\lambda_i(B) \in U^s,\ i \geq k\}.$$

Denote with μ_A and μ_B the atomic probability measures associated with A, B. Let $U \in \mathscr{B}(\mathbb{C})$ be any Borelian set and denote the cardinality of the intersection with a n-uple v as

$$Q_U(v) := \#\{i : v_i \in U\}.$$

Formally, $Q_U(v)$ is the number of elements of v inside v, counted with multiplicity. We know that

$$\begin{aligned}\mu_A(U) &= \frac{Q_U(\Lambda(A))}{n}\\ &= \frac{Q_U(\{\lambda_i(A) : i \geq k\})}{n} + \frac{Q_U(\{\lambda_i(A) : i < k\})}{n}\\ &\leq \frac{Q_{U^s}(\{\lambda_i(B) : i \geq k\})}{n} + \frac{k-1}{n}\\ &\leq \frac{Q_{U^s}(\Lambda(B))}{n} + s = \mu_B(U^s) + s.\end{aligned}$$

We symmetrically obtain also the following relation:

$$\mu_B(U) \leq \mu_A(U^s) + s.$$

As a consequence

$$\pi(A, B) = \inf\{\, \varepsilon > 0 \mid \mu_A(U) \leq \mu_B(U^\varepsilon) + \varepsilon,\ \mu_B(U) \leq \mu_A(U^\varepsilon) + \varepsilon\ \forall U \in \mathscr{B}(\mathbb{C})\,\}$$

$$\implies \pi(A, B) \leq s = d'(A, B).$$

Denote now $r = \pi(A, B)$ and let T be any sub-uple of $\Lambda(A)$. If we see T as a set, then it is a finite subset of $\mathbb{C}$, so it is a Borelian set. Given any $\varepsilon > 0$ we know that

$$\mu_A(T) = \frac{Q_T(\Lambda(A))}{n} \leq \mu_B(T^{r+\varepsilon}) + r + \varepsilon = \frac{Q_{T^{r+\varepsilon}}(\Lambda(B))}{n} + r + \varepsilon$$

so we deduce that

$$\frac{Q_T(\Lambda(A))}{n} \leq \frac{Q_{T^r}(\Lambda(B))}{n} + r \implies Q_T(\Lambda(A)) \leq Q_{T^r}(\Lambda(B)) + rn.$$

By using the fact that the map Q is integer valued, we conclude that

$$Q_T(\Lambda(A)) \leq Q_{T^r}(\Lambda(B)) + \lfloor rn \rfloor.$$

The quantity $Q_T(\Lambda(A))$ is actually the cardinality of T seen as a sub-uple of $\Lambda(A)$, so for every subset T of k eigenvalues in A, even repeated, there are at least $k - \lfloor rn \rfloor$ eigenvalues of B that have distance less than r from one of the elements of T.

Let us now build a bipartite graph, where the left set of nodes L contains the elements of $\Lambda(A)$, the right set of nodes R contains the elements of $\Lambda(B)$, and $\lfloor rn \rfloor$ additional nodes. Every additional node is connected to all the elements of L, and an element of $\Lambda(A)$ is connected to an element of $\Lambda(B)$ if and only if their distance

is less than r. If we denote E the set of edges of the graph and N the set of its nodes, then we can define the neighborhood of a subset of nodes $P \subseteq N$ as

$$N(P) := \#\{u \in N : \exists v \in P, (v, u) \in E\}.$$

By using the previous derivations, we know that for any $T \subseteq L = \Lambda(A)$ it holds

$$N(T) \geq \#T - \lfloor rn \rfloor + \lfloor rn \rfloor = \#T.$$

Thanks to Hall's marriage theorem that can be found, for example, in [7], there exists a matching for L, meaning that there exists an injective map $\alpha : L \to R$ such that

$$(u, \alpha(u)) \in E \qquad \forall\, u \in L.$$

Now let us consider the set

$$P := \{u \in L : \alpha(u) \in \Lambda(B)\}.$$

we know that $\#P \geq n - \lfloor rn \rfloor$, and we can enumerate the eigenvalues in $\Lambda(A) = L$ and $\Lambda(B)$ so that

$$\lambda_i(A) \in P, \quad \lambda_i(B) = \alpha(\lambda_i(A)) \qquad \forall\, i \leq n - \lfloor rn \rfloor.$$

Since u and $\alpha(u)$ are connected for all $u \in L$, we deduce that $\lambda_i(A)$ and $\lambda_i(B)$ are connected for at least $n - \lfloor rn \rfloor$ indices. By construction,

$$|\lambda_i(B) - \lambda_i(A)| < r \qquad \forall\, i \leq n - \lfloor rn \rfloor$$

so

$$\begin{aligned} d'(A, B) &= \min_{\sigma \in S_n} \min_{i=1,\ldots,n+1} \left\{ \frac{i-1}{n} + |\Lambda(A) - \Lambda(B)_\sigma|_i^\downarrow \right\} \\ &\leq \min_{i=1,\ldots,n+1} \left\{ \frac{i-1}{n} + |\Lambda(A) - \Lambda(B)|_i^\downarrow \right\} \\ &< \frac{\lfloor rn \rfloor}{n} + r \leq 2r = 2\pi(A, B). \end{aligned}$$

This proves that for any $A, B \in fC^{n \times n}$ we have

$$\pi(A, B) \leq d'(A, B) \leq 2\pi(A, B).$$

Given now $\{A_n\}_n, \{B_n\}_n \in \mathscr{E}$, we conclude

$$\pi(\{A_n\}_n, \{B_n\}_n) = \limsup_{n\to\infty} \pi(A_n, B_n) \leq \limsup_{n\to\infty} d'(A_n, B_n) = d'(\{A_n\}_n, \{B_n\}_n),$$

$$d'(\{A_n\}_n, \{B_n\}_n) = \limsup_{n\to\infty} d'(A_n, B_n) \leq \limsup_{n\to\infty} 2\pi(A_n, B_n) = 2\pi(\{A_n\}_n, \{B_n\}_n).$$

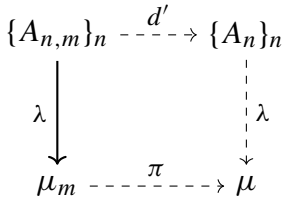
□

The two distances d' and π are equivalent, so they induce the same topology on the space $\mathscr{E}$ and they respect a property of closeness given by the following lemma.

Lemma 10 *Let $\{A_{n,m}\}_n \sim_\lambda \mu_m$, where $\{A_{n,m}\}_n \in \mathscr{E}$ and $\mu_m \in \mathbb{P}$ for every m. If we consider the statements below*

1. $\pi(\mu_m, \mu) \xrightarrow{m\to\infty} 0$,
2. $\{A_n\}_n \sim_\lambda \mu$,
3. $d'(\{A_{n,m}\}_n, \{A_n\}_n) \xrightarrow{m\to\infty} 0$,

where $\{A_n\}_n \in \mathscr{E}$ and $\mu \in \mathbb{P}$, then any two of them are true if and only if all of them are true.



Proof 1.3. $\implies$ 2.) We know that

$$\pi(A_n, \mu) \leq \pi(A_n, A_{n,m}) + \pi(A_{n,m}, \mu_m) + \pi(\mu_m, \mu) \qquad \forall n, m.$$

Given $\varepsilon > 0$, we can find M such that

$$\pi(\mu_m, \mu) \xrightarrow{m\to\infty} 0 \implies \pi(\mu_m, \mu) < \varepsilon \quad \forall m > M,$$

$$d'(\{A_{n,m}\}_n, \{A_n\}_n) \xrightarrow{m\to\infty} 0 \implies d'(\{A_{n,m}\}_n, \{A_n\}_n) < \varepsilon \quad \forall m > M.$$

Using Lemma 9, we obtain

$$\limsup_{n\to\infty} \pi(A_{n,m}, A_n) = \pi(\{A_{n,m}\}_n, \{A_n\}_n) \leq d'(\{A_{n,m}\}_n, \{A_n\}_n).$$

We can then fix $m > M$ and find $N > 0$ such that

$$\pi(A_{n,m}, A_n) \leq 2\varepsilon, \quad \pi(A_{n,m}, \mu_m) \leq \varepsilon \qquad \forall n > N.$$

We obtain that

$$\pi(A_n, \mu) \leq 2\varepsilon + \varepsilon + \varepsilon = 4\varepsilon \qquad \forall n > N,$$