UNIVERSITÀ DEGLI STUDI DI ROMA "LA SAPIENZA"
Dipartimento di Matematica, Dipartimento di Metodi e Modelli Matematici per le Scienze Applicate
ISTITUTO NAZIONALE DI ALTA MATEMATICA "FRANCESCO SEVERI"

RENDICONTI DI MATEMATICA

E DELLE SUE APPLICAZIONI

Serie VII - Volume 28 - Fascicolo I - 2008

Le modalità per la presentazione dei lavori e le condizioni di vendita si trovano alla web page.

Submission of papers and subscription rate information could be found on the web page.

ADDRESS:
RENDICONTI DI MATEMATICA
DIPARTIMENTO DI MATEMATICA - PIAZZALE A. MORO, 2 - 00185 ROMA, ITALIA

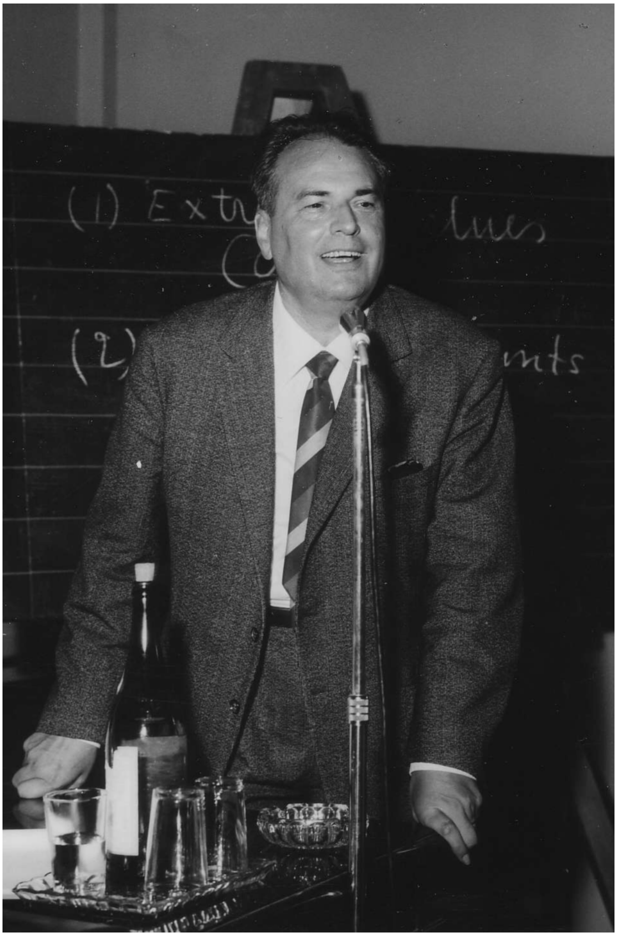
Fax.: +39 06 49913052
e-mail: rendmat@mat.uniroma1.it
web page: www.dmmm.uniroma1.it/~rendiconti

Direttore Responsabile: Prof. ALESSANDRO SILVA

Autorizzazione del Tribunale di Roma del 22-11-2005 (n. 456/05)

Composizione tipografica effettuata con il programma TEX © dalla
CompoMat – Loc. Braccone – 02040 Configni (RI) – Telefax 39 0746 672240

Stampato a cura del *Centro Stampa d'Ateneo* dell'Università di Roma *"La Sapienza"*



Rendiconti di Matematica, Serie VII
Volume 28, Roma (2008), i–ii

Foreword

Bruno de Finetti had a very versatile personality and was actively involved in many different fields of Science and Culture. In fact, not only he gave his determinant contributions to Mathematics and its applications, in particular Probability, Statistics, Financial and Actuarial Mathematics, but he was also deeply interested in Epistemology, Philosophy, Economics, Politics. Still in our days, the contributions that he gave to different fields are often revisited and, in many cases, they have been discovered to contain foresight and fertile suggestions for contemporary research.

Bruno de Finetti was born, from Italian parents, in Innsbruck on June 13, 1906. During 2006 a number of different events have been organized to celebrate the centenary of his birth. Several scientific and cultural institutions have organized scientific events with the purpose of highlighting the different facets of his scientific and human personality and to discuss his advanced contributions.

In particular, on November 15-17, 2006, in Rome, the International Symposium “Bruno de Finetti Centenary Conference” was held at the Department of Mathematics of University “La Sapienza” and Accademia dei Lincei, two institutions of which he had been a member for many years.

Already in 1981, the celebration of the 75^{th} birthday of Bruno de Finetti offered University La Sapienza and Accademia dei Lincei the opportunity for the organization of an international conference, that was specifically devoted to the theme of Exchangeability (the contributions presented at that conference were collected in the Volume *Exchangeability in Probability and Statistics* (G. Koch and F. Spizzichino, Eds), North Holland, Amsterdam, 1982).

The Symposium held on November 2006 consisted of two different parts, respectively devoted to the themes *de Finetti’s Legacy in Probability Today* and *Bruno de Finetti and Economic Analysis.*

The Organizing Committee was formed by Edoardo Vesentini (Politecnico di Torino and Accademia dei Lincei), Pierluigi Ciocca (Banca d’Italia), Giorgio

Letta (Università di Pisa and Accademia dei Lincei) Giorgio Lunghini (Università di Pavia and Accademia dei Lincei), Carlo Sbordone (Università Federico II, Napoli, and Accademia dei Lincei), and Fabio Spizzichino (Università La Sapienza, Roma). The Local Organizing Committee was formed by Giovanna Nappo, Mauro Piccioni and Fabio Spizzichino at the Department of Mathematics of Università La Sapienza.

The aim of the Symposium was just to discuss some problems, arising in the frame of recent research in the fields of Probability and Economics, respectively, and to trace the connections with ideas and anticipations contained in the work by Bruno de Finetti.

In particular the Sessions on *de Finetti's Legacy in Probability Today* were devoted to rethink some purely mathematical aspects of the contributions that he gave in the Thirties of last Century, and to present related developments that have been obtained in recent times.

These sessions were constituted by general lectures, given by a few selected speakers. Listed in the order they appeared in the Programme, the invited speakers were Eugenio Regazzini (Università di Pavia), Olav H. Kallenberg (Auburn University), Persi Diaconis (Stanford University), Steffen L. Lauritzen (Oxford University), Paul Ressel (Katholische Universität Eichstatt), Yoseph Rinott (Hebrew University, Jerusalem), and Murad Taqqu (Boston University).

The complete scientific programme of the Symposium can be found in the next pages.

More details about the Symposium can be found at the web-site:

www.mat.uniroma1.it/ricerca/convegni/2006/deFinetti/

The programme of the first session, in particular, included also an opening address by Fulvia de Finetti, the daughter of Bruno de Finetti, the presentation of *Opere di Bruno de Finetti,* and the Exhibition of Historical Books *Probability from Cardano to de Finetti.*

With her kind permission, the text of the contribution by Fulvia de Finetti is reported in this Volume.

Opere di Bruno de Finetti is a two Volumes edition published by Unione Matematica Italiana (UMI); the first Volume, in particular, is devoted to his contributions to Probability, Statistics and Decision Theory.

The book exhibition was held in the Library of the Department of Mathematics and was organized by Giovanna Nappo with the precious collaboration of the staff of the Library; more details can be found at the web-site address http :// www.mat.uniroma1.it/ricerca/convegni/2006/deFinetti/mostra/historical-books-details.html

The present Volume, edited by Giovanna Nappo, Mauro Piccioni and Fabio Spizzichino, collects the contributions that the Authors, invited to lecture on the theme *de Finetti's Legacy in Probability Today*, delivered after the Symposium.

Scientific Program
of International Symposium

Bruno de Finetti Centenary Conference

Rome, November 15-17, 2006

Università "La Sapienza"
Piazzale A. Moro, 5

Accademia Nazionale dei Lincei
Via della Lungara, 230

Wednesday, November 15 : Mathematical Session 1

De Finetti's Legacy in Probability Today

Aula Magna
Università "La Sapienza"
Piazzale A. Moro, 5

8:30 – 9:15 Registration of Participants

9:15 Opening of Conference

Chairman: Marco Scarsini

9:45 – 10:45 Eugenio Regazzini (Università di Pavia, Italy)
De Finetti's contribution to the theory of random functions

10.45 – 11.15 Coffee Break

11:15 – 12:00 Presentation of the Volumes
Opere di Bruno de Finetti Edited by Unione Matematica Italiana

Chairman: Giorgio Koch

15:00 – 16:00 Olav H. Kallenberg (Auburn University , USA)
Some highlights from the theory of multivariate symmetries

16:00 – 16:30 Coffee Break

16:30 – 17:30 Persi Diaconis (Stanford University, USA)
Exchangeability in the Twenty First Century

12:00 – 18:00 Exhibition of Historical Books
Probability from Cardano to de Finetti
Library of Department of Mathematics "G. Castelnuovo"

Thursday, November 16: Mathematical Session 2

De Finetti's Legacy in Probability Today

Accademia Nazionale dei Lincei
Via della Lungara, 230

Ore 9:00 Opening

Chairman: Carlo Sbordone

10:00 – 11:00 Steffen L. Lauritzen (Oxford University, UK)
Exchangeable Rasch Models

11:00 – 11:30 Coffee Break

11:30 – 12:30 Paul Ressel (Katholische Universitat Eichstatt, Germany)
Exchangeability and semigroups

Chairman: Wolfgang Runggaldier

15:00 – 16:00 Yoseph Rinott (Hebrew University, Israel)
Exchangeability, concepts of dependence, and statistical implications

16:00 – 16:30 Coffee Break

16:30 – 17:30 Murad Taqqu (Boston University, USA)
Dependence structures of some infinite variance stochastic processes

Friday, November 17: Economics Session

Bruno de Finetti and Economic Analysis

Accademia Nazionale dei Lincei
Via della Lungara, 230

Ore 9:00 Opening

Chairman: Siro Lombardini

9:30 – 10:30 Giorgio Lunghini (University of Pavia, Italy)
Bruno de Finetti and Economic Theory

10:30 – 11:00 Coffee Break

11:00 – 12:00 Luca Barone (Goldman-Sachs International, UK)
Mean-Variance Portfolio Selection: de Finetti scoops Markowitz

12:00 – 13:00 Flavio Pressacco (University of Udine, Italy)
B. de Finetti hero of the two worlds: (applied) mathematician and (quantitative) economist

13:00 Conclusions and Closure of Conference

Rendiconti di Matematica, Serie VII
Volume 28, Roma (2008), vii–viii

Bruno de Finetti Centenary Conference
Rome, November 15, 2006

FULVIA DE FINETTI

It is a pleasure for me to be once more in this University where my father spent many years of his academic career. As you all probably know my father left the University of Trieste in 1954 when he conquered a chair in the University of Rome and it was in 1961 that he moved from the Faculty of Economics to the Faculty of Science to teach Probability. This was the happy end of a story that goes back to 1927 when my father, just graduated, reached Rome to work as head of the Mathematical Service of the Istituto Centrale di Statistica. He had already in his mind Probability and after finishing his work at the Institute he used to attend the seminars that took place in Panisperna Street, next door to his office. He was particularly impressed by Enrico Fermi whose rapid career became a target for him. He immediately gained the attention of the great mathematicians that worked in Rome at that time: Guido Castelnuovo, Tullio Levi-Civita, Federigo Enriques.

Guido Castelnuovo at that time taught probability and so he explained why he found probability an interesting topic to teach: “Probability is a science of recent formation; hence in it, better than in other branches of mathematics, one can see the relationship between the empirical contribution and the one given by reasoning, and between the process of inductive and deductive logic used in it. The fact that it is a science in the making explains why it is appropriate to give frequent examples to show the applications of known methods or to introduce new ones.”

If you compare Castelnuovo’s sentence with what Bruno wrote to his mother when student in Milan: … Mathematics is not by now a field already explored, just to learn and pass on to posterity as it is. It is always progressing, it is enriching and lightening itself, it is a lively and vital creature, in full development and just for these reasons I love it, I study it and I wish to devote my life to it . . .

It will not surprise you to know that Castelnuovo often invited Bruno at his home to see the progresses in the works of this promising young mathematician. To open the door was a little girl wearing her hair in pigtails. She was Emma the daughter of Castelnuovo. In a letter dated July 28, 1928 Castelnuovo examines the work of Bruno, recognizes his capabilities as analyst, gives advices on how to present the work and concludes, "I feel sure that you will be able to give important contributions to Probability Calculus and its applications". And so it seems he did if important names in this field have accepted to come from U.S.A. and Europe for this Centenary Conference.

When in 1961 the Faculty of Science decided to resume for Bruno the chair of Probability that had been of Castelnuovo but extinguished when he left, the main concern of my father was that the same thing might happen when he would leave. In April 1973, Savage had already died, my father received from the University of Michigan an invitation for the year 1973-74. I think it may interest you to read part of the answer of my father to decline the invitation: "...I am very pleased and honoured for such attracting invitation and for the interest in my research . . . and in my point of view about subjective probability. I would be surely willing to support it, especially in your University where L.J. Savage spent several years of his admirable activity . . . "I am involved in many programs here, highly depending on myself (my collaborators are too young to be fully responsible for the courses)."

I am sure my father would be very happy to know that the Faculty of Science, not only preserved the chair of Probability but even increased to three the number of chairs and that some of those "too young collaborators" are continuing his teaching and research. Among them I want particularly to thank Professor Fabio Spizzichino who promoted and organized this International Centenary Conference.

Rendiconti di Matematica, Serie VII
Volume 28, Roma (2008), ix–xiii

Preface

GIOVANNA NAPPO – MAURO PICCIONI

FABIO SPIZZICHINO

De Finetti started his active research in Probability around 1927, at the age of 21. At that time, his interest was mainly addressed to the foundations of Probability. His interest on this topic focused on two different, even though related, aspects: the meaning of the term probability of an event and the axioms that should be imposed in the formalization of the mathematical theory.

In fact, the main motivations for his research can be traced back to his dissent and his objections to ideas and viewpoints about probability that were common among mathematicians and other scientists. His reflections ripped his firm opinion that probability cannot have but a subjective character. This conclusion lead him to rethink the meaning of stochastic independence and to introduce the concept of Exchangeability in the foundations of Probability and Statistics.

On the other hand he also strongly defended the idea that probability must satisfy the simple-additivity property while, in his view, the stronger assumption of countable additivity is not necessary and may give rise indeed to different drawbacks.

Another topic which attracted his interest at that time was the definition and the analysis of Stochastic Processes with Independent Increments. He first analyzed the finite-dimensional distributions of these processes, thus arriving to single out the class of infinitely divisible distributions. This piece of research was carried out, later on, by himself, P. Lévy, A. Y. Khinchin, A. N. Kolmogorov and others, in a series of papers devoted to the characterization of infinitely divisible distributions.

The two basic topics mentioned so far constitute the primary elements for the contributions collected in this Volume. The articles by Eugenio Regazzini with Federico Bassetti, Olav Kallenberg, Persi Diaconis with Svante Janson, Steffen L. Lauritzen, and Paul Ressel are related to the theme of exchangeability, whereas the article authored by Murad S. Taqqu with Joshua B. Levy, concerns the developments of the theory of stochastic processes with independent increments.

It is well known that exchangeability is nothing but a simple concept of symmetry: a (finite-dimensional) vector of random variables is exchangeable if its joint distribution is invariant under permutation of coordinates. A denumerable sequence is exchangeable if and only if any finite subsequence is exchangeable.

The simplest example of an exchangeable sequence is the case of independent, identically distributed random variables. A natural extension is the case of sequence of conditionally independent, identically distributed random variables. The celebrated de Finetti's theorem about exchangeability guarantees that denumerable sequence of exchangeable variables are necessarily conditionally i.i.d..

Of course the notion of exchangeability can be extended and generalized to cope with many other interesting situations, also in more abstract settings. The interest for this topic dates back at least to late Thirties with de Finetti's work about partial exchangeability ("equivalence partielle").

Among developments related to de Finetti's theorem, a special type of problems stimulated the interest of several probabilists: to characterize the class of exchangeable models manifesting some further conditions of invariance, under different groups of transformations. A related problem, obtained by reversing this point of view, is to find invariance properties which characterize all the joint distributions obtained as mixture of i.i.d. variables with common distribution belonging to a special class (e. g. the cases of conditionally i.i.d. Gaussian variables or of conditionally i.i.d. exponential variables). This class of problems is of basic importance in the construction and the theoretical study of statistical models (especially in a Bayesian context) and is strictly related with the theory of sufficient statistics. From the mathematical point of view this subject translates into the problem of characterizing the extremal points of some convex spaces. Results in this direction have been called de Finetti-type theorems.

The fields of exchangeability and de Finetti-type theorems have been extensively developed in the last few decades and important contributions have been given by Authors of articles that appear in the present collection. The articles presented here contain recent results and also provide a review of relevant issues from the literature.

The collection opens with the article by Eugenio Regazzini and Federico Bassetti whose main purpose is to present a detailed review of the first paper published by de Finetti on the theme of Exchangeability, which is not very well known. In this paper de Finetti performed the analysis of exchangeable events by means of the method of the characteristic functions. Since this was also the main mathematical tool used in de Finetti's studies about processes with independent increments, this paper can be considered as a bridge between exchangeability and processes with independent increments. However, the interest of the article by Regazzini and Bassetti goes beyond strictly historical aspects. In particular, following a suggestion contained in the de Finetti's paper, they obtain necessary and sufficient conditions of an algebraic type for the extendibility of a finite sequence of exchangeable events.

The article by Olav Kallenberg provides a useful review concerning the area of probabilistic symmetries, that can be seen as a natural development of the studies about exchangeability and partial exchangeability. An important part of this theory has been systematically developed by Kallenberg himself in a recent monograph (2005). In the present article representation theorems for multi-dimensional arrays of random variables, which are invariant under suitable groups of transformations, are discussed. More precisely, for two-dimensional arrays, the following results are presented: the theorems by Aldous and Hoover on the separate and joint exchangeability, the results by Kallenberg himself on the equivalence between *contractibility* (the operation of taking minors) and exchangeability, and those about invariance under rotations (*rotatability*) obtained by Aldous (in the separate case) and Kallenberg (in the joint case). In order to discuss rotatability in dimension greater than 2, the framework of continuous linear random fuctionals (CLRF) is introduced, which allows to obtain representations involving multiple Wiener-Ito integrals.

The article by Persi Diaconis and Svante Janson gives a somewhat different perspective on the concept of partial exchangeability through random graphs. The paper develops a theoretical framework with the purpose of giving a probabilistic interpretation to the notions of convergence of graphs and of infinite graph limits which have been recently investigated by Lovász and coauthors. Within this framework, the Lovász-Szegedy characterization of infinite graph limits is translated, in terms of adjacency matrix of the graph, into the Aldous-Hoover representation for two dimensional jointly exchangeable arrays. A similar analysis is performed for bipartite graphs (leading to separately exchangeable arrays) and directed graphs, showing how the results discussed by Kallenberg in the previous paper can be used in this graph theoretical setting.

Some general aspects of the field of de Finetti-type theorems are analyzed in Paul Ressel's article. This article discusses the central role of semigroups in the description of the general mathematical structure that is at the basis of the theory. Actually, the author proves a general theorem on the representation, as a mixture of *characters*, of positive definite functions defined on a semigroup. By using this result a method is provided to give alternative proofs to known theorems on exchangeability, such as the Hewitt-Savage theorem, and characterizations of mixtures of i.i.d. samples from specific parametric families. The paper presents also an application to the exchangeable partitions introduced by Kingman.

The contribution by Steffen L. Lauritzen deals with random infinite matrices which are not only row-column exchangeable (called *separately exchangeable* by Kallenberg) but also *row-column summarized* (RCES). This means that their finite-dimensional distribution is a function of the row and the column sums (these sums are actually semigroups statistics, hence they fall within the theory developed in the paper by Ressel). It is shown that, within the class of "regular" RCES random matrices, the random Rasch matrices (that were introduced in mathematical psychology as models for intelligence tests) play a special role: they constitute the extreme points. A similar analysis is performed for weakly exchangeable matrices (called *jointly exchangeable* by Kallenberg). The relations with random graphs are also addressed in the article; in particular weakly exchangeable Markov graphs, used in social network analysis, are discussed.

As mentioned, the contribution delivered by Murad S. Taqqu and Joshua B. Levy is related with developments of the original de Finetti's work on processes with independent increments. Their contribution deals with the phenomenon of long-range dependence of symmetric α- stable (S-α-S) log-fractional motions with index $\alpha \in (1,2)$. Any such process is self-similar with index $H = \frac{1}{\alpha}$ and it has infinite variance; therefore the structure of dependence between the values of the process at two different time-instants cannot be described in terms of the covariance, but rather through the use of the so-called codifferences and covariations, whose behavior is described in detail.

Bruno de Finetti gave relevant contributions in many different areas and for his contributions he has been often celebrated. But, first of all, he was a mathematician. We are particularly glad to have the possibility, with the publication of the present Volume, to honour the memory of him as a mathematician. For this reason, we would like to thank all the authors for their enthusiastic and high level collaboration.

We also thank Fulvia de Finetti for contributing this Volume with her opening address and for her stimulating and friendly suggestions during the preparation of the Symposium.

We gratefully acknowledge the kind attitude of the Scientific Board of Rendiconti di Matematica in dedicating an issue of the Journal to these Proceedings. In particular we thank our colleague Alessandro Silva, who promoted this editorial initiative and gave us a friendly support along the whole publication process.

Acknowledgements

The Symposium "Bruno de Finetti Centenary Conference" has been organized under the auspices of UMI, Accademia dei Lincei, "La Sapienza"- Università di Roma, and the Department of Mathematics. We like to thank the members of the Organizing Committee among whom, in particular, Carlo Sbordone, whose personal efforts made it possible to organize the Symposium. "La Sapienza" and the Department of Mathematics also gave a relevant support for the organization, both from a financial and a logistic point of view. University "La Sapienza" and Accademia dei Lincei provided prestigious sites for the different Sessions, while Istituto Italiano per gli Studi Filosofici provided a determinant financial and logistic support. Financial support also came from Banca d'Italia and the National Research Project "Metodi Stocastici in Finanza Matematica", PRIN 2004. The help from all these Institutions and Organizations is very gratefully acknowledged.

We also acknowledge the collaboration of cultural organizations such as Associazione Italiana per le Scienze Economiche e Sociali, Società Italiana di Statistica, Società Italiana degli Economisti, and Società Italiana per la Storia dell'Economia.

Rendiconti di Matematica, Serie VII
Volume 28, Roma (2008), 1–17

The unsung de Finetti's first paper about exchangeability

FEDERICO BASSETTI – EUGENIO REGAZZINI

Abstract: *It is a singular fact that the first and pithy de Finetti's essay on exchangeability has not earned the same reputation as that of others of his papers about the same subject. In fact, this paper contains, on the one hand, all the main results on sequences of exchangeable events, together with the right subjectivistic interpretation of the role they play in the study of the connections between probability and frequencies. On the other hand, the paper makes use of mathematical methods abandoned, immediately after its publication, by de Finetti himself. The center of this methods is the so–called characteristic function of a random phenomenon. Independently of the destiny of the paper, we think that, apart from its undoubted historical value, it contains ideas susceptible of interesting new developments. Therefore, we have deemed it suitable to give here a detailed and faithful account of its content, for the benefit of the colleagues who are not in a position to understand Italian. Moreover, to emphasize the value of the paper at issue, we develop de Finetti's brief hint to the extendibility of exchangeable sequences of events, to obtain a new explicit necessary and sufficient condition of an algebraic nature.*

1 – Introduction

Bruno de Finetti (1906–1985) is regarded as the founder of the theory of *sequences* of *exchangeable random variables* or *random exchangeable sequences* for short. His first important article about this subject dates back to 1930 (see [6]). It appears as a *Memoria*, published in the proceedings of the *Accademia*

Key Words and Phrases: *Characteristic function of a random phenomenon – De Finetti's contributions to probability – Exchangeable events – Extendibility of exchangeable sequences.*
A.M.S. Classification: 60G09, 60-03, 01A60

dei Lincei, jointly presented to the *Accademia* by two of the most outstanding Italian scientists of the time: Guido Castelnuovo (1865–1952) and Tullio Levi–Civita (1873–1941). It is clear, from his biography, that de Finetti had examined the problem of finding a probabilistic description and interpretation for *random phenomena* – those which can be repeatedly observed under homogeneous environmental conditions – ever since his early approach to probability, a couple of years before his degree in mathematics, obtained in 1927 at the University of Milan. In fact, he presented a summary of the *Memoria* in a pithy communication at the *International Mathematical Congress* held in Bologna, in September 1928. The text of such a communication was published, in 1932, in the sixth volume of the proceedings of the *Congress*. See [11].

The present paper aims at giving a precise idea of the content of the *Memoria* and, especially, of the methods used therein, since they are different from those employed in later de Finetti's contributions to the same subject. This analysis is split into six points which form Section 2. Some new developments of de Finetti's original methods are sketched in Section 4. The intermediate Section 3 reviews a paper by Jules Haag (1882–1953) in which, so far as we know, the concept of exchangeable events had been introduced and studied for the first time. A comparison between this paper and de Finetti's *Memoria* clarifies the complete independence of the two papers, and it convinces of the prominent merits of de Finetti in this field.

2 – Characteristic function of a random phenomenon

In view of the homogeneity of the environmental conditions which distinguishes random phenomena (with *equivalent* trials) from other types of phenomena, de Finetti points out that a correct probabilistic translation of such an empirical circumstance leads us to think of the probability of m successes and $(n-m)$ failures, in n trials, as invariant with respect to the order in which successes and failures alternate, whatever n and m may be. Accordingly, he defines a sequence $(E_n)_{n\geq 1}$ of events to be *equivalent* if, for every *finite* permutation π, the probability distribution of $(\mathbb{I}_{E_1}, \mathbb{I}_{E_2}, \dots)^{(1)}$ is the same as the probability distribution of $(\mathbb{I}_{E_{\pi(1)}}, \mathbb{I}_{E_{\pi(2)}}, \dots)$. So, if $\omega_k^{(n)}$ denotes the probability that the random phenomenon, taken into consideration, comes true k times in m trials, one gets

$$\omega_k^{(m)} = \sum_{h=k}^{n-m+k} \omega_h^{(n)} \frac{\binom{h}{k}\binom{n-h}{m-k}}{\binom{n}{m}} \tag{1}$$

whenever $1 \leq m \leq n$ and $k = 0, \dots, m$.

$^{(1)}$For any event E, $\mathbb{I}_E$ will stand for its indicator.

Nowadays, the term *equivalent* is replaced by the more expressive and unambiguous word *exchangeable*, proposed perhaps by Pólya (cf.[13], [14]) or by Fréchet (cf. [16]).

From (1) de Finetti derives a difference–differential equation for the probability generating function and, consequently, for the characteristic function, of the frequency of success in n trials. Such a characteristic function and its limit, as $n \to +\infty$, in the case of an infinite sequence, becomes the center of de Finetti's treatment of exchangeability.

From a methodological viewpoint, the use of characteristic functions joins the *Memoria* to the contemporaneous de Finetti's studies on processes with stationary independent increments, based on the analysis of the *derivative law* defined in terms of the characteristic function ψ_λ of the λ–th coordinate of the process; [5]. See also [24].

As already recalled in the first section, the application of the characteristic functions method to the study of exchangeable sequences is a peculiarity of the *Memoria*. In point of fact, on Khinchin's advice, in all subsequent papers on exchangeable random elements, de Finetti uses more direct tools such as probability distribution functions, moments, and so on. We have experimented that the original de Finetti approach has some remarkable merits with respect to some important problems like, for instance, concrete assessment of finitary exchangeable laws and extendibility of exchangeability. So, we believe that an accurate and faithful account of that approach could come in handy to all scholars who are unable to read Italian scientific literature.

2.1 – Fundamental recurrence relation.

Our description starts with Author's remark that (1), for $m = n-1$, reduces to

$$n\omega_k^{(n-1)} = (n-k)\omega_k^{(n)} + (k+1)\omega_{k+1}^{(n)}$$

with $\omega_0^{(0)} = 1$. Thus, for any sequence of N exchangeable events, he deduces the difference–differential equation

$$n\Omega_{n-1}(z) = n\Omega_n(z) + (1-z)\Omega_n'(z) \tag{2}$$

valid for $n = 1, \ldots, N$ and any complex number z, where Ω_n is the probability generating function defined by

$$\Omega_n(z) := \sum_{h=0}^{n} \omega_h^{(n)} z^h \qquad (n = 1, 2, \ldots, N;\ z \in \mathbb{C}) \tag{3}$$

with $\Omega_0(z) \equiv 1$.

Firstly, (2) is used to prove the identity

$$\frac{1}{m!}\Big(\frac{d^m\Omega_n}{dz^m}\Big)(1) = \binom{n}{m}\omega_m^{(m)}$$

and, consequently, to write

$$\Omega_n(1+z) = \sum_{h \geq 0} \binom{n}{h} \omega_h^{(h)} z^h. \tag{4}$$

At this stage, de Finetti defines the characteristic function of the frequency[2] $(\sum_{i=1}^{N} \mathbb{I}_{E_k}/N)$

$$t \mapsto \Psi_N(t/N) := \Omega_N(e^{it/N}) \qquad (t \in \mathbb{R})$$

to be the *characteristic function of the* (finite) *class* $\{E_1, \ldots, E_N\}$ *of exchangeable events.* Clearly, such a function characterizes the probability distribution of the random vector $(\mathbb{I}_{E_1}, \mathbb{I}_{E_2}, \ldots, \mathbb{I}_{E_N})$. Notice that this distribution is also determined by the sole knowledge of the probabilities $\omega_h^{(h)}$, $h = 0, 1, \ldots, N$, with $\omega_0^{(0)} = 1$. To see this, combine (3) with (4).

From a practical viewpoint, the following proposition – that the Author states in Section 35 of the *Memoria* – may be useful.

PROPOSITION 1. *Any sequence* $(\tilde{\omega}_h^{(N)})_{h=0,\ldots,N}$, *satisfying*

$$\tilde{\omega}_h^{(N)} \geq 0 \quad (h = 0, \ldots, N) \quad \textit{and} \quad \sum_{h=0}^{N} \tilde{\omega}_h^{(N)} = 1,$$

generates a unique exchangeable law, for the class of events $\{E_1, \ldots, E_N\}$, *according to which the probability that a random phenomenon comes true* k *times in* n *trials* $(1 \leq n \leq N, k = 0, \ldots, m)$ *is given by*

$$\sum_{h=k}^{N-n+k} \tilde{\omega}_h^{(N)} \frac{\binom{h}{k}\binom{N-h}{n-k}}{\binom{N}{n}}. \tag{5}$$

Indeed, consider the partition defined by

$$A_h := \{\sum_{k=0}^{N} \mathbb{I}_{E_k} = h\} \qquad h = 0, 1, \ldots, N$$

in a probability space such that $\tilde{\omega}_h^{(N)}$ is the probability of A_h. If the event A_h occurs, then h white balls along with $(N-h)$ black balls are placed into an urn. Now, consider an individual who just assesses the quantities $\tilde{\omega}_h^{(N)}$ as the

[2] Throughout the paper, the term frequency is used to designate what other authors call relative frequency.

probabilities for the events A_h and who randomly draws n balls from the urn ($n \leq N$), without replacement. So, if he sees the $N(N-1)\dots(N-n+1)$ possible outcomes as equally probable, whatever n may be, then the probability that the sample contains exactly k white balls is given by (5). In other words, any N–exchangeable $\{0,1\}$ sequence is a mixture of hypergeometric N–sequences.

After establishing these basic elementary facts, de Finetti moves on to the analysis of infinite sequences of exchangeable events. Such analysis is focused on the study of the pointwise limit of the characteristic function $\Psi_N(t/N)$, as $N \to +\infty$. As a matter of fact, in all later writings on exchangeability, he will consider a different approach, based on a law of large numbers for exchangeable sequences. As already mentioned, he adopter this approach following a suggestion of Alexander Khinchin (1984–1969), he met on the occasion of the *Congress* of Bologna. See [12], [20] and [21].

2.2 – Representation theorem

Given an infinite sequence $(E_n)_{n\geq 1}$ of exchangeable events, consistently with the previous notation define $\omega_h^{(h)}$ to be the probability of $E_1 \cap \dots \cap E_h$, for $h = 1, 2, \dots$, and set

$$\Omega(1+z) := \sum_{h\geq 0} \omega_h^{(h)} \frac{z^h}{h!} \qquad (z \in \mathbb{C}).$$

In Section 6 of the *Memoria* de Finetti proves the following preliminary:

Proposition 2. *For any strictly positive a and ϵ there is an integer $N_1 = N_1(a,\epsilon)$ such that*

$$\sup_{|z|\leq a} |\Omega(1+z) - \Omega_n(1+z/n)| \leq \epsilon \qquad (n \geq N_1).$$

Then, he uses this fact to prove a more important statement concerning the limiting behavior of the characteristic function $\Psi_n(t/n)$, as $n \to +\infty$:

Proposition 3. *For every $\tau > 0$ and $\epsilon > 0$, there is $N_2 = N_2(\epsilon,\tau)$ such that*

$$\sup_{|t|\leq \tau} |\Psi(t) - \Psi_n(t/n)| \leq \epsilon$$

holds true for every $n \geq N_2$ and

$$\Psi(t) := \Omega(1+it) = \sum_{h\geq 0} \omega_h^{(h)} \frac{(it)^h}{h!} \qquad (t \in \mathbb{R}).$$

It is important to note that Proposition 2 and Proposition 3 are valid uniformly with respect to Ω and Ψ, respectively. In other words, given ϵ, a and τ, N_1 and N_2 do not depend on Ω and Ψ, respectively. See next Subsection 2.5 for a different situation apropos of the connection between frequency and predictive distribution.

So, *if one assumes that* Ψ *is a characteristic function* (see the next subsection), then the corresponding random variable must take values in $[0,1]$, with probability one. Moreover, if Φ is the corresponding probability distribution function, since Ψ can be extended as an entire function, one gets

$$\Psi(t) = \int_{[0,1]} e^{it\xi} d\Phi(\xi), \qquad \omega_h^{(h)} = \int_{[0,1]} \xi^h d\Phi(\xi) \quad (h = 0, 1, \dots).$$

This, in turn, combined with (4), gives

$$\Omega_n(1+z) = \int_{[0,1]} (1 + z\xi)^n d\Phi(\xi)$$

and

$$\begin{aligned} \sum_{h\geq 0} \omega_h^{(n)} z^h = \Omega_n(z) &= \int_{[0,1]} (1 - \xi + z\xi)^n d\Phi(\xi) \\ &= \sum_{h\geq 0} \binom{n}{h} z^h \int_{[0,1]} \xi^h (1-\xi)^{n-k} d\Phi(\xi). \end{aligned}$$

This encompasses the celebrated de Finetti's *representation theorem*, viz.:

PROPOSITION 4. *The events* $(E_n)_{n\geq 1}$ *are exchangeable if and only if there is a probability distribution function* Φ *supported by* $[0,1]$ *such that the probability of* $\{\mathbb{I}_{E_1} = x_1, \dots, \mathbb{I}_{E_n} = x_n\}$ *is given by*

$$\int_{[0,1]} \xi^{\sigma_n} (1-\xi)^{n-\sigma_n} d\Phi(\xi)$$

for every $(x_1, \dots, x_n)$ *in* $\{0,1\}^n$ *for which* $x_1 + \dots + x_n = \sigma_n$*, and for every* $n = 1, 2, \dots$. *Moreover,* Φ *is the limit* (*in the sense of weak convergence of probability distributions*) *of the probability distribution function* Φ_n *of the frequency of success in the first* n *trials, as* $n \to +\infty$.

2.3 – Important remark

The previous argument is based on the presumption that the limit, Ψ, of Ψ_n is a characteristic function. Nowadays, the validity of such an assertion is proved in any good probability textbook. On the contrary, the reference books at de Finetti's disposal – [1] and [22] – although they were superb, they contained the form of the continuity theorem according to which "if Ψ_n converges to a *characteristic function*, uniformly on any compact interval, then ...". Clearly, the argument in Subsection 2.2, apart from the fact that $t \mapsto \Omega(1+it)$ is the limit – uniform on any compact interval – of a sequence of characteristic functions, does not give further indications about the fact that the limit is a characteristic function. So, to complete the proof of the representation theorem, de Finetti was obliged to check whether the above–mentioned limiting condition was enough to assert that $t \mapsto \Omega(1+it)$ was a characteristic function. He deferred the solution of the problem to the Appendix of the *Memoria*, where he proved the desired completion of the continuity theorem – perhaps for the first time – consistently with the fact that he was dealing with *finitely* (i.e., not necessarily *completely*) *additive* distributions of general real–valued random variables. In point of fact, he explicitly assumes that the sequence of distributions corresponding to $(\Psi_n)_{n\geq 1}$ is tight.

2.4 – Strong law of large numbers

In the following Sections 11 and 12, de Finetti deals with the extension of Cantelli's strong law for frequencies of Bernoulli trials to frequencies of more general exchangeable trials. Define the random frequency $\bar{f}_n$ of success in the first n trials of a random phenomenon characterized by an infinite sequence $(E_n)_{n\geq 1}$ of exchangeable events,

$$\bar{f}_n := \frac{1}{n}\sum_{k=1}^{n} \mathbb{I}_{E_k},$$

and consider the sequence $(\bar{f}_n)_{n\geq 1}$. The main result de Finetti achieves apropos of the latter sequence is *a mutual* form of the strong law of large numbers for $(\bar{f}_n)_{n\geq 1}$ that, consistently with the admissibility of simply additive probability distributions, he states correctly in the following "finitary" style.

Proposition 5. *Given strictly positive numbers ϵ and θ, there is a positive integer $N := N(\epsilon, \theta)$ such that the probability of the event*

$$\bigcap_{j=1}^{k} \{|\bar{f}_n - \bar{f}_{n+j}| \leq \epsilon\}$$

turns out to be uniformly (with respect to $k = 1, 2, \dots$) greater that $1 - \theta$, whenever $n \geq N$.

Apparently de Finetti was aware of the fact that, in a framework of completely additive probability distributions on subsets of a sample space, the above proposition holds with $k = +\infty$, and that one can assert the existence of a random number f^*, with probability distribution function Φ, that can be viewed as the almost sure limit of $(\bar{f}_n)_{n\geq 1}$. But he had at least three good reasons, from his viewpoint, to be uninterested in the "strong" formulation of his strong law of large numbers. These reasons, briefly mentioned in many points of the *Memoria*, are discussed in a more systematic way in a few contemporaneous de Finetti's papers such as [7], [8] and [9]. It is worth recalling them here, in a three-point summary. (i) Logically speaking, it is unjustifiable to speak of an *infinite* sequence of trials of a random phenomenon: The number of the trials could be arbitrarily great but, in any case, finite. (ii) Without the assumption of complete additivity and with no reference to a sample space, there is no possibility of deducing the existence of a limiting random quantity from the sole mutual convergence of a given sequence. (iii) de Finetti deduces the whole theory of probability from a very natural condition having an obvious meaning – the so-called *condition of coherence* – and shows that *complete* additivity is not necessary for a quantitative measure of probability to be coherent. See [10].

The strong law of large numbers for the frequency of success in a sequence of exchangeable events represents the last issue dealt with in Chapter 1 of the *Memoria*. Chapter 2 contains the definitions of some operators on the set of all characteristic functions, with the intention of providing a rigorous, systematic presentation, in Chapter 3, of asymptotic properties of the posterior distribution and of the merging of the predictive distribution with the frequency of success in past trials of a given random phenomenon. In view of the purely instrumental function of Chapter 2, here we jump to the more important Chapter 3.

2.5 – Probability and experience: Posterior and predictive distributions

At the time of the draft of the *Memoria*, de Finetti was unfamiliar with techniques of statistical inference, and it's amazing how he was, nevertheless, able at picking out the essence of the inductive reasoning and the tools to deal with it, from a coherent mathematical standpoint. In his view of these subjects, exchangeability is a means to study and understand the role played by the knowledge of data, gathered from experience, with regards to the evaluation of probability. In particular, he aims at clarifying how exchangeability can be employed to provide with a basis the common belief that prevision of new facts rests on the analogy with past observed facts. In the case of a random phenomenon, this belief leads to assume, although with caution, past frequency as an approximate value for probability. So, in Section 27, de Finetti provides a new rigorous description of the asymptotic behavior of the posterior distribution(3), and makes use of this statement to show the merging of the predictive

(3) In point of fact, he was unaware of [26], where a strictly related problem is studied.

distribution with frequency in past trials. Apropos of the former, he considers *infinite sequences* of exchangeable trials of a random phenomenon, generating convergent sequences of frequencies. More precisely,

PROPOSITION 6. *Let* θ_0 *be any point in the intersection of* $(0,1)$ *with the support of the distribution function,* Φ*, of the random phenomenon. Then the posterior distribution, given* $\{\bar{f}_n = \frac{\sigma_n}{n}\}$*, converges weakly to the point mass* δ_{θ_0}*, whenever* $\sigma_n/n \to \theta_0$ *as* $n \to +\infty$*, i.e.*

$$\lim_{n\to+\infty} \frac{\int_{\theta_0-\epsilon}^{\theta_0+\epsilon} \theta^{\sigma_n}(1-\theta)^{n-\sigma_n} d\Phi(\theta)}{\int_{[0,1]} \theta^{\sigma_n}(1-\theta)^{n-\sigma_n} d\Phi(\theta)} = 1 \qquad (\epsilon > 0).$$

Whence, as for the conditional probability of $\{E_{n+k}\}$ *given* $\{\bar{f}_n = \frac{\sigma_n}{n}\}$*, viz.*

$$\frac{\int_{[0,1]} \theta^{1+\sigma_n}(1-\theta)^{n-\sigma_n} d\Phi(\theta)}{\int_{[0,1]} \theta^{\sigma_n}(1-\theta)^{n-\sigma_n} d\Phi(\theta)},$$

one obtains that, for any $\epsilon > 0$*, there is* $N = N(\epsilon, \Phi)$*, such that*

$$\left| \frac{\int_{[0,1]} \theta^{1+\sigma_n}(1-\theta)^{n-\sigma_n} d\Phi(\theta)}{\int_{[0,1]} \theta^{\sigma_n}(1-\theta)^{n-\sigma_n} d\Phi(\theta)} - \frac{\sigma_n}{n} \right| \leq \epsilon$$

holds true for every $n \geq N$*.*

In Section 28, de Finetti explains the "relative" value of this proposition. Indeed, in view of the dependence of N on Φ, it does not allow a quantitative statement about the approximation of frequency to probability, independently of a complete *a priori* knowledge of the characteristic function of the random phenomenon.

Chapter 3 ends with a brief mention of the use of posterior distribution in the problem of hypothesis–testing: the sole explicit hint to a statistical technique, contained in the *Memoria*.

2.6 – Classes of exchangeable events and extension of exchangeability

The main issue dealt with in the last chapter (Chapter 4, including Sections 31–36) is extendibility of exchangeability, from a finite sequence to a "longer" sequence of events. The problem can be formulated in the following terms: *Given positive integers* n *and* k*, establish conditions on the characteristic function of a random phenomenon of* n *exchangeable events in order that they may constitute the initial* n*–segment of a random phenomenon of* $(n+k)$ *exchangeable events.*

To solve this problem, de Finetti starts from (2), viewed as a first–order linear differential equation in the dependent variable Ω_{n+1}. Since the one–parameter family of solutions of this equation is

$$\Omega_{n+1}(z) = (1-z)^{n+1}\{(n+1)\int_0^z \Omega_n(x)(1-x)^{-(n+2)}dx + c\}, \tag{6}$$

then (6) can be combined with Proposition 1 to obtain

PROPOSITION 7. *$t \mapsto \Omega_n(e^{it/n})$ is the characteristic function of the initial n-segment of a sequence of* $(n+1)$ *exchangeable events if and only if the constant c in* (6) *can be determined in such a way that all the coefficients of the polynomial* (*of degree* $(n+1)$), *defined by the right–hand side of* (6), *are nonnegative.*

Analogously, to solve the problem for some $k > 1$, one can start from (6) with $(n+2)$ in the place of $(n+1)$, consider it as an equation in the dependent variable Ω_{n+2} and, finally, substitute Ω_{n+1} with its expression in the right–hand side of (6). So, by an obvious recursive argument, de Finetti states that Ω_{n+k} can be written as

$$\Omega_{n+k}(z) = F(z) + C_1(1-z)^{n+1} + \cdots + C_k(1-z)^{n+k} \tag{7}$$

F being a polynomial, whose coefficients are completely determined by Ω_n. Then:

PROPOSITION 8. *$t \mapsto \Omega_n(e^{it/n})$ is the characteristic function of the initial n–segment of a sequence of* $(n+k)$ *exchangeable events if and only if the constants $C_1, \ldots, C_k$ can be determined in such a way that all the coefficients of the polynomial* (*of degree* $(n+k)$), *defined by the right–hand side of* (7), *are nonnegative.*

Forty years later, de Finetti came back to the problem from a new standpoint, of a geometrical nature (see [15]), followed by some Authors such as [3], [4], [17] and [27].

In Section 4 of the present paper, we will resume the original analytical de Finetti's argument, by providing an explicit form for F in (7). Our goal is to reformulate the necessary and sufficient condition in Proposition 8 in the guise of a system of linear inequalities.

De Finetti gives a complete solution of the extendibility problem when $k = +\infty$, i.e.: *To establish conditions on $t \mapsto \Omega_n(e^{it/n})$ in order that it can be viewed as characteristic function of the first n trial of a random phenomenon of infinite exchangeable events.* Resting on the representation (see Subsection 2.2) according to which $\omega_h^{(h)}$ is the h–moment of the probability distribution function

Φ of the random phenomenon, via the Hamburger solution of the problem of moments (see, e.g., [25]), de Finetti was able to state:

Proposition 9. *In order that* $t \mapsto \Omega_n(e^{it/n})$ *may be the characteristic function of the initial* n*–segment of an infinite sequence of exchangeable events it is necessary and sufficient that all the roots of a distinguished polynomial, depending on* n*, belong to the closed interval* $[0,1]$*. The polynomial* (*in* ξ) *is*

$$Det \begin{pmatrix} 1 & \xi & \xi^2 & \dots & \xi^k \\ \omega_0^{(0)} & \omega_1^{(1)} & \omega_2^{(2)} & \dots & \omega_k^{(k)} \\ \dots & \dots & \dots & \dots & \dots \\ \omega_{k-1}^{(k-1)} & \omega_k^{(k)} & \omega_{k+1}^{(k+1)} & \dots & \omega_{2k-1}^{(2k-1)} \end{pmatrix}$$

if $n = 2k-1$*, while it is*

$$Det \begin{pmatrix} 1 & \xi & \xi^2 & \dots & \xi^k \\ \omega_1^{(1)} & \omega_2^{(2)} & \omega_3^{(3)} & \dots & \omega_{k+1}^{(k+1)} \\ \dots & \dots & \dots & \dots & \dots \\ \omega_k^{(k)} & \omega_{k+1}^{(k+1)} & \omega_{k+2}^{(k+2)} & \dots & \omega_{2k}^{(2k)} \end{pmatrix}$$

if $n = 2k$.

3 – Haag's contribution to exchangeability

To our knowledge, Haag was the first Author to study sequences of exchangeable events. He publicized his conclusions during the *International Mathematical Congress* held in Toronto, August, 1924. His communication appeared in Vol. 1 of the *Proceedings*, published in 1928, the very same year of the already mentioned Bologna *Congress*. See [18]. It is highly likely that de Finetti was in the dark about the Haag contribution until the 1950s, when Edwin Hewitt and Leonard J. Savage mentioned it in a famous paper about exchangeability. See [19].

It is convenient to pause here and consider what Haag really did. In the first six brief sections, he deals with finite sequences of exchangeable events and furnishes a detailed account of the expressions of the $\omega_k^{(n)}$ both in terms of $\omega_h^{(h)}$ and in terms of $\omega_0^{(h)}$, for $h = 1, 2, \dots, n$. In Section 7, Haag attains an early version of the representation theorem, but via a rather incomplete argument. He considers a sequence of exchangeable trials with a frequency of success σ_n/n converging to x as $n \to +\infty$. By the way, Haag does not hint at any form of law of large numbers, so that the reader is not able to judge whether the convergence assumption expresses an extraordinary or, instead, a common fact. By resorting to the Stirling formula, and assuming that $f(x)dx$ provides, for some continuous function f defined on $(0,1)$, an asymptotic value for the probability that σ_n/n belongs to $(x, x+dx)$, as $n \to +\infty$, he shows that

$$\sqrt{2\pi x(1-x)}\frac{(x^x(1-x)^{1-x})^n}{\sqrt{n}}f(x)$$

represents an approximate value – for great values of n – of $\omega_{\sigma_n}^{(n)}/\binom{n}{\sigma_n}$. At this stage, by means of a heuristic argument based on formal elementary computations, he concludes that

$$\binom{p+q}{p}x^p(1-x)^q\frac{1}{n}f(x) \tag{8}$$

gives an approximate value for the probability of the event "The limiting frequency belongs to $(x, x+1/n)$ and, simultaneously, one gets p successes in $n = p+q$ trials". So, the probability of p successes in $(p+q)$ trials can be represented as limit (as $n \to +\infty$) of a sum of terms like (8), i.e.

$$\binom{p+q}{p}\int_0^1 x^p(1-x)^q f(x)dx.$$

The *assumption* that the frequency converges to a random variable, weakens the validity of the Haag argument, and emphasizes the difference between his standpoint and de Finetti's stance. Indeed, de Finetti reckons that convergence of frequency must be proved, whilst it is evident that Haag is assuming the validity of some type of empirical law which postulates convergence of frequency. So, while de Finetti introduces exchangeability to explain the role of frequency in evaluating probability – within the framework of a rigorously subjectivisitc or, on depending on taste, axiomatic conception – Haag misses out on these fundamental aspects. Moreover, while de Finetti shows to have an extraordinarily advanced command of the right mathematical apparatus to deal with probabilistic problems, Haag does not go beyond the use of the elementary combinatorial calculus. In point of fact, the final part of his paper, intitled *Applications*, includes a review of classical problems solvable by means of elementary combinatorics.

4 – Some new developments on extendibility

As anticipated in Subsection 2.6, in the last part of this paper we follow de Finetti's ideas, explained in that very same subsection, to obtain new necessary and sufficient conditions for extendibility of a given finite–dimensional exchangeable distribution. These conditions are of an algebraic nature, differently from the above–mentioned conditions derived in the frame of a geometrical approach. Taking (7) as a starting point, one first determines an explicit form for F, i.e.

$$
\begin{aligned}
F(z) = F(z; n, k) = z^k \frac{(n+k)!}{n!} \int_{(0,1)^k} \Big(\prod_{j=1}^{k} t_j^{j-1} \Big) \{1 - z(1 - t_1 \cdots t_k)\}^n \cdot \\
\cdot\, \Omega_n \left(\frac{t_1 \cdots t_k z}{1 - z(1 - t_1 \cdots t_k)} \right) dt_1 \dots dt_k \qquad (k = 1, 2, \dots).
\end{aligned} \tag{9}
$$

To prove the validity of this representation, first note that (9) with $k = 1$ is consistent with (6). Then, to complete the proof, use (6), with n replaced by $(n + k)$, and proceed by mathematical induction with respect to k.

Now, substitute expression (3) into (9) to obtain

$$
\begin{aligned}
F(z) &= \frac{(n+k)!}{n!} \sum_{l=0}^{n} \omega_l^{(n)} z^{l+k} \int_{(0,1)^k} (t_1 \cdots t_k)^l \cdot \\
&\qquad\qquad (1 - z(1 - t_1 \cdots t_k))^{n-l} t_2 t_3^2 \dots t_k^{k-1} dt_1 \dots dt_k \\
&= \frac{(n+k)!}{n!} \sum_{l=0}^{n} \omega_l^{(n)} z^{l+k} \frac{1}{\Gamma(k)} \int_0^1 x^l (1-x)^{k-1} \{1 - z(1-x)\}^{n-l} dx \\
&\qquad \text{(see, for example, 3.3.5.11 in [23])} \\
&= \frac{(n+k)!}{n!(k-1)!} \sum_{l=0}^{n} \omega_l^{(n)} \sum_{h=0}^{n-l} \binom{n-l}{h} z^{h+l+k} (-1)^h B(l+1, k+h)
\end{aligned}
$$

with $B(\alpha, \beta) := \int_0^1 x^{\alpha-1}(1-x)^{\beta-1} dx$. Whence, from (6),

$$
\begin{aligned}
\Omega_{n+k}(z) = \sum_{i=0}^{k-1} (-1)^i z^i \sum_{j=1 \vee (i-1)} \binom{n+j}{i} C_j + \sum_{i=k}^{n+k} (-1)^i z^i \Big\{ \sum_{j=1 \vee (i-n)} \binom{n+j}{i} C_j \\
+ \frac{(n+k)!}{n!(k-1)!} (-1)^k \sum_{l=0}^{i-k} \omega_l^{(n)} (-1)^l \binom{n-l}{n-i+k} B(l+1, i-l) \Big\}.
\end{aligned}
$$

Then, Proposition 8 can be restated as

Proposition 10. *$t \mapsto \Omega_n(e^{it/n}) := \sum_{h=0}^{n} \omega_h^{(n)} e^{iht/n}$ is the characteristic function of the initial n–segment of a sequence of $(n + k)$ exchangeable events if* *and only if the constants $C_1, \ldots, C_k$ can be determined in such a way that*

$$
\begin{aligned}
&0 \le \rho_i := (-1)^i \sum_{j=1\vee(i-n)}^{k} \binom{n+j}{i} C_j \qquad i = 0, \ldots, k-1 \\
(10)\qquad &0 \le \rho_i := (-1)^i \Big\{ \sum_{j=1\vee(i-n)} \binom{n+j}{i} C_j \\
&\qquad + \frac{(n+k)!}{n!(k-1)!} (-1)^k \sum_{l=0}^{i-k} \omega_l^{(n)} (-1)^l \binom{n-l}{n-i+k} B(l+1, i-l) \Big\} \\
&\qquad i = k, \ldots, n+k.
\end{aligned}
$$

Moreover, if this system of linear inequalities is consistent, then for each of the solutions $(C_1, \ldots, C_k)$, the vector $(\rho_0, \ldots, \rho_{n+k})$ represents an exchangeable assessment for $(\omega_0^{(n+k)}, \ldots, \omega_{n+k}^{(n+k)})$, consistent with the initial segment $(\omega_0^{(n)}, \ldots, \omega_n^{(n)})$.

The research of conditions for consistency of systems like (10) originated a wealth of literature on the subject. Here, we propose a solution derived from [2]. In matrix form, (10) becomes $Ax \le b$ where

$$
A = [a_{ij}]_{1\le i\le n+k+1, 1\le j\le k}, \quad b' = (b_1, \ldots, b_{n+k+1}), \quad x' = (C_1, \ldots, C_k),
$$

with

$$
\begin{aligned}
&a_{ij} := (-1)^i \binom{n+j}{i-1}, \qquad b_1 = 0, \ldots, b_k = 0, \\
&b_i = (-1)^{i+k-1} \frac{(n+k)!}{n!(k-1)!} \sum_{l=0}^{i-1-k} (-1)^l \omega_l^{(n)} \binom{n-l}{n-i+k+1} B(l+1, i-1-l) \\
&\qquad i = k+1, \ldots, n+k+1.
\end{aligned}
$$

Since, as it is easy to show, the rank of A is k, Theorem 3 in [2] yields

PROPOSITION 11. *$\Omega_n(e^{it/n}) := \sum_{h=0}^{n} \omega_h^{(n)} e^{iht/n}$ is the characteristic function of the initial n–segment of a sequence of $(n+k)$ exchangeable events if, and only if, there exist $1 \le i_1 < i_2 < \cdots < i_k \le n+k+1$ such that*

$$
\begin{pmatrix}
a_{i_1 1} & \ldots & a_{i_1 k} \\
\ldots & \ldots & \ldots \\
a_{i_k 1} & \ldots & a_{i_k k}
\end{pmatrix}
$$

has a nonvanishing determinant Δ, *and*

$$\frac{1}{\Delta}\, det \begin{pmatrix} a_{i_1 1} & \dots & a_{i_1 k} & b_{i_1} \\ \dots & \dots & \dots & \dots \\ a_{i_k 1} & \dots & a_{i_k k} & b_{i_k} \\ a_{j1} & \dots & a_{jk} & b_j \end{pmatrix}$$

turns out to be nonnegative for every $j = 1, \dots, n+k+1$.

In the particular case of $k = 1$, *this necessary and sufficient condition reduces to require that* $\{\omega_h^{(n)} : h = 0, \dots, n\}$ *satisfies*

$$\max\{\beta_i : \textit{for any even integer} \le n+1\} \le \min\{\beta_i : \textit{for any odd integer} \le n+1\}$$

where

$$\beta_i = \sum_{l=0}^{i-1} (-1)^l B(l+1, n-l+1)\omega_l^{(n)} \qquad (i = 1, \dots, n+1).$$

In fact, this result could be obtained by direct inspection of (10) with $k = 1$.

To conclude, let us remark that Proposition 11 is susceptible of interesting geometrical interpretations that one can deduce directly from the above–mentioned Cernikov paper. It would be interesting to compare them with the geometrical arguments in [15] and developed by other Authors already mentioned in Subsection 2.6.

Acknowledgements

The authors are very grateful to the organizing committee of the *Bruno de Finetti Centenary conference* (Rome, November 2006) and to referees for their time, comments and suggestions.

REFERENCES

[1] G. Castelnuovo: *Calcolo delle Probabilità*, Vol.1. 2nd ed. Zanichelli, Bologna, 1925

[2] S. N. Cernikov: *Systems of linear inequalities*, Amer. Math. Soc. Transl., **26** (1963), 11-86.

[3] L. Crisma, L.: *Alcune valutazioni quantitative interessanti la proseguibilità di processi aleatori scambiabili*, Rend. Ist. Mat. Univ. Trieste, **3** (1971), 96-124.

[4] L. Crisma: *Quantitative analysis of exchangeability in alternative processes.*, pp. 207–216 In *Exchangeability in probability and statistics*, G.Koch F.Spizzichino (eds.), North-Holland, Amsterdam, 1982.

[5] B. de Finetti: *Sulle funzioni a incremento aleatorio*, Atti Reale Accademia Nazionale dei Lincei. Rend. Serie VI, **10** (1929), 163-168.

[6] B. de Finetti: *Funzione caratteristica di un fenomeno aleatorio*, Atti Reale Accademia Nazionale dei Lincei. Mem. Serie VI., **4** (1930), 86-133.

[7] B. de Finetti: *Sui passaggi al limite nel calcolo delle probabilità*, Rendiconti Reale Istituto Lombardo di Scienze e Lettere, **63** (1930), 155-166.

[8] B. de Finetti: *A proposito dell'estensione del teorema delle probabilità totali alle classi numerabili*, Rendiconti Reale Istituto Lombardo di Scienze e Lettere, **63** (1930), 901-905.

[9] B. de Finetti: *Ancora sull'estensione alle classi numerabili del teorema delle probabilità totali*, Rendiconti del Reale Istituto Lombardo di Scienze e Lettere, **63** (1930), 1063-1069.

[10] B. de Finetti: *Sul significato soggettivo della probabilità*, Fundamenta Mathematicae **17** (1931), 298-329.

[11] B. de Finetti: *Funzione caratterisitca di un fenomeno aleatorio*, Atti del Congresso Internazionale dei Matematici, (Bologna, 3–10 Settembre, 1928), Zanichelli, Bologna, pp. 179-190.

[12] B. de Finetti: *La prévision : ses lois logiques, ses sources subjectives*, Ann. Inst. H. Poincaré, **7** (1937), 1-68.

[13] B. de Finetti: *Resoconto critico del colloquio di Ginevra intorno alla teoria delle probabilità*, Giorn. Istit. Ital. Attuari, **9** (1938), 1-40.

[14] B. de Finetti: *Compte rendu critique du colloque de Génève sur la théorie des probabilités*, In *Actualités Scientifiques et Industrielles* Vol. 766 Hermann, Paris, 1939.

[15] B. de Finetti: *Sulla proseguibilità di processi aleatori scambiabili*, Rend. Ist. Mat. Univ. Trieste, **1** (1969), 53–67.

[16] B. de Finetti: *La probabilità: guardarsi dalle contraffazioni!*, Scientia, **111** (1976), 255-303.

[17] P. Diaconis: *Finite forms of de Finetti's theorem on exchangeability*, Synthese, **36** (1977), 271-281.

[18] J. Haag: *Sur un probléme général de probabilités et ses diverses applications*, In *Proceedings of the International Congress of Mathematicians* (Toronto, August 11-26, 1924) **1** (1928), 659–674, Univ. Toronto Press.

[19] E. Hewitt – L. J. Savage: *Symmetric measures on Cartesian products*, Trans. Amer. Math. Soc. **80** (1955), 470–501.

[20] A. Khinchin: *Sur les classes d'événements équivalentes*, Mat.Sb. **39** (1932), 40-43.

[21] A. Khinchin: *Remarques sur les suites d'événements obéssants á la loi des grands nombres*, Mat.Sb., **39** (1932), 115-119.

[22] P. Lévy: *Calcul des Probabilités*, Gauthier–Villars, Paris, 1925.

[23] A. P. Prudnikov – Y. A. Brychkov – O. I. Marichev: *Integrals and series.*, vol. 1 Gordon & Breach Science Publishers, New York, 1926.

[24] E. Regazzini: *de Finetti's contribution to the theory of random functions*, *International Journal of Approximate Reasoning* **47** (2008), 387–393.

[25] J. A. Shohat – J. D. Tamarkin: *The Problem of Moments*, American Mathematical Society, Providence, R.I., 1943

[26] R. von Mises: *Fundamentalsaetze der wahrscheinlick keitsrechnung*, Math. Zeit., **4** (1919), 1-97.

[27] G. R. Wood: *Binomial mixtures and finite exchangeability*, Ann. Probab., **20** (1992), 1167-1173.

Lavoro pervenuto alla redazione il 20 dicemnvre 2007
ed accettato per la pubblicazione il 27 dicembre 2007.
Bozze licenziate il 9 aprile 2008

Indirizzo degli Autori:

Eugenio Regazzini – Dipartimento di Matematica – via Ferrata 1 – 27100 Pavia, Italy
E-mail: eugenio.regazzini@unipv.it

Federico Bassetti – Dipartimento di Matematica – via Ferrata 1 – 27100 Pavia, Italy
E-mail: federico.bassetti@unipv.it

Supported in part by MIUR grant 2006/134526.

Rendiconti di Matematica, Serie VII
Volume 28, Roma (2008), 19–32

Some highlights from the theory of multivariate symmetries

OLAV KALLENBERG

Abstract: *We explain how invariance in distribution under separate or joint contractions, permutations, or rotations can be defined in a natural way for d-dimensional arrays of random variables. In each case, the distribution is characterized by a general representation formula, often easy to state but surprisingly complicated to prove. Comparing the representations in the first two cases, one sees that an array on a tetrahedral index set is contractable iff it admits an extension to a jointly exchangeable array on the full rectangular index set.*

Multivariate rotatability is defined most naturally for continuous linear random functionals on tensor products of Hilbert spaces. Here the simplest examples are the multiple Wiener–Itô integrals, which also form the basic building blocks of the general representations. The rotatable theory can be used to derive similar representations for separately or jointly exchangeable or contractable random sheets. The present paper provides a non-technical survey of the mentioned results, the complete proofs being available elsewhere. We conclude with a list of open problems.

1 – Basic symmetries and classical results

Many basic ideas in the area of probabilistic symmetries can be traced back to the pioneering work of Bruno de Finetti. After establishing, in 1930–37, his celebrated representation theorem for exchangeable sequences, he proposed in de Finetti (1938) the study of *partial exchangeability* of a random sequence, in the sense of invariance in distribution under a proper subgroup of permutations

---

Key Words and Phrases: *Contractable – Exchangeable and rotatable random arrays – Functionals and sheets – Multiple Wiener-Itô integrals.*
A.M.S. Classification: 60G09

of the elements. (A permutation on an infinite set is defined as a bijective map.) For a basic example, we may arrange the elements in a doubly infinite array and require invariance in distribution under permutations of all rows and all columns. This leads to the notion of *row-column* or *separate exchangeability*, considered below.

Many other probabilistic symmetries of interest can be described in terms of higher-dimensional arrays, processes, measures, or functionals. Their study leads to an extensive theory, whose current state is summarized in the last three chapters of the monograph KALLENBERG (2005)(1). Our present aim is to give an informal introduction to some basic notions and results in the area. No novelty is claimed, apart from some open problems listed at the end of the paper. Before introducing the multivariate symmetries, we need to consider the one-dimensional case. For infinite sequences $X = (X_j)$ of random variables, we have the following basic symmetries, listed in the order of increasing strength. (Here X has the property on the left iff its distribution is invariant under the transformations on the right.)

stationary	*shifts*
contractable	*contractions*
exchangeable	*permutations*
rotatable	*rotations*

Thus, X is *contractable* if all subsequences have the same distribution, *exchangeable* if the joint distribution is invariant under arbitrary permutations, and *rotatable* if the distribution is invariant under any orthogonal transformation applied to finitely many elements. The notion of stationarity is well-known and will not be considered any further in this paper.

Sequences with the last three symmetry properties are characterized by the following classical results. Letting $X = (X_j)$ be an infinite sequence of random variables, we have:

- (DE FINETTI (1930, 1937)): *X is exchangeable iff it is mixed* (*or conditionally*) *i.i.d.*,
- (RYLL-NARDZEWSKI (1957)): *X is contractable iff it is exchangeable, hence mixed i.i.d.*,
- (FREEDMAN (1962)): *X is rotatable iff it is mixed i.i.d. centered Gaussian.*

For processes X on $\mathbb{R}_+$, the notions of contractability, exchangeability, and rotatability are defined in terms of the increments over any set of disjoint intervals of equal length. We may also assume that X is continuous in probability and starts at 0. Then the first two properties are again equivalent, and the three cases are characterized as follows:

(1) Henceforth abbreviated as K(2005).

- (Bühlmann (1960)): *X is exchangeable iff it is a mixture of Lévy processes,*
- (Freedman (1963)): *X is rotatable iff it is a mixture of centered Brownian motions with different rates.*

In particular, we see from the former result that a continuous process X is exchangeable iff it is a mixture of Brownian motions with arbitrary rate and drift coefficients. Thus, in the continuous case, the exchangeable and rotatable processes differ only by a random centering. This observation plays an important role for the analysis of exchangeable and contractable random sheets. Together with Ryll-Nardzewski's theorem, it also signifies a close relationship between the various symmetry properties in our basic hierarchy.

Proceeding to two-dimensional random arrays $X = (X_{ij};\, i,j \ge 1)$ indexed by $\mathbb{N}^2$, we may define the permuted arrays $X \circ (p,q)$ and $X \circ p$ by

$$(X \circ (p,q))_{ij} = X_{p_i,q_j}, \qquad (X \circ p)_{ij} = X_{p_i,p_j}\,,$$

where $p = (p_i)$ and $q = (q_j)$ are permutations on $\mathbb{N}$. Then X is said to be *separately exchangeable* if $X \circ (p,q) \stackrel{d}{=} X$ for all permutations p and q on $\mathbb{N}$ and *jointly exchangeable* if $X \circ p \stackrel{d}{=} X$ for any such permutation p. Note that the latter property is weaker, so that every separately exchangeable array is also jointly exchangeable. The definitions in higher dimensions are similar.

The contractable case is similar. Thus, X is said to be *separately contractable* if $X \circ (p,q) \stackrel{d}{=} X$ for all subsequences p and q of $\mathbb{N}$ and *jointly contractable* if $X \circ p \stackrel{d}{=} X$ for any such subsequence p. However, only the joint version is of interest, since the separate notions of exchangeability and contractability are equivalent by Ryll-Nardzewski's theorem above (applied to random elements in $\mathbb{R}^\infty$).

To define rotatability in higher dimensions, consider arrays $U = (U_{ij})$ such that, for some $n \in \mathbb{N}$, the restriction to the square $\{1,\ldots,n\}^2$ is orthogonal and otherwise $U_{ij} = \delta_{ij}$. For any such arrays U and V, we may define the array $X \circ (U \otimes V)$ by

$$(X \circ (U \otimes V))_{ij} = \sum\nolimits_{h,k} X_{hk}\, U_{hi}\, V_{kj}, \quad i,j \in \mathbb{N}\,,$$

and put $U^{\otimes 2} = U \otimes U$. Then X is said to be *separately rotatable* if $X \circ (U \otimes V) \stackrel{d}{=} X$ for all orthogonal arrays U and V as above and *jointly rotatable* if $X \circ U^{\otimes 2} \stackrel{d}{=} X$ for any such array U. Even these properties extend immediately to arbitrary dimensions.

The natural index set of a jointly exchangeable array X is not $\mathbb{N}^2$ but rather $\mathbb{N}^{(2)} = \{(i,j) \in \mathbb{N}^2;\, i \ne j\}$. In fact, an array X on $\mathbb{N}^2$ is jointly exchangeable iff the same property holds for the non-diagonal array

$$Y_{ij} = (X_{ij}, X_{ii}), \quad i \ne j\,.$$

(The latter property makes sense, since every permutation on $\mathbb{N}$ induces a joint permutation on $\mathbb{N}^{(2)}$.) Similarly, the natural index set for a d-dimensional, jointly exchangeable array is the set $\mathbb{N}^{(d)}$, consisting of all d-tuples $(i_1,\dots,i_d)$ with distinct entries $i_1,\dots,i_d$.

In the contractable case, we can go even further. Thus, an array X on $\mathbb{N}^2$ is jointly contractable iff the same property holds for the sub-diagonal (or super-diagonal, depending on the geometrical representation) array

$$Z_{ij} = (X_{ij}, X_{ji}, X_{ii}), \quad i < j\,.$$

Similarly, the natural index set for a d-dimensional, jointly contractable array is the *tetrahedral* index set

$$\Delta_d = \{(i_1,\dots,i_d) \in \mathbb{N}^d;\, i_1 < \cdots < i_d\}\,.$$

It is often convenient to identify Δ_d with the class $\widetilde{\mathbb{N}}_d$, consisting of all subsets of $\mathbb{N}$ of cardinality d.

To summarize, we are led to consider exchangeable arrays on $\overline{\mathbb{N}} = \bigcup_d \mathbb{N}^{(d)}$ and contractable arrays on $\widetilde{\mathbb{N}} = \bigcup_d \widetilde{\mathbb{N}}_d$, where the qualification "jointly" is understood. The natural setting for the rotatable case will be discussed later.

2 – Exchangeable and contractable arrays

The aim of this section is to explain how separately or jointly exchangeable or contractable arrays of arbitrary dimension can be characterized by some general functional representations. In order to fully understand those formulas, it is useful to begin with the one-dimensional case. Write $U(0,1)$ for the uniform distribution on $[0,1]$.

- *An infinite random sequence $X = (X_j)$ is contractable (hence exchangeable) iff there exist a measurable function f on $[0,1]^2$ and some i.i.d. $U(0,1)$ random variables α and $\xi_1, \xi_2, \dots$ such that a.s.*

$$X_j = f(\alpha, \xi_j), \quad j \ge 1\,.$$

This is just another way of stating de Finetti's theorem. In particular, we see directly from this formula that the X_j are conditionally i.i.d. given α. This formulation has the disadvantage that the function f is not unique, and further that an independent randomization variable may be needed to construct the associated coding variables α and ξ_j.

For exchangeable arrays of higher dimension, the characterization problem is much harder. Here the first breakthrough came with Aldous' intricate proof (later simplified by Kingman) of the following result. (See also the elementary discussion in Aldous (1985).)

- (Aldous (1981)): *An array $X = (X_{ij})$ on $\mathbb{N}^2$ is separately exchangeable iff there exist a measurable function f on $[0,1]^4$ and some i.i.d. $U(0,1)$ random variables α, ξ_i, η_j, and ζ_{ij} such that a.s.*

$$X_{ij} = f(\alpha, \xi_i, \eta_j, \zeta_{ij}), \quad i, j \geq 1 .$$

Representations of this type can also be deduced from certain results in formal logic, going back to the 1960's. (Here an elementary discussion appears in Hoover (1982).) Combining related methods with the techniques of non-standard analysis, Hoover found some general representations characterizing separately or jointly exchangeable arrays of arbitrary dimension. In the two-dimensional, jointly exchangeable case, his representation reduces to the following:

- (Hoover (1979)[(2)]): *An array $X = (X_{ij})$ on $\mathbb{N}^{(2)}$ is jointly exchangeable iff there exist a measurable function f on $[0,1]^4$ and some i.i.d. $U(0,1)$ random variables α, ξ_i, and $\zeta_{\{i,j\}}$ such that a.s.*

$$X_{ij} = f(\alpha, \xi_i, \xi_j, \zeta_{\{i,j\}}), \quad i \neq j .$$

Note that the representation in the separately exchangeable case follows as an easy corollary. Still deeper is the corresponding representation in the contractable case:

- (K (1992)): *An array $X = (X_{ij})$ on Δ_2 is jointly contractable iff there exist a measurable function f on $[0,1]^4$ and some i.i.d. $U(0,1)$ random variables α, ξ_i, and ζ_{ij} such that a.s.*

$$X_{ij} = f(\alpha, \xi_i, \xi_j, \zeta_{ij}), \quad i < j .$$

Comparing with the result in the jointly exchangeable case, we get the following rather surprising extension theorem:

- (K (1992)): *An array $X = (X_{ij})$ on Δ_2 is jointly contractable iff it can be extended to a jointly exchangeable array on $\mathbb{N}^{(2)}$.*

In fact, the last two results are clearly equivalent, given Hoover's representation in the jointly exchangeable case. No direct proof is known. As already mentioned, the representing function f in the quoted theorems is far from unique.

---

(2) The reason for the earlier date is that Hoover's long and difficult paper, written at the Institute of Advanced Study at Princeton, was never published.

To illustrate the possibilities, we may quote an equivalence criterion in the contractable case:

- (K (1992)): *Two measurable functions f and f' on $[0,1]^2$ can be used to represent the same contractable array X on Δ_2 iff there exist some measurable functions g_0, g_0' on $[0,1]$, g_1, g_1' on $[0,1]^2$, and g_2, g_2' on $[0,1]^4$, each measure preserving in the last argument, such that a.s., for any i.i.d. $U(0,1)$ random variables α, ξ_i, and ζ_{ij},*

$$\begin{aligned} &f(g_0(\alpha), g_1(\alpha,\xi_i), g_1(\alpha,\xi_j), g_2(\alpha,\xi_i,\xi_j,\zeta_{ij})) \\ &\quad = f'(g_0'(\alpha), g_1'(\alpha,\xi_i), g_1'(\alpha,\xi_j), g_2'(\alpha,\xi_i,\xi_j,\zeta_{ij})), \quad i<j\,. \end{aligned}$$

To state the higher-dimensional results in a concise form, we may introduce an array of i.i.d. $U(0,1)$ random variables (or *U-array*) $\xi = (\xi_J)$ indexed by $\widetilde{\mathbb{N}}$, and write

$$\hat{\xi}_J = (\xi_I;\, I \subset J), \quad J \in \widetilde{\mathbb{N}}\,.$$

Similarly, for any $k \in \overline{\mathbb{N}}$, we may form the associated set $\tilde{k} = \{k_1, k_2, \dots\}$ and write

$$\hat{\xi}_k = (\xi_I;\, I \subset \tilde{k}), \quad k \in \overline{\mathbb{N}}\,.$$

To be precise, we also need to specify an order *among* (not *within*) the sets $I \subset \tilde{k}$, which is determined in an obvious way by the order within k of the elements k_j.

Using the previous terminology and notation and writing 2^n for the class of subsets of $\{1,\dots,n\}$, we may state the general representations as follows:

- (Hoover (1979)): *An array X on $\overline{\mathbb{N}}$ is exchangeable iff there exist a measurable function f on $\bigcup_n [0,1]^{2^n}$ and a U-array ξ on $\widetilde{\mathbb{N}}$ such that a.s.*

$$X_k = f(\hat{\xi}_k), \quad k \in \overline{\mathbb{N}}\,.$$

- (K (1992)): *An array X on $\widetilde{\mathbb{N}}$ is contractable iff there exist a measurable function f on $\bigcup_n [0,1]^{2^n}$ and a U-array ξ on $\widetilde{\mathbb{N}}$ such that a.s.*

$$X_J = f(\hat{\xi}_J), \quad J \in \widetilde{\mathbb{N}}\,.$$

As before, the last result yields an associated extension theorem:

- (K (1992)): *An array X on $\widetilde{\mathbb{N}}$ is contractable iff it can be extended to an exchangeable array on $\overline{\mathbb{N}}$.*

3 – Rotatable arrays and functionals

Even separately or jointly rotatable arrays may be characterized in terms of a.s. representations. Here we may again begin with the one-dimensional case:

- (FREEDMAN (1962)): *An infinite random sequence $X = (X_j)$ is rotatable iff there exist some i.i.d. $N(0,1)$ random variables $\zeta_1, \zeta_2, \dots$ and an independent random variable $\sigma \geq 0$ such that a.s.*

$$X_j = \sigma\zeta_j, \quad j \geq 1\,.$$

This is clearly equivalent to the previous characterization of rotatable sequences as mixed i.i.d. centered Gaussian. The two-dimensional case is again a lot harder. The following result, originally conjectured by DAWID (1978), was proved (under a moment condition) by an intricate argument based on the representation theorem for separately exchangeable arrays:

- (ALDOUS (1981)): *An array $X = (X_{ij})$ on $\mathbb{N}^2$ is separately rotatable iff there exist some i.i.d. $N(0,1)$ random variables ξ_{ki}, η_{kj}, and ζ_{ij}, along with an independent set of random coefficients σ and α_k satisfying $\sum_k \alpha_k^2 < \infty$, such that a.s.*

$$X_{ij} = \sigma\zeta_{ij} + \sum\nolimits_k \alpha_k\, \xi_{ki}\, \eta_{kj}, \quad i,j \geq 1\,.$$

For jointly rotatable arrays, we have instead:

- (K (1988)): *An array $X = (X_{ij})$ on $\mathbb{N}^2$ is jointly rotatable iff there exist some i.i.d. $N(0,1)$ random variables ξ_{ki} and ζ_{ij}, along with an independent set of random coefficients ρ, σ, σ', and α_{hk} satisfying $\sum_{h,k} \alpha_{hk}^2 < \infty$, such that a.s.*

$$X_{ij} = \rho\delta_{ij} + \sigma\zeta_{ij} + \sigma'\zeta_{ji} + \sum\nolimits_{h,k} \alpha_{hk}\,(\xi_{hi}\,\xi_{kj} - \delta_{ij}\delta_{hk}), \quad i,j \geq 1\,.$$

Here the centering terms $\delta_{ij}\delta_{hk}$ are needed, in general, to ensure convergence of the double series on the right.

The higher-dimensional representations are stated most conveniently in a Hilbert space setting. Here we consider any real, separable, infinite-dimensional Hilbert space H. By a *continuous linear random functional* (*CLRF*) on H we mean a real-valued process X on H such that

- $h_n \to 0$ in H implies $Xh_n \xrightarrow{P} 0$,
- $X(ah + bk) = aXh + bXk$ a.s. for all $h, k \in H$ and $a, b \in \mathbb{R}$.

For a simple example, we may consider an *isonormal Gaussian process* (*G-process*) on H, defined as a centered Gaussian process X on H such that

$$\mathrm{Cov}(Xh, Xk) = \langle h, k\rangle, \quad h, k \in H\,.$$

By a *unitary operator* on H we mean a linear isometry U of H onto itself. A CLRF X on H is said to be *rotatable* if $X \circ U \stackrel{d}{=} X$ for all unitary operators U on H, where $(X \circ U)h = X(Uh)$. Taking $H = l^2$, we get the following equivalent version of Freedman's theorem:

- (Freedman (1962–63)): *A CLRF X on H is rotatable iff $X = \sigma\eta$ a.s. for some G-process η on H and an independent random variable $\sigma \geq 0$.*

This formulation has the advantage of also containing the corresponding continuous–time representation mentioned earlier–the fact that a continuous process X on $\mathbb{R}_+$ with $X_0 = 0$ is rotatable iff $X = \sigma B$ a.s. for some Brownian motion B and an independent random variable $\sigma \geq 0$. This amounts to choosing $H = L^2(\lambda)$, where λ denotes Lebesgue measure on $\mathbb{R}_+$.

The higher-dimensional representations are stated, most conveniently, in terms of rotations on tensor products of Hilbert spaces H_k. The latter are best understood when $H_k = L^2(\mu_k)$ for some σ-finite measures $\mu_1, \ldots, \mu_n$ on measurable spaces $S_1, \ldots, S_n$. The *tensor product* $\bigotimes_k H_k = H_1 \otimes \cdots \otimes H_n$ of the spaces H_k can then be defined by

$$H_1 \otimes \cdots \otimes H_n = L^2(\mu_1 \otimes \cdots \otimes \mu_n)\,,$$

where $\mu_1 \otimes \cdots \otimes \mu_n$ denotes the product measure of $\mu_1, \ldots, \mu_n$ on $S_1 \times \cdots \times S_n$. For any elements $h_k \in H_k$, we define the tensor product $\bigotimes_k h_k = h_1 \otimes \cdots \otimes h_n$ in $\bigotimes_k H_k$ by

$$(h_1 \otimes \cdots \otimes h_n)(s_1, \ldots, s_n) = h_1(s_1) \cdots h_n(s_n)\,,$$

for any $s_k \in S_k$, $k = 1, \ldots, n$. Choosing an orthonormal basis (ONB) $h_{k1}, h_{k2}, \ldots$ in H_k for every k, we note that the tensor products $\bigotimes_k h_{k,j_k}$ for arbitrary $j_1, \ldots, j_n \in \mathbb{N}$ form an ONB in $\bigotimes_k H_k$.

Given any unitary operators U_k on H_k, $k = 1, \ldots, n$, there exists a unique unitary operator $\bigotimes_k U_k = U_1 \otimes \cdots \otimes U_n$ on $\bigotimes_k H_k$ such that, for any elements $h_k \in H_k$,

$$(U_1 \otimes \cdots \otimes U_n)(h_1 \otimes \cdots \otimes h_n) = U_1 h_1 \otimes \cdots \otimes U_n h_n\,.$$

When $H_k = H$ or $U_k = U$ for all k, we may write $H^{\otimes n} = \bigotimes_k H_k$ or $U^{\otimes n} = \bigotimes_k U_k$, respectively. A CLRF X on $H^{\otimes n}$ is said to be *separately rotatable* if $X \circ \bigotimes_k U_k \stackrel{d}{=} X$ for all unitary operators $U_1, \ldots, U_n$ on H and *jointly rotatable* if $X \circ U^{\otimes n} \stackrel{d}{=} X$ for any such operator U. Basic examples are the *multiple Wiener–Itô integrals* (*WI-integrals*), defined most easily, as in K (2002), through the following characterizations (as opposed to the traditional lengthy constructions):

- *For any independent G-processes η_k on H_k, $k = 1, \ldots, n$, there exists an a.s. unique CLRF $\bigotimes_k \eta_k$ on $\bigotimes_k H_k$ such that, a.s. for any elements $h_k \in H_k$,*

$$(\eta_1 \otimes \cdots \otimes \eta_n)(h_1 \otimes \cdots \otimes h_n) = \eta_1 h_1 \cdots \eta_n h_n\,.$$

- *For any G-process η on H and any $n \in \mathbb{N}$, there exists an a.s. unique CLRF $\eta^{\otimes n}$ on $H^{\otimes n}$ such that, a.s. for any orthogonal elements $h_k \in H_k$,*

$$\eta^{\otimes n}(h_1 \otimes \cdots \otimes h_n) = \eta h_1 \cdots \eta h_n \,.$$

Similarly, we may define the multiple integral $\bigotimes_k \eta_k^{\otimes r_k}$ as a CLRF on $\bigotimes_k H_k^{\otimes r_k}$ for any $r_1, \ldots, r_n \in \mathbb{N}$. It is easily seen that $\bigotimes_k \eta_k$ is separately rotatable while $\eta^{\otimes n}$ is jointly rotatable. Note that the product formula for $\eta^{\otimes n}$ fails when the elements $h_1, \ldots, h_n$ are not orthogonal. In particular, we have the a.s. representation (due to Itô (1951))

$$\eta^{\otimes n} h^{\otimes n} = \|h\|^n p_n(\eta h / \|h\|), \quad h \in H, \ n \in \mathbb{N} \,,$$

where p_n denotes the n-th degree Hermite polynomial with leading coefficient 1.

To state the representation of separately rotatable random functionals, let $\mathcal{P}_d$ denote the set of partitions π of $\{1, \ldots, d\}$, and write $H^{\otimes J} = \bigotimes_{j \in J} H$ and $H^{\otimes \pi} = \bigotimes_{J \in \pi} H$.

- (K (1995)): *A CLRF X on $H^{\otimes d}$ is separately rotatable iff there exist some independent G-processes η_J on $H \otimes H^{\otimes J}$, $J \in 2^d \setminus \{\emptyset\}$, and an independent set of random elements $\alpha_\pi \in H^{\otimes \pi}$, $\pi \in \mathcal{P}_d$, such that a.s.*

$$Xf = \sum_{\pi \in \mathcal{P}_d} \left(\bigotimes_{J \in \pi} \eta_J \right) (\alpha_\pi \otimes f), \quad f \in H^{\otimes d} \,.$$

The last formula exhibits X as a finite sum of randomized multiple WI-integrals. Introducing an ONB $h_1, h_2, \ldots$ in H and writing

$$X_{k_1,\ldots,k_d} = X(h_{k_1} \otimes \cdots \otimes h_{k_d}), \quad k_1, \ldots, k_d \in \mathbb{N} \,,$$

we may write the previous representation in coordinate form as

$$X_k = \sum_{\pi \in \mathcal{P}_d} \sum_{l \in \mathbb{N}^\pi} \alpha_l^\pi \prod_{J \in \pi} \eta^J_{k_J, l_J}, \quad k \in \mathbb{N}^d \,,$$

for some i.i.d. $N(0,1)$ random variables η^J_{kl} and an independent collection of random elements α_l^π satisfying $\sum_l (\alpha_l^\pi)^2 < \infty$ a.s. Any separately rotatable array on $\mathbb{N}^d$ can be represented in this form. Note that in the functional version, the coefficients α_l^π have been combined into random elements α_π of H, which explains the role of the extra dimension of the G-processes η_J.

We turn to the more complicated jointly rotatable case. Here we write $\mathcal{O}_d$ for the class of partitions of $\{1, \ldots, d\}$ into *ordered* subsets $k = (k_1, \ldots, k_r) \in \mathbb{N}^{(r)}$, $1 \le r \le d$. The dimension r of k is denoted by $|k|$.

- (K (1995)): *A CLRF X on $H^{\otimes d}$ is jointly rotatable iff there exist some independent G-processes η_r on $H^{\otimes(r+1)}$, $r = 1, \ldots, d$, and an independent set of random elements $\alpha_\pi \in H^{\otimes \pi}$, $\pi \in \mathcal{O}_d$, such that a.s.*

$$Xf = \sum_{\pi \in \mathcal{O}_d} \left(\bigotimes_{k \in \pi} \eta_{|k|} \right) (\alpha_\pi \otimes f), \quad f \in H^{\otimes d}.$$

Here the multiple integral $\bigotimes_{k\in\pi} \eta_{|k|}$ is understood to depend, in an obvious way, on the order of the elements within each sequence k.

The displayed formula may again be stated in basis form, using the mentioned expression of WI-integrals in terms of Hermite polynomials. However, the representation of jointly rotatable arrays is more complicated, as it also includes diagonal terms of different order. For example, the term $\rho\delta_{ij}$ in the quoted representation on $\mathbb{N}^2$ has no extension to a CLRF on $H^{\otimes 2}$. This shows another advantage of the Hilbert space setting, apart from the avoidance of infinite series involving Hermite polynomials.

4 – Exchangeable random sheets

We have already noted the close relationship between exchangeability and rotatability for continuous processes on $\mathbb{R}_+$. Exploring this connection, we may derive representations of certain separately or jointly exchangeable or contractable processes on $\mathbb{R}_+^d$ (and occasionally on $[0,1]^d$). By a *random sheet* on $\mathbb{R}_+^d$ we mean a continuous process $X = (X_t)$ that vanishes on all coordinate hyperplanes, so that $X_t = 0$ when $\bigwedge_j t_j = 0$. Note that exchangeability and rotatability may now be defined in an obvious way in terms of the multivariate increments.

To understand the higher-dimensional formulas, we may first consider the case of separately or jointly rotatable random sheets on $\mathbb{R}_+^2$. The following representations follow easily from the previous results for rotatable arrays.

- *A random sheet X on $\mathbb{R}_+^2$ is separately rotatable iff there exist some independent Brownian motions $B^1, B^2, \ldots$ and $C^1, C^2, \ldots$ and an independent Brownian sheet Z, along with an independent set of random coefficients σ and α_k with $\sum_k \alpha_k^2 < \infty$ a.s., such that a.s.*

$$X_{s,t} = \sigma Z_{s,t} + \sum\nolimits_k \alpha_k \, B_s^k \, C_t^k, \quad s, t \geq 0.$$

- *A random sheet X on $\mathbb{R}_+^2$ is jointly rotatable iff there exist some independent Brownian motions $B^1, B^2, \ldots$ and an independent Brownian sheet Z,* *along with an independent set of random coefficients* ρ, σ, σ', *and* α_k *with* $\sum_k \alpha_k^2 < \infty$ *a.s., such that a.s.*

$$X_{s,t} = \rho(s \wedge t) + \sigma Z_{s,t} + \sigma' Z_{t,s} + \\ + \sum\nolimits_{h,k} \alpha_{hk} \, (B_s^h \, B_t^k - \delta_{hk}(s \wedge t)), \quad s, t \geq 0 \, .$$

The representations in the exchangeable case are similar, apart from some additional centering terms.

- (K (1988)): *A random sheet* X *on* $\mathbb{R}_+^2$ *is separately exchangeable iff there exist some independent Brownian motions* B^k *and* C^k *and an independent Brownian sheet* Z*, along with an independent set of random coefficients* ϑ, σ, *and* $\alpha_k, \beta_k, \gamma_k$ *with* $\sum_k (\alpha_k^2 + \beta_k^2 + \gamma_k^2) < \infty$ *a.s., such that a.s.*

$$X_{s,t} = \vartheta st + \sigma Z_{s,t} + \sum\nolimits_k (\alpha_k \, B_s^k \, C_t^k + \beta_k \, t B_s^k + \gamma_k \, s C_t^k), \quad s, t \geq 0 \, .$$

- (K (1988)): *A random sheet* X *on* $\mathbb{R}_+^2$ *is jointly exchangeable iff there exist some independent Brownian motions* B^k *and an independent Brownian sheet* Z*, along with an independent set of random coefficients* ρ, ϑ, σ, σ', *and* $\alpha_k, \beta_k, \beta'_k, \gamma_k$ *with* $\sum_k (\alpha_k^2 + \beta_k^2 + {\beta'}_k^2 + \gamma_k^2) < \infty$ *a.s., such that a.s.*

$$X_{s,t} = \rho(s \wedge t) + \vartheta st + \sigma Z_{s,t} + \sigma' Z_{t,s} + \\ + \sum\nolimits_{h,k} \alpha_{hk} \, (B_s^h \, B_t^k - \delta_{hk}(s \wedge t)) + \\ + \sum\nolimits_k (\beta_k \, t B_s^k + \beta'_k \, s B_t^k + \gamma_k \, B_{s \wedge t}^k), \quad s, t \geq 0 \, .$$

Partial results of this type were also obtained, independently, in an unpublished thesis of Hestir (1986).

The higher-dimensional representations may again be stated, most conveniently, in terms of multiple WI-integrals. Here we write $\hat{\mathcal{P}}_d = \bigcup_I \mathcal{P}_I$, where $\mathcal{P}_I$ denotes the class of partitions π of $I \in 2^d \setminus \{\emptyset\}$. For $\pi \in \mathcal{P}_I$, we write $\pi^c = I^c$. Let λ^I denote Lebesgue measure on $\mathbb{R}_+^I$. For notational convenience, we may identify a set A with its indicator function 1_A.

- (K (1995)): *A random sheet* X *on* $\mathbb{R}_+^d$ *is separately exchangeable iff there exist some independent G-processes* η_I *on* $H \otimes L^2(\lambda^I)$, $I \in 2^d \setminus \{\emptyset\}$*, along with an independent set of random coefficients* $\alpha_\pi \in H^{\otimes \pi}$, $\pi \in \hat{\mathcal{P}}_d$*, such that a.s.*

$$X_t = \sum_{\pi \in \hat{\mathcal{P}}_d} \left(\lambda^{\pi^c} \otimes \bigotimes_{I \in \pi} \eta_I \right) (\alpha_\pi \otimes [0, t]), \quad t \in \mathbb{R}_+^d \, .$$

A similar representation holds for separately exchangeable random sheets on $[0,1]^d$, except that the G-processes η_J then need to be replaced by suitably "tied-down" versions.

The jointly exchangeable case is even more complicated and requires some further notation. Given a finite set J, we define $\hat{\mathcal{O}}_J = \bigcup_{I \subset J} \mathcal{O}_I$, where $\mathcal{O}_I$ denotes the class of partitions of I into ordered subsets k of size $|k| \geq 1$. In this definition, we may take J to be an arbitrary partition $\pi \in \mathcal{P}_d$, regarded as a finite collection of sets $\{J_1, \dots, J_r\}$. For any $t \in \mathbb{R}_+^d$ and $\pi \in \mathcal{P}_d$, we introduce the vector $\hat{t}_\pi \in \mathbb{R}_+^\pi$ with components $\hat{t}_{\pi,J} = \bigwedge_{j \in J} t_j$, $J \in \pi$.

- (K (1995)): *A random sheet X on $\mathbb{R}_+^d$ is jointly exchangeable iff there exist some independent G-processes η_r on $H \otimes L^2(\lambda^r)$, $1 \leq r \leq d$, along with an independent set of random coefficients $\alpha_{\pi,\kappa} \in H^{\otimes \kappa}$, $\kappa \in \hat{\mathcal{O}}_\pi$, such that a.s.*

$$X_t = \sum_{\pi \in \mathcal{P}_d} \sum_{\kappa \in \hat{\mathcal{O}}_\pi} \Big(\lambda^{\kappa^c} \otimes \bigotimes_{k \in \kappa} \eta_{|k|} \Big) (\alpha_{\pi,\kappa} \otimes [0, \hat{t}_\pi]), \quad t \in \mathbb{R}_+^d \,.$$

One would expect the last representation to remain valid for jointly exchangeable sheets on $[0,1]^d$, with the G-processes η_r replaced by their tied-down versions. However, the status of this conjecture is still open. We may also mention some similar but still more complicated representations, available for jointly contractable sheets on $\mathbb{R}_+^d$ (cf. K (2005), p. 398). Finally, there exists an extensive theory of exchangeable random measures in the plane, covered by K (1990, 2005) but not included in the present survey.

In summary, the previous representations illustrate the amazing unity of the subject: using representations of contractable or exchangeable arrays, we may derive representations of rotatable random functionals in terms of multiple WI-integrals, which can then be used to obtain representations of exchangeable or contractable random sheets.

5 – Some open problems

The theory of multivariate symmetries is still incomplete. We conclude with a list of open problems in the area.

- Give a direct proof of the extension theorem for contractable arrays. This would provide an alternative approach to the deep representation theorem for such arrays, given Hoover's representation in the jointly exchangeable case. Some difficulties are likely to arise from the non-uniqueness, the fact that different representations may lead to different extensions. Is there a natural choice?
- Characterize the jointly exchangeable random sheets on $[0,1]^d$. One expects the representation on $\mathbb{R}_+^d$ to remains valid with the G-processes η_r replaced

by their tied-down versions. Hence, the question is whether there exist exchangeable sheets that are not given by this formula.

- Find representations of arrays and processes with different symmetries (contractable, exchangeable, or rotatable) in different indices or variables.
- Characterize the classes of separately or jointly exchangeable random measures on $\mathbb{R}_+^d$ for $d \geq 3$. The complexity of such representations, already for $d = 2$ (cf. K (1990, 2005)), suggests that one should first look for a compact way of writing these formulas, starting perhaps with the special case of simple point processes.
- Extend BÜHLMANN'S (1960) theorem to higher dimensions, by characterizing processes on $\mathbb{R}_+^d$ with separately or jointly exchangeable increments. It seems reasonable to begin with the case of signed random measures on $\mathbb{R}_+^2$.
- Multiple Wiener-Itô integrals constitute the basic examples of rotatable arrays and functionals in higher dimensions. Are there any natural symmetries leading to multiple p-stable integrals for $p < 2$? For the one-dimensional case, see e.g. DIACONIS and FREEDMAN (1987).
- Can the semigroup methods of RESSEL (1985) be used to derive the representations of separately or jointly rotatable arrays and functionals? In view of the complexity of the current proofs, it seems worthwhile looking for alternative approaches.

Acknowledgements

My sincere thanks to the Editors for providing a long list of helpful comments and suggestions.

REFERENCES

[1] D. ALDOUS: *Representations for partially exchangeable arrays of random variables*, J. Multivar. Anal., **11** (1981), 581–598.

[2] D. ALDOUS: *Exchangeability and related topics*, In "École d'Été de Probabilités de Saint-Flour XIII - 1983"; Lect. Notes in Math., **1117**, pp. 1-198, Springer, Berlin, 1985.

[3] H. BÜHLMANN: *Austauschbare stochastische Variabeln und ihre Grenzwertsätze*, Univ. Calif. Publ. Statist., **3** (1960), 1–35.

[4] A. P. DAWID: *Extendibility of spherical matrix distributions*, J. Multivar. Anal., **8** (1978), 559–566.

[5] P. DIACONIS – D. FREEDMAN: *A dozen de Finetti-style results in search of a theory*, Ann. Inst. H. Poincaré, **23** (1987), 397–423.

[6] DE FINETTI: *Funzione caratteristica di un fenomeno aleatorio*, Mem. R. Acc. Lincei, **4** (6) (1930), 86–133.

[7] de Finetti: *La prévision: ses lois logiques, ses sources subjectives*, Ann. Inst. H. Poincaré, **7** (1937), 1–68.

[8] de Finetti: *Sur la condition d'équivalence partielle*, Act. Sci. Ind., **739** (1938), 5–18.

[9] D. Freedman: *Invariants under mixing which generalize de Finetti's theorem*, Ann. Math. Statist., **33** (1962), 916–923; **34** (19-63), 1194–1216.

[10] K. Hestir: *The Aldous representation theorem and weakly exchangeable non-negative definite arrays*, Ph.D. Thesis, Statist. Dept., Univ. of California, Berkeley, 1986.

[11] D. N. Hoover: *Relations on probability spaces and arrays of random variables*, Preprint. Institute of Advanced Study, Princeton, 1979.

[12] D. N. Hoover: *Row-column exchangeability and a generalized model for probability*, In "Exchangeability in Probability and Statistics", G. Koch – F. Spizzichino (eds.), pp. 281–291; North-Holland, Amsterdam, 1982.

[13] K. Itô: *Multiple Wiener integral*, J. Math. Soc. Japan, **3** (1951), 157–169.

[14] O. Kallenberg: *Some new representations in bivariate exchangeability*, Probab. Th. Rel. Fields, **77** (1988), 415–455.

[15] O. Kallenberg: *Exchangeable random measures in the plane*, J. Theor. Probab., **3** (1990), 81–136.

[16] O. Kallenberg: *Symmetries on random arrays and set-indexed processes*, J. Theor. Probab., **5** (1992), 727–765.

[17] O. Kallenberg: *Random arrays and functionals with multivariate rotational symmetries*, Probab. Th. Rel. Fields, **103** (1995), 91–141.

[18] O. Kallenberg: *Foundations of Modern Probability*, 2nd ed., Springer, New York, 2002.

[19] O. Kallenberg: *Probabilistic Symmetries and Invariance Principles*, Springer, New York, 2005.

[20] P. Ressel: *De Finetti-type theorems: an analytical approach*, Ann. Probab., **13** (1985), 898–922.

[21] C. Ryll-Nardzewski: *On stationary sequences of random variables and the de Finetti equivalence*, Colloq. Math., **4** (1957), 149–156.

Lavoro pervenuto alla redazione il 12 ottobre 2007
ed accettato per la pubblicazione il 18 ottobre 2007.
Bozze licenziate il 9 aprile 2008

Indirizzo dell'Autore:

Olav Kallenberg – Department of Mathematics and Statistics – 221 Parker Hall, Auburn University – Auburn, AL 36849-5310, USA
E-mail: kalleoh@auburn.edu

Rendiconti di Matematica, Serie VII
Volume 28, Roma (2008), 33–61

Graph limits and exchangeable random graphs

PERSI DIACONIS – SVANTE JANSON

Abstract: *We develop a clear connection between de Finetti's theorem for exchangeable arrays (work of Aldous–Hoover–Kallenberg) and the emerging area of graph limits (work of Lovász and many coauthors). Along the way, we translate the graph theory into more classical probability.*

1 – Introduction

De Finetti's profound contributions are now woven into many parts of probability, statistics and philosophy. Here we show how developments from de Finetti's work on partial exchangeability have a direct link to the recent development of a limiting theory for large graphs. This introduction first recalls the theory of exchangeable arrays (Section 1.1). Then, the subject of graph limits is outlined (Section 1.2). Finally, the link between these ideas, which forms the bulk of this paper, is outlined (Section 1.3).

1.1 – Exchangeability, partial exchangeability and exchangeable arrays

Let $\{X_i\}$, $1 \le i < \infty$, be a sequence of binary random variables. They are *exchangeable* if

$$\mathbb{P}(X_1 = e_1, \dots, X_n = e_n) = \mathbb{P}(X_1 = e_{\sigma(1)}, \dots, X_n = e_{\sigma(n)})$$

for all n, permutations $\sigma \in \mathfrak{S}_n$ and all $e_i \in \{0, 1\}$. The celebrated representation theorem [10, 11] says

Key Words and Phrases: *Graph limit – Exchangeable array – de Finetti's theorem.*
A.M.S. Classification: 60G09, 05C80, 05C62

Theorem 1.1. (de Finetti) *If* $\{X_i\}$, $1 \le i < \infty$, *is a binary exchangeable sequence, then:*

(i) *With probability* 1, $X_\infty = \lim \frac{1}{n}(X_1 + \cdots + X_n)$ *exists.*
(ii) *If* $\mu(A) = P\{X_\infty \in A\}$, *then for all* n *and* e_i, $1 \le i \le n$,

$$\mathbb{P}(X_1 = e_1, \dots, X_n = e_n) = \int_0^1 x^s (1-x)^{n-s} \mu(dx) \tag{1.1}$$

for $s = e_1 + \cdots + e_n$.

It is natural to refine and extend de Finetti's theorem to allow more general observables (X_i with values in a Polish space) and other notions of symmetry (partial exchangeability). A definitive treatment of these developments is given in Kallenberg [17]. Of interest here is the extension of de Finetti's theorem to two-dimensional arrays.

Definition. Let $\{X_{ij}\}$, $1 \le i, j < \infty$, be binary random variables. They are *separately exchangeable* if

$$\mathbb{P}(X_{ij} = e_{ij},\ 1 \le i, j \le n) = \mathbb{P}(X_{ij} = e_{\sigma(i)\tau(j)},\ 1 \le i, j \le n) \tag{1.2}$$

for all n, all permutations $\sigma, \tau \in \mathfrak{S}_n$ and all $e_{ij} \in \{0, 1\}$. They are (*jointly*) *exchangeable* if (1.2) holds in the special case $\tau = \sigma$.

Equivalently, the array $\{X_{ij}\}$ is jointly exchangeable if the array $\{X_{\sigma(i)\sigma(j)}\}$ has the same distribution as $\{X_{ij}\}$ for every permutation σ of $\mathbb{N}$, and similarly for separate exchangeability.

The question of two-dimensional versions of de Finetti's theorem under (separate) exchangeability arose from the statistical problems of two-way analysis of variance. Early workers expected a version of (1.1) with perhaps a two-dimensional integral. The probabilist David Aldous [1] and the logician Douglas Hoover [16] found that the answer is more complicated.

Define a random binary array $\{X_{ij}\}$ as follows: Let U_i, V_j, $1 \le i, j < \infty$, be independent and uniform in $[0, 1]$. Let $W(x, y)$ be a function from $[0, 1]^2$ to $[0, 1]$. Let X_{ij} be 1 or 0 as a $W(U_i, V_j)$-coin comes up heads or tails. Let P_W be the probability distribution of $\{X_{ij}\}$, $1 \le i, j < \infty$. The family $\{X_{ij}\}$ is separately exchangeable because of the symmetry of the construction. The Aldous–Hoover theorem says that any separately exchangeable binary array is a mixture of such P_W:

Theorem 1.2. (Aldous–Hoover) *Let* $X = \{X_{ij}\}$, $1 \le i, j < \infty$, *be a separately exchangeable binary array. Then, there is a probability* μ *such that*

$$\mathbb{P}\{X \in A\} = \int P_W(A) \mu(dW).$$

There is a similar result for jointly exchangeable arrays.

The uniqueness of μ resisted understanding; if $\widehat{W}$ is obtained from W by a measure-preserving change of each variable, clearly the associated process $\{\widehat{X}_{ij}\}$ has the same joint distribution as $\{X_{ij}\}$. Using model theory, Hoover [16] was able to show that this was the only source of non-uniqueness. A 'probabilist's proof' was finally found by Kallenberg, see [17, Sect. 7.6] for details and references.

These results hold for higher dimensional arrays with X_{ij} taking values in a Polish space with minor change [17, Chap. 7]. The description above has not mentioned several elegant results of the theory. In particular, Kallenberg's 'spreadable' version of the theory replaces invariance under a group by invariance under subsequences. A variety of tail fields may be introduced to allow characterizing when W takes values in $\{0, 1\}$ [12, Sect. 4]. Much more general notions of partial exchangeability are studied in [13].

1.2 – Graph limits

Large graphs, both random and deterministic, abound in applications. They arise from the internet, social networks, gene regulation, ecology and in mathematics. It is natural to seek an approximation theory: What does it mean for a sequence of graphs to converge? When can a large complex graph be approximated by a small graph?

In a sequence of papers [6, 7, 8, 9, 15, 18, 19, 20, 23, 22, 24, 21] Laszlo Lovász with coauthors (listed here in order of frequency) V. T. Sós, B. Szegedy, C. Borgs, J. Chayes, K. Vesztergombi, A. Schrijver, M. Freedman have developed a beautiful, unifying limit theory. This sheds light on topics such as graph homomorphisms, Szemeredi's regularity lemma, quasi-random graphs, graph testing and extremal graph theory. Their theory has been developed for dense graphs (number of edges comparable with the square of number of vertices) but parallel theories for sparse graphs are beginning to emerge [4].

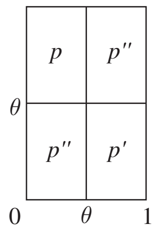
Roughly, a growing sequence of finite graphs G_n converges if, for any fixed graph F, the proportion of copies of F in G_n converges. Section 2 below has precise definitions.

Example 1.3 Define a probability distribution on graphs on n-vertices as follows. Flip a θ-coin for each vertex (dividing vertices into 'boys' and 'girls'). Connect two boys with probability p. Connect two girls with probability p'. Connect a boy and a girl with probability p''. Thus, if $p = p' = 0$, $p'' = 1$, we have a random bipartite graph. If $p = p' = 1$, $p'' = 0$, we have two disjoint complete graphs. If $p = p' = p''$, we have the Erdös–Renyi model. As n grows, these models generate a sequence of random graphs which converge almost surely to a limiting object described below.

More substantial examples involving random threshold graphs are in [14].

If a sequence of graphs converges, what does it converge to? For exchangeable random graphs (defined below), there is a limiting object which may be thought of as a probability measure on infinite random graphs. Suppose $W(x, y) = W(y, x)$ is a function from $[0, 1]^2 \to [0, 1]$. Choose $\{U_i\}$, $1 \le i < \infty$, independent uniformly distributed random variables on $[0, 1]$. Form an infinite random graph by putting an edge from i to j with probability $W(U_i, U_j)$. This measure on graphs (or alternatively W) is the limiting object.

For the "boys and girls" example above, W may be pictured as



The theory developed shows that various properties of G_n can be well approximated by calculations with the limiting object. There is an elegant characterization of these 'continuous graph properties' with applications to algorithms for graph testing (Does this graph contain an Eulerian cycle?) or parameter estimation (What is an approximation to the size of the maximum cut?). There is a practical way to find useful approximations to a large graph by graphs of fixed size [6]. This paper also contains a useful review of the current state of the theory with proofs and references.

We have sketched the theory for unweighted graphs. There are generalizations to graphs with weights on vertices and edges, to bipartite, directed and hypergraphs. The sketch leaves out many nice developments. For example, the useful cut metric between graphs [21] and connections to statistical physics [9].

1.3 – Overview of the present paper

There is an apparent similarity between the measure P_W of the Aldous–Hoover theorem and the limiting object W from graph limits. Roughly, working with symmetric W gives the graph limit theory; working with general W gives directed graphs. The main results of this paper make these connections precise.

Basic definitions are in Section 2 which introduces a probabilist's version of graph convergence equivalent to the definition using graph homomorphisms. Section 3 uses the well-established theory of weak convergence of a sequence of probability measures on a metric space to get properties of graph convergence. Section 4 carries things over to infinite graphs.

The main results appear in Section 5. This introduces exchangeable random graphs and gives a one-to-one correspondence between infinite exchangeable random graphs and distributions on the space of proper graph limits (Theorem 5.3), which specializes to a one-to-one correspondence between proper graph limits and extreme points in the set of distributions of exchangeable random graphs (Corollary 5.4).

A useful characterization of the extreme points of the set of exchangeable random graphs is in Theorem 5.5. These results are translated to the equivalence between proper graph limits and the Aldous–Hoover theory in Section 6. The non-uniqueness of the representing W, for exchangeable random graphs and for graph limits, is discussed in Section 7.

The equivalence involves symmetric $W(x,y)$ and a single permutation σ taking $W(U_i,U_j)$ to $W(U_{\sigma(i)},U_{\sigma(j)})$. The original Aldous–Hoover theorem, with perhaps non-symmetric $W(x,y)$ and $W(U_i,V_j)$ to $W(U_{\sigma(i)},V_{\tau(j)})$ translates to a limit theorem for bipartite graphs. This is developed in Section 8. The third case of the Aldous–Hoover theory for two-dimensional arrays, perhaps non-symmetric $W(x,y)$ and a single permutation σ, corresponds to directed graphs; this is sketched in Section 9.

The extensions to weighted graphs are covered by allowing X_{ij} to take general values in the Aldous–Hoover theory. The extension to hypergraphs follows from the Aldous–Hoover theory for higher-dimensional arrays. (The details of these extensions are left to the reader.)

Despite these parallels, the theories have much to contribute to each other. The algorithmic, graph testing, Szemeredi partitioning perspective is new to exchangeability theory. Indeed, the “boys and girls” random graph was introduced to study the psychology of vision in Diaconis–Freedman (1981). As far as we know, its graph theoretic properties have not been studied. The various developments around shell-fields in exchangeability, which characterize zero/one $W(x,y)$, have yet to be translated into graph-theoretic terms.

2 – Definitions and basic properties

All graphs will be simple, without multiple edges or loops. Infinite graphs will be important in later sections, but will always be clearly stated to be infinite; otherwise, graphs will be finite. We denote the vertex and edge sets of a graph G by $V(G)$ and $E(G)$, and the numbers of vertices and edges by $v(G) := |V(G)|$ and $e(G) := |E(G)|$. We consider both labelled and unlabelled graphs; the labels will be the integers $1,\dots,n$, where n is the number of vertices in the graph. A labelled graph is thus a graph with vertex set $[n] := \{1,\dots,n\}$ for some $n \ge 1$; we let $\mathcal{L}_n$ denote the set of the $2^{\binom{n}{2}}$ labelled graphs on $[n]$ and let $\mathcal{L} := \bigcup_{n=1}^{\infty}\mathcal{L}_n$. An unlabelled graph can be regarded as a labelled graph where we ignore the labels; formally, we define $\mathcal{U}_n$, the set of unlabelled graphs of order n, as the quotient set $\mathcal{L}_n/\cong$ of labelled graphs modulo isomorphisms. We let $\mathcal{U} := \bigcup_{n=1}^{\infty}\mathcal{U}_n = \mathcal{L}/\cong$, the set of all unlabelled graphs.

Note that we can, and often will, regard a labelled graph as an unlabelled graph.

If G is an (unlabelled) graph and $v_1, \dots, v_k$ is a sequence of vertices in G, then $G(v_1, \dots, v_k)$ denotes the labelled graph with vertex set $[k]$ where we put an edge between i and j if v_i and v_j are adjacent in G. We allow the possibility that $v_i = v_j$ for some i and j. (In this case, there is no edge ij because there are no loops in G.)

We let $G[k]$, for $k \geq 1$, be the random graph $G(v_1, \dots, v_k)$ obtained by sampling $v_1, \dots, v_k$ uniformly at random among the vertices of G, with replacement. In other words, $v_1, \dots, v_k$ are independent uniformly distributed random vertices of G.

For $k \leq v(G)$, we further let $G[k]'$ be the random graph $G(v'_1, \dots, v'_k)$ where we sample $v'_1, \dots, v'_k$ uniformly at random without replacement; the sequence $v'_1, \dots, v'_k$ is thus a uniformly distributed random sequence of k distinct vertices.

The graph limit theory in [21] and subsequent papers is based on the study of the functional $t(F, G)$ which is defined for two graphs F and G as the proportion of all mappings $V(F) \to V(G)$ that are graph homomorphisms $F \to G$, i.e., map adjacent vertices to adjacent vertices. In probabilistic terms, $t(F, G)$ is the probability that a uniform random mapping $V(F) \to V(G)$ is a graph homomorphism. Using the notation introduced above, we can, equivalently, write this as, assuming that F is labelled and $k = v(F)$,

$$t(F, G) := \mathbb{P}\big(F \subseteq G[k]\big). \tag{2.1}$$

Note that both F and $G[k]$ are graphs on $[k]$, so the relation $F \subseteq G[k]$ is well-defined as containment of labelled graphs on the same vertex set, i.e. as $E(F) \subseteq E(G[k])$. Although the relation $F \subseteq G[k]$ may depend on the labelling of F, the probability in (2.1) does not, by symmetry, so $t(F, G)$ is really well defined by (2.1) for unlabelled F and G.

With F, G and k as in (2.1), we further define, again following [21] (and the notation of [8]) but stating the definitions in different but equivalent forms,

$$t_{\text{inj}}(F, G) := \mathbb{P}\big(F \subseteq G[k]'\big) \tag{2.2}$$

and

$$t_{\text{ind}}(F, G) := \mathbb{P}\big(F = G[k]'\big), \tag{2.3}$$

provided F and G are (unlabelled) graphs with $v(F) \leq v(G)$. If $v(F) > v(G)$ we set $t_{\text{inj}}(F, G) := t_{\text{ind}}(F, G) := 0$.

Since the probability that a random sample $v_1, \dots, v_k$ of vertices in G contains some repeated vertex is $\leq k^2/(2v(G))$, it follows that [21]

$$\big|t(F, G) - t_{\text{inj}}(F, G)\big| \leq \frac{v(F)^2}{2v(G)}. \tag{2.4}$$

Hence, when considering asymptotics with $v(G) \to \infty$, it does not matter whether we use t or t_{inj}. Moreover, if $F \in \mathcal{L}_k$, then, as pointed out in [8] and [21],

$$t_{\mathrm{inj}}(F,G) = \sum_{F' \in \mathcal{L}_k,\ F' \supseteq F} t_{\mathrm{ind}}(F',G) \tag{2.5}$$

and, by inclusion-exclusion,

$$t_{\mathrm{ind}}(F,G) = \sum_{F' \in \mathcal{L}_k,\ F' \supseteq F} (-1)^{e(F')-e(F)} t_{\mathrm{inj}}(F',G). \tag{2.6}$$

Hence, the two families $\{t_{\mathrm{inj}}(F,\cdot)\}_{F\in\mathcal{U}}$ and $\{t_{\mathrm{ind}}(F,\cdot)\}_{F\in\mathcal{U}}$ of graph functionals contain the same information and can replace each other.

The basic definition of Lovász and Szegedy [21] and Borgs, Chayes, Lovász, Sós and Vesztergombi [8] is that a sequence (G_n) of graphs converges if $t(F, G_n)$ converges for every graph F. We can express this by considering the map $\tau : \mathcal{U} \to [0,1]^{\mathcal{U}}$ defined by

$$\tau(G) := (t(F,G))_{F\in\mathcal{U}} \in [0,1]^{\mathcal{U}}. \tag{2.7}$$

Then (G_n) converges if and only if $\tau(G_n)$ converges in $[0,1]^{\mathcal{U}}$, equipped with the usual product topology. Note that $[0,1]^{\mathcal{U}}$ is a compact metric space; as is well known, a metric can be defined by, for example,

$$d\big((x_F),(y_F)\big) := \sum_{i=0}^{\infty} 2^{-i}|x_{F_i} - y_{F_i}|, \tag{2.8}$$

where $F_1, F_2, \dots$ is some enumeration of all unlabelled graphs.

We define $\mathcal{U}^* := \tau(\mathcal{U}) \subseteq [0,1]^{\mathcal{U}}$ to be the image of $\mathcal{U}$ under this mapping τ, and let $\overline{\mathcal{U}^*}$ be the closure of $\mathcal{U}^*$ in $[0,1]^{\mathcal{U}}$. Thus $\overline{\mathcal{U}^*}$ is a compact metric space. (For explicit descriptions of the subset $\overline{\mathcal{U}^*}$ of $[0,1]^{\mathcal{U}}$ as a set of graph functionals, see Lovász and Szegedy [21].)

As pointed out in [21] and [8] (in equivalent terminology), τ is not injective; for example, $\tau(K_{n,n})$ is the same for all complete bipartite graphs $K_{n,n}$. Nevertheless, as in [21] and [8], we can consider a graph G as an element of $\mathcal{U}^*$ by identifying G and $\tau(G)$ (thus identifying graphs with the same $\tau(G)$), and then convergence of (G_n) as defined above is equivalent to convergence in $\overline{\mathcal{U}^*}$. The limit is thus an element of $\overline{\mathcal{U}^*}$, but typically not a graph in $\mathcal{U}^*$. The main result of Lovász and Szegedy [21] is a representation of the elements in $\overline{\mathcal{U}^*}$ to which we will return in Section 6.

Remark 2.1. As said above, $\overline{\mathcal{U}^*}$ is a compact metric space, and it can be given several equivalent metrics. One metric is the metric (2.8) inherited from

$[0,1]^{\mathcal{U}}$, which for graphs becomes $d(G,G') = \sum_i 2^{-i}|t(F_i,G) - t(F_i,G')|$. Another metric, shown by Borgs, Chayes, Lovász, Sós and Vesztergombi [8] to be equivalent, is the cut-distance $\delta_\square$, see [8] for definitions. Further characterizations of convergence of sequences of graphs in $\overline{\mathcal{U}}$ are given in [8, 9].

The identification of graphs with the same image in $\mathcal{U}^*$ (i.e., with the same $t(F,\cdot)$ for all F) is sometimes elegant but at other times inconvenient. It can be avoided if we instead let $\mathcal{U}^+$ be the union of $\mathcal{U}$ and some one-point set $\{*\}$ and consider the mapping $\tau^+ : \mathcal{U} \to [0,1]^{\mathcal{U}^+} = [0,1]^{\mathcal{U}} \times [0,1]$ defined by

$$\tau^+(G) = \big(\tau(G),\, v(G)^{-1}\big). \tag{2.9}$$

Then τ^+ is injective, because if $\tau(G_1) = \tau(G_2)$ for two graphs G_1 and G_2 with the same number of vertices, then G_1 and G_2 are isomorphic and thus $G_1 = G_2$ as unlabelled graphs. (This can easily be shown directly: it follows from (2.1) that $G_1[k] \overset{\mathrm{d}}{=} G_2[k]$ for every k, which implies $G_1[k]' \overset{\mathrm{d}}{=} G_2[k]'$ for every $k \le v(G_1) = v(G_2)$; now take $k = v(G_1)$. It is also a consequence of [8, Theorem 2.7 and Theorem 2.3 or Lemma 5.1].)

Consequently, we can identify $\mathcal{U}$ with its image $\tau^+(\mathcal{U}) \subseteq [0,1]^{\mathcal{U}^+}$ and define $\overline{\mathcal{U}} \subseteq [0,1]^{\mathcal{U}^+}$ as its closure. It is easily seen that a sequence (G_n) of graphs converges in $\overline{\mathcal{U}}$ if and only if either $v(G_n) \to \infty$ and (G_n) converges in $\overline{\mathcal{U}^*}$, or the sequence (G_n) is constant from some n_0 on. Hence, convergence in $\overline{\mathcal{U}}$ is essentially the same as the convergence considered by by Lovász and Szegedy [21], but without any identification of non-isomorphic graphs of different orders.

Alternatively, we can consider τ_{inj} or τ_{ind} defined by

$$\begin{aligned}\tau_{\mathrm{inj}}(G) &:= (t_{\mathrm{inj}}(F,G))_{F\in\mathcal{U}} \in [0,1]^{\mathcal{U}},\\ \tau_{\mathrm{ind}}(G) &:= (t_{\mathrm{ind}}(F,G))_{F\in\mathcal{U}} \in [0,1]^{\mathcal{U}}.\end{aligned}$$

It is easy to see that both τ_{inj} and τ_{ind} are injective mappings $\mathcal{U} \to [0,1]^{\mathcal{U}}$. (If $t_{\mathrm{inj}}(F,G_1) = t_{\mathrm{inj}}(F,G_2)$ for all F, we take $F = G_1$ and $F = G_2$ and conclude $G_1 = G_2$, using our special definition of t_{inj} when $v(F) > v(G)$.) Hence, we can again identify $\mathcal{U}$ with its image and consider its closure $\overline{\mathcal{U}}$ in $[0,1]^{\mathcal{U}}$. Moreover, using (2.4), (2.5), and (2.6), it is easily shown that if (G_n) is a sequence of unlabelled graphs, then

$$\tau^+(G_n) \text{ converges} \iff \tau_{\mathrm{ind}}(G_n) \text{ converges} \iff \tau_{\mathrm{inj}}(G_n) \text{ converges.} \tag{2.10}$$

Hence, the three compactifications $\overline{\tau^+(\mathcal{U})}$, $\overline{\tau_{\mathrm{inj}}(\mathcal{U})}$, $\overline{\tau_{\mathrm{ind}}(\mathcal{U})}$ are homeomorphic and we can use any of them for $\overline{\mathcal{U}}$. We let $\mathcal{U}_\infty := \overline{\mathcal{U}} \setminus \mathcal{U}$; this is the set of all limit objects of sequences (G_n) in $\mathcal{U}$ with $v(G_n) \to \infty$. (I.e., it is the set of all proper graph limits.)

We will in the sequel prefer to use $\overline{\mathcal{U}}$ rather than $\overline{\mathcal{U}^*}$, thus not identifying some graphs of different orders, nor identifying finite graphs with some limit objects in $\mathcal{U}_\infty$.

For every fixed graph F, the functions $t(F,\cdot)$, $t_{\mathrm{inj}}(F,\cdot)$ and $t_{\mathrm{ind}}(F,\cdot)$ have unique continuous extensions to $\overline{\mathcal{U}}$, for which we use the same notation. We similarly extend $v(\cdot)^{-1}$ continuously to $\overline{\mathcal{U}}$ by defining $v(G)=\infty$ and thus $v(G)^{-1}=0$ for $G\in\mathcal{U}_\infty := \overline{\mathcal{U}}\setminus\mathcal{U}$. Then (2.4), (2.5) and (2.6) hold for all $G\in\overline{\mathcal{U}}$, where (2.4) means that

$$t_{\mathrm{inj}}(F,G)=t(F,G), \qquad G\in\mathcal{U}_\infty. \tag{2.10}$$

Note that $\overline{\mathcal{U}}$ is a compact metric space. Different, equivalent, metrics are given by the embeddings τ^+, τ_{inj}, τ_{ind} into $[0,1]^{\mathcal{U}^+}$ and $[0,1]^{\mathcal{U}}$. Another equivalent metric is, by Remark 2.1 and the definition of τ^+, $\delta_\square(G_1,G_2)+|v(G_1)^{-1}-v(G_2)^{-1}|$.

We summarize the results above on convergence.

Theorem 2.1. *A sequence (G_n) of graphs converges in the sense of Lovász and Szegedy [21] if and only if it converges in the compact metric space $\overline{\mathcal{U}^*}$. Moreover, if $v(G_n)\to\infty$, the sequence (G_n) converges in this sense if and only if it converges in $\overline{\mathcal{U}}$.*

The projection $\pi:[0,1]^{\mathcal{U}^+}=[0,1]^{\mathcal{U}}\times[0,1]\to[0,1]^{\mathcal{U}}$ maps $\tau^+(G)$ to $\tau(G)$ for every graph G, so by continuity it maps $\overline{\mathcal{U}}$ into $\overline{\mathcal{U}^*}$. For graph $G\in\mathcal{U}$, $\pi(G)=\tau(G)$ is the object in $\overline{\mathcal{U}^*}$ corresponding to G considered above, and we will in the sequel denote this object by $\pi(G)$; recall that this projection $\mathcal{U}\to\overline{\mathcal{U}^*}$ is not injective. (We thus distinguish between a graph G and its "ghost" $\pi(G)$ in $\overline{\mathcal{U}^*}$. Recall that when graphs are considered as elements of $\overline{\mathcal{U}^*}$ as in [21] and [8], certain graphs are identified with each other; we avoid this.) On the other hand, an element G of $\overline{\mathcal{U}}$ is by definition determined by $\tau(G)$ and $v(G)^{-1}$, cf. (2.9), so the restriction $\pi:\mathcal{U}_n\to\overline{\mathcal{U}^*}$ is injective for each $n\le\infty$. In particular, $\pi:\mathcal{U}_\infty\to\overline{\mathcal{U}^*}$ is injective. Moreover, this map is surjective because every element $G\in\overline{\mathcal{U}^*}$ is the limit of some sequence (G_n) of graphs in $\mathcal{U}$ with $v(G_n)\to\infty$; by Theorem 2.1, this sequence converges in $\overline{\mathcal{U}}$ to some element G', and then $\pi(G')=G$. Since $\mathcal{U}_\infty$ is compact, the restriction of π to $\mathcal{U}_\infty$ is thus a homeomorphism, and we have the following theorem, saying that we can identify the set $\mathcal{U}_\infty$ of proper graph limits with $\overline{\mathcal{U}^*}$.

Theorem 2.2. *The projection π maps the set $\mathcal{U}_\infty := \overline{\mathcal{U}}\setminus\mathcal{U}$ of proper graph limits homeomorphically onto $\overline{\mathcal{U}^*}$.*

3 – Convergence of random graphs

A *random unlabelled graph* is a random element of $\mathcal{U}$ (with any distribution; we do not imply any particular model). We consider convergence of a sequence (G_n) of random unlabelled graphs in the larger space $\overline{\mathcal{U}}$; recall that this is a compact metric space so we may use the general theory set forth in, for example, Billingsley [2].

We use the standard notations $\xrightarrow{\text{d}}$, $\xrightarrow{\text{p}}$, $\xrightarrow{\text{a.s.}}$ for convergence in distribution, probability, and almost surely, respectively. We will only consider the case when $v(G_n) \to \infty$, at least in probability. (The reader may think of the case when G_n has n vertices, although that is not necessary in general.)

We begin with convergence in distribution.

Theorem 3.1. *Let G_n, $n \geq 1$, be random unlabelled graphs and assume that $v(G_n) \xrightarrow{\text{p}} \infty$. The following are equivalent, as $n \to \infty$.*

(i) $G_n \xrightarrow{\text{d}} \Gamma$ *for some random* $\Gamma \in \overline{\mathcal{U}}$.
(ii) *For every finite family $F_1, \dots, F_m$ of (non-random) graphs, the random variables $t(F_1, G_n), \dots, t(F_m, G_n)$ converge jointly in distribution.*
(iii) *For every (non-random) $F \in \mathcal{U}$, the random variables $t(F, G_n)$ converge in distribution.*
(iv) *For every (non-random) $F \in \mathcal{U}$, the expectations $\mathbb{E}\, t(F, G_n)$ converge.*

If these properties hold, then the limits in (ii), (iii) and (iv) are $\big(t(F_i, \Gamma)\big)_{i=1}^m$, $t(F, \Gamma)$ and $\mathbb{E}\, t(F, \Gamma)$, respectively. Furthermore, $\Gamma \in \mathcal{U}_\infty$ a.s.

The same results hold if t is replaced by t_{inj} or t_{ind}.

Proof. (i) $\iff$ (ii). Since $\overline{\mathcal{U}}$ is a closed subset of $[0,1]^{\mathcal{U}^+}$, convergence in distribution in $\overline{\mathcal{U}}$ is equivalent to convergence of $\tau^+(G_n) = \big((t(F, G_n))_{F \in \mathcal{U}}, v(G_n)^{-1}\big)$ in $[0,1]^{\mathcal{U}^+}$, Since we assume $v(G_n)^{-1} \xrightarrow{\text{p}} 0$, this is equivalent to convergence of $(t(F, G_n))_{F \in \mathcal{U}}$ in $[0,1]^{\mathcal{U}}$ [2, Theorem 4.4], which is equivalent to convergence in distribution of all finite families $(t(F_i, G_n))_{i=1}^m$.

(ii) $\Longrightarrow$ (iii). Trivial.

(iii) $\Longrightarrow$ (iv). Immediate, since t is bounded (by 1).

(iv) $\Longrightarrow$ (ii). Let $F_1, \dots, F_m$ be fixed graphs and let $\ell_1, \dots, \ell_m$ be positive integers. Let F be the disjoint union of ℓ_i copies of F_i, $i = 1, \dots, m$. Then, for every $G \in \mathcal{U}$, from the definition of t,

$$t(F, G) = \prod_{i=1}^m t(F_i, G)^{\ell_i},$$

and hence

$$\mathbb{E}\prod_{i=1}^{m} t(F_i, G)^{\ell_i} = \mathbb{E}\, t(F, G). \tag{3.1}$$

Consequently, if (iv) holds, then every joint moment $\mathbb{E}\prod_{i=1}^{m} t(F_i, G)^{\ell_i}$ of $t(F_1, G_n), \dots, t(F_m, G_n)$ converges. Since $t(F_i, G_n)$ are bounded (by 1), this implies joint convergence in distribution by the method of moments.

The identification of the limits is immediate. Since $v(G_n) \xrightarrow{\text{p}} \infty$,(i) implies that $v(\Gamma) = \infty$ a.s., and thus $\Gamma \in \mathcal{U}_\infty$.

Finally, it follows from (2.4), (2.5) and (2.6) that we can replace t by t_{inj} or t_{ind} in (ii) and (iv), and the implications (ii) $\Longrightarrow$ (iii) and (iii) $\Longrightarrow$ (iv) are immediate for t_{inj} and t_{ind} too. ∎

Specializing to the case of a non-random limit $G \in \mathcal{U}_\infty$, we obtain the corresponding result for convergence in probability.

Corollary 3.2. *Let G_n, $n \geq 1$, be random unlabelled graphs such that $v(G_n) \xrightarrow{\text{p}} \infty$, and let $G \in \mathcal{U}_\infty$. The following are equivalent, as $n \to \infty$.*

(i) $G_n \xrightarrow{\text{p}} G$.
(ii) $t(F, G_n) \xrightarrow{\text{p}} t(F, G)$ *for every (non-random)* $F \in \mathcal{U}$.
(iii) $\mathbb{E}\, t(F, G_n) \to t(F, G)$ *for every (non-random)* $F \in \mathcal{U}$.

The same result holds if t is replaced by t_{inj} or t_{ind}.

Note further that under the same assumptions, it follows directly from Theorem 2.1 that $G_n \xrightarrow{\text{a.s.}} G$ if and only if $t(F, G_n) \xrightarrow{\text{a.s.}} t(F, G)$ for every $F \in \mathcal{U}$.

We observe another corollary to Theorem 3.1 (and its proof).

Corollary 3.3. *If Γ is a random element of $\mathcal{U}_\infty = \overline{\mathcal{U}} \setminus \mathcal{U} \cong \overline{\mathcal{U}^*}$, then, for every sequence $F_1, \dots, F_m$ of graphs, possibly with repetitions,*

$$\mathbb{E}\prod_{i=1}^{m} t(F_i, \Gamma) = \mathbb{E}\, t\left(\oplus_{i=1}^{m} F_i, \Gamma\right), \tag{3.2}$$

where $\oplus_{i=1}^{m} F_i$ denotes the disjoint union of $F_1, \dots, F_m$. As a consequence, the distribution of Γ is uniquely determined by the numbers $\mathbb{E}\, t(F, \Gamma)$, $F \in \mathcal{U}$. Alternatively, the distribution of Γ is uniquely determined by the numbers $\mathbb{E}\, t_{\text{ind}}(F, \Gamma)$, $F \in \mathcal{U}$.

Proof. Since $\mathcal{U}$ is dense in $\overline{\mathcal{U}} \supseteq \mathcal{U}_\infty$, there exists random unlabelled graphs G_n such that $G_n \xrightarrow{\text{a.s.}} \Gamma$. In particular, $G_n \xrightarrow{\text{d}} \Gamma$ and $v(G_n) \xrightarrow{\text{p}} \infty$ (in fact, we may assume $v(G_n) = n$), so Theorem 3.1 and its proof apply, and (3.2) follows from (3.1) applied to G_n by letting $n \to \infty$.

For the second statement, note that (3.2) shows that the expectations $\mathbb{E}\, t(F,\Gamma)$, $F \in \mathcal{U}$, determine all moments $\mathbb{E} \prod_{i=1}^m t(F_i,\Gamma)$, and thus the joint distribution of $t(F,\Gamma)$, $F \in \mathcal{U}$, which is the same as the distribution of $\tau(\Gamma) = \big(t(F,\Gamma)\big)_{F\in\mathcal{U}} \in [0,1]^{\mathcal{U}}$, and we have defined $\mathcal{U}_\infty$ such that we identify Γ and $\tau(\Gamma)$. Finally, the numbers $\mathbb{E}\, t_{\text{ind}}(F,\Gamma)$, $F \in \mathcal{U}$, determine all $\mathbb{E}\, t(F,\Gamma)$ by (2.5), recalling that $t_{\text{inj}}(F,\Gamma) = t(F,\Gamma)$ by (2.10). ∎

Remark 3.1. The numbers $\mathbb{E}\, t(F,\Gamma)$ for a random $\Gamma \in \mathcal{U}_\infty$ thus play a role similar to the one played by moments for a random variable. (And the relation between $\mathbb{E}\, t(F,\Gamma)$ and $\mathbb{E}\, t_{\text{ind}}(F,\Gamma)$ has some resemblance to the relation between moments and cumulants.)

4 – Convergence to infinite graphs

We will in this section consider also labelled *infinite* graphs with the vertex set $\mathbb{N} = \{1,2,\dots\}$. Let $\mathcal{L}_\infty$ denote the set of all such graphs. These graphs are determined by their edge sets, so $\mathcal{L}_\infty$ can be identified with the power set $\mathcal{P}(E(K_\infty))$ of all subsets of the edge set $E(K_\infty)$ of the complete infinite graph K_∞, and thus with the infinite product set $\{0,1\}^{E(K_\infty)}$. We give this space, and thus $\mathcal{L}_\infty$, the product topology. Hence, $\mathcal{L}_\infty$ is a compact metric space.

It is sometimes convenient to regard $\mathcal{L}_n$ for a finite n as a subset of $\mathcal{L}_\infty$: we can identify graphs in $\mathcal{L}_n$ and $\mathcal{L}_\infty$ with the same edge set. In other words, if $G \in \mathcal{L}_n$ is a graph with vertex set $[n]$, we add an infinite number of isolated vertices $n+1, n+2, \dots$ to obtain a graph in H.

Conversely, if $H \in \mathcal{L}_\infty$ is an infinite graph, we let $H|_{[n]} \in \mathcal{L}_n$ be the induced subgraph of H with vertex set $[n]$.

If G is a (finite) graph, let $\widehat{G}$ be the random labelled graph obtained by a random labelling of the vertices of G by the numbers $1,\dots,v(G)$. (If G is labelled, we thus ignore the labels and randomly relabel.) Thus $\widehat{G}$ is a random finite graph with the same number of vertices as G, but as just said, we can (and will) also regard $\widehat{G}$ as a random graph in $\mathcal{L}_\infty$.

We use the same notation $\widehat{G}$ also for a random (finite) graph G given a random labelling.

Theorem 4.1. *Let (G_n) be a sequence of random graphs in $\mathcal{U}$ and assume that $v(G_n) \xrightarrow{\text{p}} \infty$. Then the following are equivalent.*

(i) $G_n \xrightarrow{\text{d}} \Gamma$ *in* $\overline{\mathcal{U}}$ *for some random* $\Gamma \in \overline{\mathcal{U}}$.
(ii) $\widehat{G_n} \xrightarrow{\text{d}} H$ *in* $\mathcal{L}_\infty$ *for some random* $H \in \mathcal{L}_\infty$.

If these hold, then $\mathbb{P}(H|_{[k]} = F) = \mathbb{E}\, t_{\mathrm{ind}}(F, \Gamma)$ *for every* $F \in \mathcal{L}_k$. *Furthermore,* $\Gamma \in \mathcal{U}_\infty$ *a.s.*

PROOF. Let G be a labelled graph and consider the graph $\widehat{G}|_{[k]}$, assuming $k \le v(G)$. This random graph equals $G[k]' = G(v_1', \dots, v_k')$, where $v_1', \dots, v_k'$ are k vertices sampled at random without replacement as in Section 2. Hence, by (2.3), for every $F \in \mathcal{L}_k$,

$$\mathbb{P}(\widehat{G}|_{[k]} = F) = t_{\mathrm{ind}}(F, G), \qquad \text{if } k \le v(G).$$

Applied to the random graph G_n, this yields

$$\mathbb{E}\, t_{\mathrm{ind}}(F, G_n) \le \mathbb{P}(\widehat{G_n}|_{[k]} = F) \le \mathbb{E}\, t_{\mathrm{ind}}(F, G_n) + \mathbb{P}\big(v(G_n) < k\big). \tag{4.1}$$

By assumption, $\mathbb{P}\left(v(G_n) < k\right) \to 0$ as $n \to \infty$, and it follows from (4.1) and Theorem 3.1 that $G_n \xrightarrow{\mathrm{d}} \Gamma$ in $\overline{\mathcal{U}}$ if and only if

$$\mathbb{P}(\widehat{G_n}|_{[k]} = F) \to \mathbb{E}\, t_{\mathrm{ind}}(F, \Gamma) \tag{4.2}$$

for every $k \ge 1$ and every $F \in \mathcal{L}_k$.

Since $\mathcal{L}_k$ is a finite set, (4.2) says that, for every k, $\widehat{G_n}|_{[k]} \xrightarrow{\mathrm{d}} H_k$ for some random graph $H_k \in \mathcal{L}_k$ with $\mathbb{P}(H_k = F) = \mathbb{E}\, t_{\mathrm{ind}}(F, \Gamma)$ for $F \in \mathcal{L}_k$. Since $\mathcal{L}_\infty$ has the product topology, this implies $\widehat{G_n} \xrightarrow{\mathrm{d}} H$ in $\mathcal{L}_\infty$ for some random $H \in \mathcal{L}_\infty$ with $H|_{[k]} \overset{\mathrm{d}}{=} H_k$.

Conversely, if $\widehat{G_n} \xrightarrow{\mathrm{d}} H$ in $\mathcal{L}_\infty$, then $\widehat{G_n}|_{[k]} \xrightarrow{\mathrm{d}} H|_{[k]}$ so the argument above shows that

$$\mathbb{E}\, t_{\mathrm{ind}}(F, G_n) = \mathbb{P}(\widehat{G_n}|_{[k]} = F) + o(1) \to \mathbb{P}(H|_{[k]} = F)$$

as $n \to \infty$, for every $F \in \mathcal{L}_k$, and Theorem 3.1 yields the existence of some random $\Gamma \in \mathcal{U}_\infty \subset \overline{\mathcal{U}}$ with $G_n \xrightarrow{\mathrm{d}} \Gamma$ and $\mathbb{E}\, t_{\mathrm{ind}}(F, \Gamma) = \mathbb{P}(H|_{[k]} = F)$. ∎

5 – Exchangeable random graphs

DEFINITION. A random infinite graph $H \in \mathcal{L}_\infty$ is *exchangeable* if its distribution is invariant under every permutation of the vertices. (It is well-known that it is equivalent to consider only finite permutations, i.e., permutations σ of $\mathbb{N}$ that satisfy $\sigma(i) = i$ for all sufficiently large i, so σ may be regarded as a permutation in $\mathfrak{S}_n$ for some n.)

Equivalently, if $X_{ij} := \mathbf{1}[ij \in H]$ is the indicator of there being an edge ij in H, then the array $\{X_{ij}\}$, $1 \le i, j \le \infty$, is (jointly) exchangeable as defined in Section 1.

LEMMA 5.1. *Let H be a random infinite graph in $\mathcal{L}_\infty$. Then the following are equivalent.*

(i) *H is exchangeable.*
(ii) *$H|_{[k]}$ has a distribution invariant under all permutations of $[k]$, for every $k \geq 1$.*
(iii) *$\mathbb{P}\big(H|_{[k]} = F\big)$ depends only on the isomorphism type of F, and can thus be seen as a function of F as an unlabelled graph in $\mathcal{U}_k$, for every $k \geq 1$.*

PROOF. (i) $\Longrightarrow$ (ii). Immediate.

(ii) $\Longrightarrow$ (i). If σ is a finite permutation of $\mathbb{N}$, then σ restricts to a permutation of $[k]$ for every large k, and it follows that if $H \circ \sigma$ is H with the vertices permuted by σ, then, for all large k $H \circ \sigma|_{[k]} = H|_{[k]} \circ \sigma \overset{\mathrm{d}}{=} H|_{[k]}$, which implies $H \circ \sigma \overset{\mathrm{d}}{=} H$.

(ii) $\Longleftrightarrow$ (iii). Trivial. ∎

THEOREM 5.2. *The limit H in Theorem 4.1 is exchangeable.*

PROOF. H satisfies Lemma 5.1(iii). ∎

Moreover, Theorem 4.1 implies the following connection with random elements of $\mathcal{U}_\infty$.

THEOREM 5.2. *There is a one-to-one correspondence between distributions of random elements $\Gamma \in \mathcal{U}_\infty$ (or $\overline{\mathcal{U}^*}$) and distributions of exchangeable random infinite graphs $H \in \mathcal{L}_\infty$ given by*

$$\mathbb{E}\, t_{\mathrm{ind}}(F, \Gamma) = \mathbb{P}(H|_{[k]} = F) \tag{5.1}$$

for every $k \geq 1$ and every $F \in \mathcal{L}_k$, or, equivalently,

$$\mathbb{E}\, t(F, \Gamma) = \mathbb{P}(H \supset F) \tag{5.2}$$

for every $F \in \mathcal{L}$. Furthermore, $H|_{[n]} \xrightarrow{\mathrm{d}} \Gamma$ in $\overline{\mathcal{U}}$ as $n \to \infty$.

PROOF. Note first that (5.1) and (5.2) are equivalent by (2.5) and (2.6), since $t(F, \Gamma) = t_{\mathrm{inj}}(F, \Gamma)$ by (2.10), and $H \supset F$ if and only if $H|_{[k]} \supseteq F$ when $F \in \mathcal{L}_k$.

Suppose that Γ is a random element of $\mathcal{U}_\infty \subset \overline{\mathcal{U}}$. Since $\mathcal{U}$ is dense in $\overline{\mathcal{U}}$, there exist (as in the proof of Corollary 3.3) random unlabelled graphs G_n such that $G_n \xrightarrow{\mathrm{a.s.}} \Gamma$ in $\overline{\mathcal{U}}$ and thus $v(G_n) \xrightarrow{\mathrm{a.s.}} \infty$ and $G_n \xrightarrow{\mathrm{d}} \Gamma$. Hence, Theorems 4.1 and 5.2 show that $\widehat{G_n} \xrightarrow{\mathrm{d}} H$ for some random exchangeable infinite graph H

satisfying (5.1). Furthermore, (5.1) determines the distribution of $H|_{[k]}$ for every k, and thus the distribution of H.

Conversely, if H is an exchangeable random infinite graph, let $G_n = H|_{[n]}$. By Lemma 5.1(ii), the distribution of each G_n is invariant under permutations of the vertices, so if $\widehat{G_n}$ is G_n with a random (re)labelling, we have $\widehat{G_n} \overset{\mathrm{d}}{=} G_n$. Since $G_n \xrightarrow{\mathrm{d}} H$ in $\mathcal{L}_\infty$ (because $\mathcal{L}_\infty$ has a product topology), we thus have $\widehat{G_n} \xrightarrow{\mathrm{d}} H$ in $\mathcal{L}_\infty$, so Theorem 4.1 applies and shows the existence of a random $\Gamma \in \mathcal{U}_\infty$ such that $G_n \xrightarrow{\mathrm{d}} \Gamma$ and (5.1) holds. Finally (5.1) determines the distribution of Γ by Corollary 3.3. ∎

Remark 5.1. Moreover, $H|_{[n]}$ converges a.s. to some random variable $\Gamma \in \mathcal{U}_\infty$, because $t_{\mathrm{ind}}(F, H|_{[n]})$, $n \geq v(F)$, is a reverse martingale for every $F \in \Gamma$. Alternatively, this follows by concentration estimates from the representation in Section 6, see Lovász and Szegedy [21,Theorem 2.5].

Corollary 5.4. *There is a one-to-one correspondence between elements Γ of $\mathcal{U}_\infty \cong \overline{\mathcal{U}^*}$ and extreme points of the set of distributions of exchangeable random infinite graphs $H \in \mathcal{L}_\infty$. This correspondence is given by*

$$t(F, \Gamma) = \mathbb{P}(H \supset F) \tag{5.3}$$

for every $F \in \mathcal{L}$. Furthermore, $H|_{[n]} \xrightarrow{\text{a.s.}} \Gamma$ in $\overline{\mathcal{U}}$ as $n \to \infty$.

Proof. The extreme points of the set of distributions on $\mathcal{U}_\infty$ are the point masses, which are in one-to-one correspondence with the elements of $\mathcal{U}_\infty$. ∎

We can characterize these extreme point distributions of exchangeable random infinite graphs as follows.

Theorem 5.5. *Let H be an exchangeable random infinite graph. Then the following are equivalent.*

(i) *The distribution of H is an extreme point in the set of exchangeable distributions in $\mathcal{L}_\infty$.*
(ii) *If F_1 and F_2 are two (finite) graphs with disjoint vertex sets $V(F_1)$, $V(F_2) \subset \mathbb{N}$, then*
$$\mathbb{P}(H \supset F_1 \cup F_2) = \mathbb{P}(H \supset F_1)\, \mathbb{P}(H \supset F_2).$$
(iii) *The restrictions $H|_{[k]}$ and $H|_{[k+1,\infty)}$ are independent for every k.*
(iv) *Let $\mathcal{F}_n$ be the σ-field generated by $H|_{[n,\infty)}$. Then the tail σ-field $\bigcap_{n=1}^\infty \mathcal{F}_n$ is trivial, i.e., contains only events with probability 0 or 1.*

Proof. (i) $\Longrightarrow$ (ii). By Corollary 5.4, H corresponds to some (non-random) $\Gamma \in \mathcal{U}_\infty$ such that

$$\mathbb{P}(H \supset F) = t(F, \Gamma) \tag{5.4}$$

for every $F \in \mathcal{L}$. We have defined $\mathcal{L}$ such that a graph $F \in \mathcal{L}$ is labelled by $1, \dots, v(F)$, but both sides of (5.4) are invariant under relabelling of F by arbitrary positive integers; the left hand side because H is exchangeable and the right hand side because $t(F, \Gamma)$ only depends on F as an unlabelled graph. Hence (5.4) holds for every finite graph F with $V(F) \subset \mathbb{N}$.

Furthermore, since Γ is non-random, Corollary 3.3 yields $t(F_1 \cup F_2, \Gamma) = t(F_1, \Gamma)t(F_2, \Gamma)$. Hence,

$$\mathbb{P}(H \supset F_1 \cup F_2) = t(F_1 \cup F_2, \Gamma) = t(F_1, \Gamma)t(F_2, \Gamma) = \mathbb{P}(H \supset F_1)\,\mathbb{P}(H \supset F_2).$$

(ii) $\Longrightarrow$ (iii). By inclusion–exclusion, as for (2.6), (ii) implies that if $1 \le k < l < \infty$, then for any graphs F_1 and F_2 with $V(F_1) = \{1, \dots, k\}$ and $V(F_2) = \{k+1, \dots, k+l\}$, the events $H|_{[k]} = F_1$ and $H|_{\{k+1,\dots,l\}} = F_2$ are independent. Hence $H|_{[k]}$ and $H|_{\{k,\dots,l\}}$ are independent for every $l > k$, and the result follows.

(iii) $\Longrightarrow$ (iv). Suppose A is an event in the tail σ-field $\bigcap_{n=1}^{\infty} \mathcal{F}_n$. Let $\mathcal{F}_n^*$ be the σ-field generated by $H|_{[n]}$. By (iii), A is independent of $\mathcal{F}_n^*$ for every n, and thus of the σ-field $\mathcal{F}$ generated by $\bigcup \mathcal{F}_n^*$, which equals the σ-field $\mathcal{F}_1$ generated by H. However, $A \in \mathcal{F}_1$, so A is independent of itself and thus $\mathbb{P}(A) = 0$ or 1.

(iv) $\Longrightarrow$ (i). Let $F \in \mathcal{L}_k$ for some k and let F_n be F with all vertices shifted by n. Consider the two indicators $I = \mathbf{1}[H \supseteq F]$ and $I_n = \mathbf{1}[H \supseteq F_n]$. Since I_n is $\mathcal{F}_n$-measurable,

$$\mathbb{P}(H \supset F \cup F_n) = \mathbb{E}(I I_n) = \mathbb{E}\big(\mathbb{E}(I \mid \mathcal{F}_n) I_n\big). \tag{5.5}$$

Moreover, $\mathbb{E}(I \mid \mathcal{F}_n)$, $n = 1, 2, \dots$, is a reverse martingale, and thus a.s.

$$\mathbb{E}(I \mid \mathcal{F}_n) \to \mathbb{E}\Big(I \mid \bigcap_{n=1}^{\infty} \mathcal{F}_n\Big) = \mathbb{E}\, I,$$

using (iv). Hence, $\big(\mathbb{E}(I \mid \mathcal{F}_n) - \mathbb{E}\, I\big) I_n \to 0$ a.s., and by dominated convergence

$$\mathbb{E}\Big(\big(\mathbb{E}(I \mid \mathcal{F}_n) - \mathbb{E}\, I\big) I_n\Big) \to 0.$$

Consequently, (5.5) yields

$$\mathbb{P}(H \supset F \cup F_n) = \mathbb{E}\, I\, \mathbb{E}\, I_n + o(1) = \mathbb{P}(H \supset F)\,\mathbb{P}(H \supset F_n) + o(1).$$

Moreover, since H is exchangeable, $\mathbb{P}(H \supset F \cup F_n)$ (for $n \geq v(F)$) and $\mathbb{P}(H \supset F_n)$ do not depend on n, and we obtain as $n \to \infty$

$$\mathbb{P}(H \supset F \cup F_k) = \mathbb{P}(H \supset F)^2. \tag{5.6}$$

Let Γ be a random element of $\mathcal{U}_\infty$ corresponding to H as in Theorem 5.3. By (5.2) and (3.2), (5.6) can be written

$$\mathbb{E}\, t(F,\Gamma)^2 = \bigl(\mathbb{E}\, t(F,\Gamma)\bigr)^2.$$

Hence the random variable $t(F,\Gamma)$ has variance 0 so it is a.s. constant. Since this holds for every $F \in \mathcal{L}$, it follows that Γ is a.s. constant, i.e., we can take Γ non-random, and (i) follows by Corollary 5.4. ∎

6 – Representations of graph limits and exchangeable graphs

As said in the introduction, the exchangeable infinite random graphs were characterized by Aldous [1] and Hoover [16], see also Kallenberg [17], and the graph limits in $\mathcal{U}_\infty \cong \overline{\mathcal{U}^*}$ were characterized in a very similar way by Lovász and Szegedy [21]. We can now make the connection between these two characterizations explicit.

Let $\mathcal{W}$ be the set of all measurable functions $W : [0,1]^2 \to [0,1]$ and let $\mathcal{W}_\mathsf{s}$ be the subset of symmetric functions. For every $W \in \mathcal{W}_\mathsf{s}$, we define an infinite random graph $G(\infty, W) \in \mathcal{L}_\infty$ as follows: we first choose a sequence $X_1, X_2, \dots$ of i.i.d. random variables uniformly distributed on $[0,1]$, and then, given this sequence, for each pair (i,j) with $i < j$ we draw an edge ij with probability $W(X_i, X_j)$, independently for all pairs (i,j) with $i < j$ (conditionally given $\{X_k\}$). Further, let $G(n,W)$ be the restriction $G(\infty,W)|_{[n]}$, which is obtained by the same construction with a finite sequence $X_1, \dots, X_n$.

It is evident that $G(\infty,W)$ is an exchangeable infinite random graph, and the result by Aldous and Hoover is that every exchangeable infinite random graph is obtained as a mixture of such $G(\infty,W)$; in other words as $G(\infty,W)$ with a random W.

Considering again a deterministic $W \in \mathcal{W}_\mathsf{s}$, it is evident that Theorem 5.5(ii) holds, and thus Theorem 5.5 and Corollary 5.4 show that $G(\infty,W)$ corresponds to an element $\Gamma_W \in \mathcal{U}_\infty$. Moreover, by Theorem 5.3 and Remark 5.1, $G(n,W) \to \Gamma_W$ a.s. as $n \to \infty$, and (5.3) shows that if $F \in \mathcal{L}_k$, then

$$t(F,\Gamma_W) = \mathbb{P}\bigl(F \subseteq G(k,W)\bigr) = \int_{[0,1]^k} \prod_{ij \in \mathbb{E}(F)} W(x_i,x_j)\, \mathrm{d}x_1 \dots \mathrm{d}x_k. \tag{6.1}$$

The main result of Lovász and Szegedy [21] is that every element of $\mathcal{U}_\infty \cong \overline{\mathcal{U}^*}$ can be obtained as Γ_W satisfying (6.1) for some $W \in \mathcal{W}_\mathsf{s}$.

It is now clear that the representation theorems of Aldous–Hoover [1, 16] and Lovász and Szegedy [21] are connected by Theorem 5.3 and Corollary 5.4 above, and that one characterization easily follows from the other.

Remark 6.1. The representations by W are far from unique, see Section 7. Borgs, Chayes, Lovász, Sós and Vesztergombi [8] call an element $W \in \mathcal{W}_{\mathsf{s}}$ a *graphon*. They further define a pseudometric (called the *cut-distance*) on $\mathcal{W}_{\mathsf{s}}$ and show that if we consider the quotient space $\widehat{\mathcal{W}_{\mathsf{s}}}$ obtained by identifying elements with cut-distance 0, we obtain a compact metric space, and the mapping $W \mapsto \Gamma_W$ yields a bijection $\widehat{\mathcal{W}_{\mathsf{s}}} \to \overline{\mathcal{U}^*} \cong \mathcal{U}_\infty$, which furthermore is a homeomorphism.

Remark 6.2. As remarked in Lovász and Szegedy [21], we can more generally consider a symmetric measurable function $W : \mathcal{S}^2 \to [0,1]$ for any probability space $(\mathcal{S}, \mu)$, and define $G(\infty, W)$ as above with X_i i.i.d. random variables in $\mathcal{S}$ with distribution μ. This does not give any new limit objects $G(\infty, W)$ or Γ_W, since we just said that every limit object is obtained from some $W \in \mathcal{W}_{\mathsf{s}}$, but they can sometimes give useful representations.

An interesting case is when W is the adjacency matrix of a (finite) graph G, with $\mathcal{S} = V(G)$ and μ the uniform measure on $\mathcal{S}$; we thus let X_i be i.i.d. random vertices of G and $G(n, W)$ equals the random graph $G[n]$ defined in Section 2. It follows from (6.1) and (2.1) that $t(F, \Gamma_W) = t(F, G)$ for every $F \in \mathcal{U}$, and thus $\Gamma_W = G$ as elements of $\overline{\mathcal{U}^*}$. In other words, $\Gamma_W \in \mathcal{U}_\infty$ equals $\pi(G)$, the "ghost" of G in $\mathcal{U}_\infty \cong \overline{\mathcal{U}^*}$.

Remark 6.3. For the asymptotic behavior of $G(n, W)$ in another, sparse, case, with W depending on n, see [3].

7 – Non-uniqueness

The functions W on $[0,1]^2$ used to represent graph limits or exchangeable arrays are far from unique. (For a special case when there is a natural canonical choice, which much simplifies and helps applications, see [14].) For example, it is obvious that if $\varphi : [0,1] \to [0,1]$ is any measure preserving map, then W and $W \circ \varphi$, defined by $W \circ \varphi(x, y) := W\big(\varphi(x), \varphi(y)\big)$, define the same graph limit and the same (in distribution) exchangeable array.

Although in principle, this is the only source on non-uniqueness, the details are more complicated, mainly because the measure preserving map φ does not have to be a bijection, and thus the relation $W' = W \circ \varphi$ is not symmetric: it can hold without there being a measure preserving map φ' such that $W = W' \circ \varphi'$. (For a 1-dimensional example, consider $f(x) = x$ and $f'(x) = \varphi(x) = 2x \bmod 1$; for a 2-dimensional example, let $W(x, y) = f(x)f(y)$ and $W'(x, y) = f'(x)f'(y)$.)

For exchangeable arrays, the equivalence problem was solved by Hoover [16], who gave a criterion which in our case reduces to (vi) below; this criterion involves an auxiliary variable, and can be interpreted as saying $W = W' \circ \varphi'$ for a random

φ'. This work was continued by Kallenberg, see [17], who gave a probabilistic proof and added criterion (v). For graph limits, Borgs, Chayes, Lovász, Sós and Vesztergombi [8] gave the criterion (vii) in terms of the cut-distance, and Bollobás and Riordan [4] found the criterion (v) in this context. Further, Borgs, Chayes, Lovász, Sós and Vesztergombi [8] announced the related criterion that there exists a measurable function $U : [0,1]^2 \to [0,1]$ and two measure preserving maps $\varphi, \varphi' : [0,1] \to [0,1]$ such that $W = U \circ \varphi$ and $W' = U \circ \varphi'$ a.e.; the proof of this will appear in [5].

As in Section 6, these two lines of work are connected by the results in Section 5, and we can combine the previous results as follows.

THEOREM 7.1. *Let $W, W' \in \mathcal{W}_s$. Then the following are equivalent.*

(i) $\Gamma_W = \Gamma_{W'}$ *for the graph limits* $\Gamma_W, \Gamma_{W'} \in \mathcal{U}_\infty$.
(ii) $t(F, \Gamma_W) = t(F, \Gamma_{W'})$ *for every graph* F.
(iii) *The exchangeable random infinite graphs* $G(\infty, W)$ *and* $G(\infty, W')$ *have the same distribution.*
(iv) *The random graphs* $G(n, W)$ *and* $G(n, W')$ *have the same distribution for every finite* n.
(v) *There exist measure preserving maps* $\varphi, \varphi' : [0,1] \to [0,1]$ *such that* $W \circ \varphi = W' \circ \varphi'$ *a.e. on* $[0,1]^2$, *i.e.,* $W\big(\varphi(x), \varphi(y)\big) = W'\big(\varphi'(x), \varphi'(y)\big)$ *a.e.*
(vi) *There exists a measure preserving map* $\psi : [0,1]^2 \to [0,1]$ *such that* $W(x_1, x_2) = W'\big(\psi(x_1, y_1), \psi(x_2, y_2)\big)$ *a.e. on* $[0,1]^4$.
(vii) $\delta_\square(W, W') = 0$, *where* $\delta_\square$ *is the cut-distance defined in [8].*

PROOF. (i) $\iff$ (ii). By our definition of $\mathcal{U}_\infty \subset \overline{\mathcal{U}}$.

(i) $\iff$ (iii). By Corollary 5.4.

(iii) $\iff$ (ii). Obvious.

(v) $\implies$ (iii). If $X_1, X_2, \dots$ are i.i.d. random variables uniformly distributed on $[0,1]$, then so are $\varphi(X_1), \varphi(X_2), \dots$, and thus $G(\infty, W) \overset{\mathrm{d}}{=} G(\infty, W \circ \varphi) = G(\infty, W' \circ \varphi') \overset{\mathrm{d}}{=} G(\infty, W')$.

(iii) $\implies$ (v). The general form of the representation theorem as stated in [17, Theorem 7.15, see also p. 304] is (in our two-dimensional case) $X_{ij} = f(\xi_\emptyset, \xi_i, \xi_j, \xi_{ij})$ for a function $f : [0,1]^4 \to [0,1]$, symmetric in the two middle variables, and independent random variables $\xi_\emptyset$, ξ_i $(1 \le i)$ and ξ_{ij} $(1 \le i < j)$, all uniformly distributed on [0,1], and where we further let $\xi_{ji} = \xi_{ij}$ for $j > i$. We can write the construction of $G(\infty, W)$ in this form with

$$f(\xi_\emptyset, \xi_i, \xi_j, \xi_{ij}) = \mathbf{1}[\xi_{ij} \le W(\xi_i, \xi_j)]. \tag{7.1}$$

Note that this f does not depend on $\xi_\emptyset$. (In general, $\xi_\emptyset$ is needed for the case of a random W, which can be written as a deterministic function of $\xi_\emptyset$, but this is not needed in the present theorem.)

Suppose that $G(\infty, W) \overset{\mathrm{d}}{=} G(\infty, W')$, let f be given by W by (7.1), and let similarly f' be given by W'; for notational convenience we write $W_1 := W$, $W_2 := W'$, $f_1 := f$ and $f_2 := f'$. The equivalence theorem [17, Theorem 7.28] takes the form, using (7.1), that there exist measurable functions $g_{k,0} : [0,1] \to [0,1]$, $g_{k,1} : [0,1]^2 \to [0,1]$ and $g_{k,2} : [0,1]^4 \to [0,1]$, for $k = 1,2$, that are measure preserving in the last coordinate for any fixed values of the other coordinates, and such that the two functions (for $k = 1$ and $k = 2$)

$$
\begin{aligned}
& f_k\big(g_{k,0}(\xi_\emptyset), g_{k,1}(\xi_\emptyset, \xi_1), g_{k,1}(\xi_\emptyset, \xi_2), g_{k,2}(\xi_\emptyset, \xi_1, \xi_2, \xi_{12})\big) = \\
& = \mathbf{1}\big[W_k\big(g_{k,1}(\xi_\emptyset, \xi_1), g_{k,1}(\xi_\emptyset, \xi_2)\big) \ge g_{k,2}(\xi_\emptyset, \xi_1, \xi_2, \xi_{12})\big]
\end{aligned}
$$

are a.s. equal. Conditioned on $\xi_\emptyset, \xi_1$ and ξ_2, the random variable $g_{k,2}(\xi_\emptyset, \xi_1, \xi_2, \xi_{12})$ is uniformly distributed on $[0,1]$, and it follows (e.g., by taking the conditional expectation) that a.s.

$$
W_1\big(g_{1,1}(\xi_\emptyset, \xi_1), g_{1,1}(\xi_\emptyset, \xi_2)\big) = W_2\big(g_{2,1}(\xi_\emptyset, \xi_1), g_{2,1}(\xi_\emptyset, \xi_2)\big).
$$

For a.e. value x_0 of $\xi_\emptyset$, this thus holds for a.e. values of ξ_1 and ξ_2, and we may choose $\varphi(x) = g_{1,1}(x_0, x)$ and $\varphi'(x) := g_{2,1}(x_0, x)$ for some such x_0.

(iii) $\iff$ (vi). Similar, using [17, Theorem 7.28(iii)].

(ii) $\iff$ (vii). See [8]. ∎

8 – Bipartite graphs

The definitions and results above have analogues for bipartite graphs, which we give in this section, leaving some details to the reader. The proofs are straightforward analogues of the ones given above and are omitted. Applications of the results of this section to random difference graphs are in [14].

A *bipartite graph* will be a graph with an explicit bipartition; in other words, a bipartite graph G consists of two vertex sets $V_1(G)$ and $V_2(G)$ and an edge set $E(G) \subseteq V_1(G) \times V_2(G)$; we let $v_1(G) := |V_1(G)|$ and $v_2(G) := |V_2(G)|$ be the numbers of vertices in the two sets. Again we consider both the labelled and unlabelled cases; in the labelled case we assume the labels of the vertices in $V_j(G)$ are $1, \dots, v_j(G)$ for $j = 1,2$. Let $\mathcal{B}^L_{n_1 n_2}$ be the set of the $2^{n_1 n_2}$ labelled bipartite graphs with vertex sets $[n_1]$ and $[n_2]$, and let $\mathcal{B}_{n_1 n_2}$ be the quotient set $\mathcal{B}^L_{n_1 n_2}/\cong$ of unlabelled bipartite graphs with n_1 and n_2 vertices in the two parts; further, let $\mathcal{B}^L := \bigcup_{n_1,n_2 \ge 1} \mathcal{B}^L_{n_1 n_2}$ and $\mathcal{B} := \bigcup_{n_1,n_2 \ge 1} \mathcal{B}_{n_1 n_2}$.

We let $G[k_1, k_2]$ be the random graph in $\mathcal{B}^L_{k_1 k_2}$ obtained by sampling k_j vertices from $V_j(G)$ $(j = 1,2)$, uniformly with replacement, and let, provided $k_j \le v_j(G)$, $G[k_1, k_2]'$ be the corresponding random graph obtained by sampling without replacement. We then define $t(F,G)$, $t_{\mathrm{inj}}(F,G)$ and $t_{\mathrm{ind}}(F,G)$ for (unlabelled) bipartite graphs F and G in analogy with (2.1)–(2.3). Then (2.4)–(2.6) still hold, *mutatis mutandis*; for example,

$$\left|t(F,G) - t_{\mathrm{inj}}(F,G)\right| \le \frac{v_1(F)^2}{2v_1(G)} + \frac{v_2(F)^2}{2v_2(G)}. \tag{8.1}$$

In analogy with (2.7), we now define $\tau : \mathcal{B} \to [0,1]^{\mathcal{B}}$ by

$$\tau(G) := (t(F,G))_{F\in\mathcal{B}} \in [0,1]^{\mathcal{B}}. \tag{8.2}$$

We define $\mathcal{B}^* := \tau(\mathcal{B}) \subseteq [0,1]^{\mathcal{B}}$ to be the image of $\mathcal{B}$ under this mapping τ, and let $\overline{\mathcal{B}^*}$ be the closure of $\mathcal{B}^*$ in $[0,1]^{\mathcal{B}}$; this is a compact metric space.

Again, τ is not injective; we may consider a graph G as an element of $\mathcal{B}^*$ by identifying G and $\tau(G)$, but this implies identification of some graphs of different orders and we prefer to avoid it. We let $\mathcal{B}^+$ be the union of $\mathcal{B}$ and some two-point set $\{*_1, *_2\}$ and consider the mapping $\tau^+ : \mathcal{B} \to [0,1]^{\mathcal{B}^+} = [0,1]^{\mathcal{B}} \times [0,1] \times [0,1]$ defined by

$$\tau^+(G) = \bigl(\tau(G),\, v_1(G)^{-1},\, v_2(G)^{-1}\bigr). \tag{8.3}$$

Then τ^+ is injective and we can identify $\mathcal{B}$ with its image $\tau^+(\mathcal{B}) \subseteq [0,1]^{\mathcal{B}^+}$ and define $\overline{\mathcal{B}} \subseteq [0,1]^{\mathcal{B}^+}$ as its closure; this is a compact metric space.

The functions $t(F,\cdot)$, $t_{\mathrm{inj}}(F,\cdot)$, $t_{\mathrm{ind}}(F,\cdot)$ and $v_j(\cdot)^{-1}$, for $F \in \mathcal{B}$ and $j = 1,2$, have unique continuous extensions to $\overline{\mathcal{B}}$.

We let $\mathcal{B}_{\infty\infty} := \{G \in \overline{\mathcal{B}} : v_1(G) = v_2(G) = \infty\}$; this is the set of all limit objects of sequences (G_n) in $\mathcal{B}$ with $v_1(G_n), v_2(G_n) \to \infty$. By (8.1), $t_{\mathrm{inj}}(F,G) = t(F,G)$ for every $G \in \mathcal{B}_{\infty\infty}$ and every $F \in \mathcal{B}$. The projection $\pi : \overline{\mathcal{B}} \to \overline{\mathcal{B}^*}$ restricts to a homeomorphism $\mathcal{B}_{\infty\infty} \cong \overline{\mathcal{B}^*}$.

Remark 8.1. Note that in the bipartite case there are other limit objects too in $\overline{\mathcal{B}}$; in fact, $\overline{\mathcal{B}}$ can be partitioned into $\mathcal{B}$, $\mathcal{B}_{\infty\infty}$, and the sets $\mathcal{B}_{n\infty}$, $\mathcal{B}_{\infty n}$, for $n = 1,2,\dots$, where, for example, $\mathcal{B}_{n_1\infty}$ is the set of limits of sequences (G_n) of bipartite graphs such that $v_2(G_n) \to \infty$ but $v_1(G_n) = n_1$ is constant. We will not consider such degenerate limits further here, but we remark that in the simplest case $n_1 = 1$, a bipartite graph in $\mathcal{B}^L_{1n_2}$ can be identified with a subset of $[n_2]$, and an unlabelled graph in $\mathcal{B}_{1n_2}$ thus with a number in $m \in \{0,\dots,n_2\}$, the number of edges in the graph, and it is easily seen that a sequence of such unlabelled graphs with $n_2 \to \infty$ converges in $\overline{\mathcal{B}}$ if and only if the proportion m/n_2 converges; hence we can identify $\mathcal{B}_{1\infty}$ with the interval [0,1].

We have the following basic result, cf. Theorem 2.1.

Theorem 8.1. *Let (G_n) be a sequence of bipartite graphs with $v_1(G_n)$, $v_2(G_n) \to \infty$. Then the following are equivalent.*

(i) *$t(F, G_n)$ converges for every $F \in \mathcal{B}$.*
(ii) *$t_{\mathrm{inj}}(F, G_n)$ converges for every $F \in \mathcal{B}$.*
(iii) *$t_{\mathrm{ind}}(F, G_n)$ converges for every $F \in \mathcal{B}$.*
(iv) *G_n converges in $\overline{\mathcal{B}}$.*

In this case, the limit G of G_n belongs to $\mathcal{B}_{\infty\infty}$ and the limits in (i), (ii) and (iii) are $t(F, G)$, $t_{\mathrm{inj}}(F, G)$ and $t_{\mathrm{ind}}(F, G)$.

For convergence of random unlabelled bipartite graphs, the results in Section 3 hold with trivial changes.

Theorem 8.2. *Let G_n, $n \geq 1$, be random unlabelled bipartite graphs and assume that $v_1(G_n), v_2(G_n) \xrightarrow{\mathrm{p}} \infty$. The following are equivalent, as $n \to \infty$.*

(i) *$G_n \xrightarrow{\mathrm{d}} \Gamma$ for some random $\Gamma \in \overline{\mathcal{B}}$.*
(ii) *For every finite family $F_1, \dots, F_m$ of (non-random) bipartite graphs, the random variables $t(F_1, G_n), \dots, t(F_m, G_n)$ converge jointly in distribution.*
(iii) *For every (non-random) $F \in \mathcal{B}$, the random variables $t(F, G_n)$ converge in distribution.*
(iv) *For every (non-random) $F \in \mathcal{B}$, the expectations $\mathbb{E}\, t(F, G_n)$ converge.*

If these properties hold, then the limits in (ii), (iii) and (iv) are $\big(t(F_i, \Gamma)\big)_{i=1}^{m}$, $t(F, \Gamma)$ and $\mathbb{E}\, t(F, \Gamma)$, respectively. Furthermore, $\Gamma \in \mathcal{B}_{\infty\infty}$ a.s.

The same results hold if t is replaced by t_{inj} or t_{ind}.

Corollary 8.3. *Let G_n, $n \geq 1$, be random unlabelled bipartite graphs such that $v_1(G_n), v_2(G_n) \xrightarrow{\mathrm{p}} \infty$, and let $G \in \mathcal{B}_{\infty\infty}$. The following are equivalent, as $n \to \infty$.*

(i) *$G_n \xrightarrow{\mathrm{p}} G$.*
(ii) *$t(F, G_n) \xrightarrow{\mathrm{p}} t(F, G)$ for every (non-random) $F \in \mathcal{B}$.*
(iii) *$\mathbb{E}\, t(F, G_n) \to t(F, G)$ for every (non-random) $F \in \mathcal{B}$.*

The same result holds if t is replaced by t_{inj} or t_{ind}.

As above, the distribution of Γ is uniquely determined by the numbers $\mathbb{E}\, t(F, \Gamma)$, $F \in \mathcal{B}$.

Let $\mathcal{B}^L_{\infty\infty}$ denote the set of all labelled infinite bipartite graphs with the vertex sets $V_1(G) = V_2(G) = \mathbb{N}$. $\mathcal{B}^L_{\infty\infty}$ is a compact metric space with the natural product topology.

If G is a bipartite graph, let $\widehat{G}$ be the random labelled bipartite graph obtained by random labellings of the vertices in $V_j(G)$ by the numbers $1, \dots, v_j(G)$, for $j = 1, 2$. This is a random finite bipartite graph, but we can also regard it as a random element of $\mathcal{B}^L_{\infty\infty}$ by adding isolated vertices.

DEFINITION. A random infinite bipartite graph $H \in \mathcal{B}^L_{\infty\infty}$ is *exchangeable* if its distribution is invariant under every pair of finite permutations of $V_1(H)$ and $V_2(H)$.

THEOREM 8.4. *Let (G_n) be a sequence of random graphs in $\mathcal{B}$ and assume that $v_1(G_n), v_2(G_n) \xrightarrow{\text{p}} \infty$. Then the following are equivalent.*

(i) $G_n \xrightarrow{\text{d}} \Gamma$ *in* $\overline{\mathcal{B}}$ *for some random* $\Gamma \in \overline{\mathcal{B}}$.
(ii) $\widehat{G_n} \xrightarrow{\text{d}} H$ *in* $\mathcal{B}^L_{\infty\infty}$ *for some random* $H \in \mathcal{B}^L_{\infty\infty}$.

If these hold, then $\mathbb{P}(H|_{[k_1]\times[k_2]} = F) = \mathbb{E}\, t_{\text{ind}}(F, \Gamma)$ for every $F \in \mathcal{B}^L_{k_1 k_2}$. Furthermore, $\Gamma \in \mathcal{B}_{\infty\infty}$ a.s., and H is exchangeable.

THEOREM 8.5. *There is a one-to-one correspondence between distributions of random elements $\Gamma \in \mathcal{B}_{\infty\infty}$ (or $\overline{\mathcal{B}^*}$) and distributions of exchangeable random infinite graphs $H \in \mathcal{B}^L_{\infty\infty}$ given by*

$$\mathbb{E}\, t_{\text{ind}}(F, \Gamma) = \mathbb{P}(H|_{[k_1]\times[k_2]} = F) \tag{8.4}$$

for every $k_1, k_2 \geq 1$ and every $F \in \mathcal{B}^L_{k_1 k_2}$, or, equivalently,

$$\mathbb{E}\, t(F, \Gamma) = \mathbb{P}(H \supset F) \tag{8.5}$$

for every $F \in \mathcal{B}^L$. Furthermore, $H|_{[n_1]\times[n_2]} \xrightarrow{\text{d}} \Gamma$ in $\overline{\mathcal{B}}$ as $n_1, n_2 \to \infty$.

COROLLARY 8.6. *There is a one-to-one correspondence between elements Γ of $\mathcal{B}_{\infty\infty} \cong \overline{\mathcal{B}^*}$ and extreme points of the set of distributions of exchangeable random infinite graphs $H \in \mathcal{B}^L_{\infty\infty}$. This correspondence is given by*

$$t(F, \Gamma) = \mathbb{P}(H \supset F) \tag{8.6}$$

for every $F \in \mathcal{B}^L$. Furthermore, $H|_{[n_1]\times[n_2]} \xrightarrow{\text{p}} \Gamma$ in $\overline{\mathcal{B}}$ as $n_1, n_2 \to \infty$.

REMARK 8.2. We have not checked whether $H|_{[n_1]\times[n_2]} \xrightarrow{\text{a.s.}} \Gamma$ in $\overline{\mathcal{B}}$ as $n_1, n_2 \to \infty$. This holds at least for a subsequence $(n_1(m), n_2(m))$ with both $n_1(m)$ and $n_2(m)$ non-decreasing because then $t_{\text{inj}}(F, H|_{[n_1]\times[n_2]})$ is a reverse martingale.

Theorem 8.7. *Let H be an exchangeable random infinite bipartite graph. Then the following are equivalent.*

(i) *The distribution of H is an extreme point in the set of exchangeable distributions in $\mathcal{B}^L_{\infty\infty}$.*
(ii) *If F_1 and F_2 are two (finite) bipartite graphs with the vertex sets $V_j(F_1)$ and $V_j(F_2)$ disjoint subsets of $\mathbb{N}$ for $j = 1, 2$, then*

$$\mathbb{P}(H \supset F_1 \cup F_2) = \mathbb{P}(H \supset F_1)\,\mathbb{P}(H \supset F_2).$$

The construction in Section 6 takes the following form; note that there is no need to assume symmetry of W. For every $W \in \mathcal{W}$, we define an infinite random bipartite graph $G(\infty, \infty, W) \in \mathcal{B}^L_{\infty\infty}$ as follows: we first choose two sequence $X_1, X_2, \dots$ and $Y_1, Y_2, \dots$ of i.i.d. random variables uniformly distributed on $[0, 1]$, and then, given these sequences, for each pair $(i, j) \in \mathbb{N} \times \mathbb{N}$ we draw an edge ij with probability $W(X_i, Y_j)$, independently for all pairs (i, j). Further, let $G(n_1, n_2, W)$ be the restriction $G(\infty, \infty, W)|_{[n_1]\times[n_2]}$, which is obtained by the same construction with finite sequences $X_1, \dots, X_{n_1}$ and $Y_1, \dots, Y_{n_2}$.

It is evident that $G(\infty, \infty, W)$ is an exchangeable infinite random bipartite graph. Furthermore, it satisfies Theorem 8.7(ii). Theorem 8.5 and Corollary 8.6 yield a corresponding element $\Gamma''_W \in \mathcal{B}_{\infty\infty} \cong \overline{\mathcal{B}^*}$ such that $G(n_1, n_2, W) \xrightarrow{\mathrm{p}} \Gamma''_W$ as $n_1, n_2 \to \infty$ and, for every $F \in \mathcal{B}^L_{k_1k_2}$,

$$t(F, \Gamma''_W) = \int_{[0,1]^{k_1+k_2}} \prod_{ij\in\mathbb{E}(F)} W(x_i, y_j)\,\mathrm{d}x_1 \dots \mathrm{d}x_{\mathrm{k}_1}\,\mathrm{d}y_1 \dots \mathrm{d}y_{\mathrm{k}_2}.$$

The result by Aldous [1] in the non-symmetric case is that every exchangeable infinite random bipartite graph is obtained as a mixture of such $G(\infty, \infty, W)$; in other words as $G(\infty, \infty, W)$ with a random W.

By Theorem 8.5 and Corollary 8.6 above, this implies (and is implied by) the fact that every element of $\overline{\mathcal{B}}$ equals Γ''_W for some (non-unique) $W \in \mathcal{W}$; the bipartite version of the characterization by Lovász and Szegedy [21].

9 – Directed graphs

A *directed graph* G consists of a vertex set $V(G)$ and an edge set $E(G) \subseteq V(G) \times V(G)$; the edge indicators thus form an arbitrary zero–one matrix $\{X_{ij}\}$, $i, j \in V(G)$. Note that we allow loops, corresponding to the diagonal indicators X_{ii}. The definitions and results above have analogues for directed graphs too, with mainly notational differences. We sketch these in this section, leaving the details to the reader.

Let $\mathcal{D}_n^L$ be the set of the 2^{n^2} labelled directed graphs with vertex set $[n]$ and let $\mathcal{D}_n$ be the quotient set $\mathcal{D}_n^L/\cong$ of unlabelled directed graphs with n vertices; further, let $\mathcal{D}^L := \bigcup_{n\geq 1} \mathcal{D}_n^L$ and $\mathcal{D} := \bigcup_{n\geq 1} \mathcal{D}_n$.

The definitions in Section 2 apply to directed graphs too, with at most notational differences. $G[k]$ and $G[k]'$ now are random directed graphs and $t(F, G)$, $t_{\mathrm{inj}}(F, G)$ and $t_{\mathrm{ind}}(F, G)$ are defined for (unlabelled) directed graphs F and G by (2.1)–(2.3). We now define $\tau : \mathcal{D} \to [0, 1]^{\mathcal{D}}$ by, cf. (2.7),

$$\tau(G) := (t(F, G))_{F\in\mathcal{D}} \in [0, 1]^{\mathcal{D}}. \tag{9.1}$$

We define $\mathcal{D}^* := \tau(\mathcal{D}) \subseteq [0, 1]^{\mathcal{D}}$ to be the image of $\mathcal{D}$ under this mapping τ, and let $\overline{\mathcal{D}^*}$ be the closure of $\mathcal{D}^*$ in $[0, 1]^{\mathcal{D}}$; this is a compact metric space.

Again, τ is not injective. We let $\mathcal{D}^+$ be the union of $\mathcal{D}$ and some one-point set $\{*\}$ and consider the mapping $\tau^+ : \mathcal{D} \to [0, 1]^{\mathcal{D}^+} = [0, 1]^{\mathcal{D}} \times [0, 1]$ defined by (2.9) as before. Then τ^+ is injective and we can identify $\mathcal{D}$ with its image $\tau^+(\mathcal{D}) \subseteq [0, 1]^{\mathcal{D}^+}$ and define $\overline{\mathcal{D}} \subseteq [0, 1]^{\mathcal{D}^+}$ as its closure; this is a compact metric space. The functions $t(F, \cdot)$, $t_{\mathrm{inj}}(F, \cdot)$, $t_{\mathrm{ind}}(F, \cdot)$ and $v(\cdot)^{-1}$, for $F \in \mathcal{D}$, have unique continuous extensions to $\overline{\mathcal{D}}$.

We let $\mathcal{D}_\infty := \{G \in \overline{\mathcal{D}} : v(G) = \infty\}$; this is the set of all limit objects of sequences (G_n) in $\mathcal{D}$ with $v(G_n) \to \infty$. Analogously to (2.10), $t_{\mathrm{inj}}(F, G) = t(F, G)$ for every $G \in \mathcal{D}_\infty$ and every $F \in \mathcal{D}$. The projection $\pi : \overline{\mathcal{D}} \to \overline{\mathcal{D}^*}$ restricts to a homeomorphism $\mathcal{D}_\infty \cong \overline{\mathcal{D}^*}$.

All results in Sections 2–5 are valid for directed graphs too, with at most notational differences.

The main difference for the directed case concerns the representations discussed in Section 6. Since two vertices may be connected by up to two directed edges (in opposite directions), and the events that the two possible edges occur typically are dependent, a single function W is no longer enough. Instead, we have a representation using several functions as follows.

Let $\mathcal{W}_5$ be the set of quintuples $\mathbf{W} = (W_{00}, W_{01}, W_{10}, W_{11}, w)$ where $W_{\alpha\beta} : [0, 1]^2 \to [0, 1]$ and $w : [0, 1] \to \{0, 1\}$ are measurable functions such that $\sum_{\alpha,\beta=0}^{1} W_{\alpha\beta}(x, y) = 1$ and $W_{\alpha\beta}(x, y) = W_{\beta\alpha}(y, x)$ for $\alpha, \beta \in \{0, 1\}$ and $x, y \in [0, 1]$. For $\mathbf{W} \in \mathcal{W}_5$, we define a random infinite directed graph $G(\infty, \mathbf{W})$ by specifying its edge indicators X_{ij} as follows: we first choose a sequence $Y_1, Y_2, \dots$ of i.i.d. random variables uniformly distributed on $[0, 1]$, and then, given this sequence, let $X_{ii} = w(Y_i)$ and for each pair (i, j) with $i < j$ choose X_{ij} and X_{ji} at random such that

$$\mathbb{P}(X_{ij} = \alpha \text{ and } X_{ji} = \beta) = W_{\alpha\beta}(Y_i, Y_j), \qquad \alpha, \beta \in \{0, 1\}; \tag{9.2}$$

this is done independently for all pairs (i, j) with $i < j$ (conditionally given $\{Y_k\}$). In other words, for every i we draw a loop at i if $w(Y_i) = 1$ and for each pair (i, j) with $i < j$ we draw edges ij and ji at random such that (9.2) holds.

Further, let $G(n, \mathbf{W})$ be the restriction $G(\infty, \mathbf{W})|_{[n]}$, which is obtained by the same construction with a finite sequence $Y_1, \dots, Y_n$.

In particular, note that the loops appear independently, each with probability $p = \mathbb{P}\big(w(Y_1) = 1\big)$. We may specify the loops more clearly by the following alternative version of the construction. Let $\mathcal{S} := [0,1] \times \{0,1\}$ and let $\mathcal{W}_4$ be the set of quadruples $\mathbf{W} = (W_{00}, W_{01}, W_{10}, W_{11})$ where $W_{\alpha\beta} : \mathcal{S}^2 \to [0,1]$ are measurable functions such that $\sum_{\alpha,\beta=0}^{1} W_{\alpha\beta}(x,y) = 1$ and $W_{\alpha\beta}(x,y) = W_{\beta\alpha}(y,x)$ for $\alpha, \beta \in \{0,1\}$ and $x, y \in \mathcal{S}$. For every $\mathbf{W} \in \mathcal{W}_4$ and $p \in [0,1]$, we define a random infinite directed graph $G(\infty, \mathbf{W}, p)$ by specifying its edge indicators X_{ij} as follows: We first choose sequences $\xi_1, \xi_2, \dots$ and $\zeta_1, \zeta_2, \dots$ of random variables, all independent, with $\xi_i \sim U(0,1)$ and $\zeta_i \sim \mathrm{Be}(p)$, i.e., $\zeta_i \in \{0,1\}$ with $\mathbb{P}(\zeta_i = 1) = p$; we let $Y_i := (\xi_i, \zeta_i) \in \mathcal{S}$. Then, given these sequences, let $X_{ii} = \zeta_i$ and for each pair (i,j) with $i < j$ choose X_{ij} and X_{ji} at random according to (9.2), independently for all pairs (i,j) with $i < j$ (conditionally given $\{Y_k\}$). In other words, ζ_i is the indicator of a loop at i. Further, let $G(n, \mathbf{W}, p)$ be the restriction $G(\infty, \mathbf{W}, p)|_{[n]}$, which is obtained by the same construction with a finite sequence $Y_1, \dots, Y_n$.

It is obvious from the symmetry of the construction that the random infinite directed graphs $G(\infty, \mathbf{W})$ and $G(\infty, \mathbf{W}, p)$ are exchangeable. Further, using Theorem 5.5, their distributions are extreme points, so by Corollary 5.4 they correspond to directed graph limits, i.e., elements of $\mathcal{D}_\infty$, which we denote by $\Gamma_{\mathbf{W}}$ and $\Gamma_{\mathbf{W},p}$, respectively; (5.3) shows that if $F \in \mathcal{D}_k$, then

$$t(F, \Gamma_{\mathbf{W}}) = \mathbb{P}\big(F \subseteq G(k, \mathbf{W})\big), \quad t(F, \Gamma_{\mathbf{W},p}) = \mathbb{P}\big(F \subseteq G(k, \mathbf{W}, p)\big).$$

By Theorem 5.3 and Remark 5.1, $G(n, \mathbf{W}) \to \Gamma_{\mathbf{W}}$ and $G(n, \mathbf{W}, p) \to \Gamma_{\mathbf{W},p}$ a.s. as $n \to \infty$.

We can show a version of the representation results in Section 6 for directed graphs.

THEOREM 9.1. *An exchangeable random infinite directed graph is obtained as a mixture of $G(\infty, \mathbf{W})$; in other words, as $G(\infty, \mathbf{W})$ with a random $\mathbf{W}$. Alternatively, it is obtained as a mixture of $G(\infty, \mathbf{W}, p)$; in other words, as $G(\infty, \mathbf{W}, p)$ with a random $(\mathbf{W}, p)$.*

Every directed graph limit, i.e., every element of $\mathcal{D}_\infty$, is $\Gamma_{\mathbf{W}}$ for some $\mathbf{W} \in \mathcal{W}_5$, or equivalently $\Gamma_{\mathbf{W},p}$ for some $\mathbf{W} \in \mathcal{W}_4$ and $p \in [0,1]$.

PROOF. For jointly exchangeable random arrays $\{X_{ij}\}$ of zero–one variables, the Aldous–Hoover representation theorem takes the form [17, Theorem 7.22]

$$\begin{aligned} X_{ii} &= f_1(\xi_\emptyset, \xi_i), \\ X_{ij} &= f_2(\xi_\emptyset, \xi_i, \xi_j, \xi_{ij}), \qquad i \neq j, \end{aligned}$$

where $f_1 : [0,1]^2 \to \{0,1\}$ and $f_2 : [0,1]^4 \to \{0,1\}$ are two measurable functions, $\xi_{ji} = \xi_{ij}$, and $\xi_\emptyset$, ξ_i $(1 \le i)$ and ξ_{ij} $(1 \le i < j)$ are independent random variables uniformly distributed on $[0,1]$ (as in the proof of Theorem 7.1). If further the distribution of the array $\{X_{ij}\}$ is an extreme point in the set of exchangeable distributions, then by Theorem 5.5 and [17, Lemma 7.35], there exists such a representation where f_1 and f_2 do not depend on $\xi_\emptyset$, so $X_{ii} = f_1(\xi_i)$ and $X_{ij} = f_2(\xi_i, \xi_j, \xi_{ij})$, $i \ne j$. In this case, define $w = f_1$ and

$$W_{\alpha\beta}(x,y) := \mathbb{P}\big(f_2(x,y,\xi) = \alpha \text{ and } f_2(y,x,\xi) = \beta\big), \qquad \alpha,\beta \in \{0,1\},$$

where $\xi \sim U(0,1)$. This defines a quintuple $\mathbf{W} \in \mathcal{W}_5$, such that the edge indicators X_{ij} of $G(\infty, \mathbf{W})$ have the desired distribution.

In general, the variable $\xi_\emptyset$ can be interpreted as making $\mathbf{W}$ random.

To obtain the alternative representation, let $\zeta_i := w(\xi_i) = X_{ii}$ and $p := \mathbb{P}(\zeta_i = 1)$. There exists a measure preserving map $\phi : (\mathcal{S}, \mu_p) \to ([0,1], \lambda)$, where λ is the Lebesgue measure and $\mu_p := \lambda \times \mathrm{Be}(p)$, such that $[0,1] \times \{j\}$ is mapped onto $\{x \in [0,1] : w(x) = j\}$ for $j = 0,1$ (i.e., $w \circ \phi(x,\zeta) = \zeta$), and we can use the quadruple $(W_{\alpha\beta} \circ \phi)_{\alpha,\beta}$.

The representations for graph limits follow by Corollary 5.4 as discussed above. ∎

Example 9.2. A *random tournament* T_n is a random directed graph on n vertices without loops where each pair of vertices is connected by exctly one edge, with random direction (with equal probabilities for the two directions, and independent of all other edges). This equals $G(n, \mathbf{W})$ or $G(n, \mathbf{W}, p)$ with $W_{00} = W_{11} = 0$, $W_{01} = W_{10} = 1/2$, and $w = 0$ or $p = 0$, and converges thus a.s. to the limit $\Gamma_{\mathbf{W},0}$ for $\mathbf{W} = (W_{\alpha\beta})_{\alpha,\beta}$.

Note that if $\{X_{ij}\}$ are the edge indicators of an exchangeable random infinite directed graph, then the loop indicators $\{X_{ii}\}$ form a binary exchangeable sequence, and the representation as $G(\infty, \mathbf{W}, p)$ in Theorem 9.1 exhibits them as a mixture of i.i.d. $\mathrm{Be}(p)$ variable, which has brought us back to de Finetti's theorem 1.1.

Acknowledgements

This lecture is an extended version of a talk presented by PD at the 100th anniversary of de Finetti's birth in Rome, 2006. We thank the organizers. This work was partially funded by the French ANR's Chaire d'excellence grant to PD.

SJ thanks Christian Borgs and Jennifer Chayes for inspiration from lectures and discussions during the Oberwolfach meeting 'Combinatorics, Probability and Computing', held in November, 2006. Parts of the research were completed during a visit by SJ to the Université de Nice - Sophia Antipolis in January 2007.

REFERENCES

[1] D. Aldous: *Representations for partially exchangeable arrays of random variables*, J. Multivar. Anal. **11** (1981), 581-598 .

[2] P. Billingsley: *Convergence of Probability Measures*, Wiley, New York, 1968.

[3] B. Bollobás – S. Janson – O. Riordan: *The phase transition in inhomogeneous random graphs*, Random Struct. Alg. **31** (2007), 3–122.

[4] B. Bollobás – O. Riordan: *Sparse graphs: metrics and random models.*, Preprint, 2007. http://arxiv.org/0708.1919v1

[5] C. Borgs – J. T. Chayes – L. Lovász: *Unique limits of dense graph sequences*, In preparation.

[6] C. Borgs – J. T. Chayes – L. Lovász– V. T. Sós – B. Szegedy – K. Vesztergombi: *Graph limits and parameter testing*, STOC'06: Proceedings of the 38th Annual ACM Symposium on Theory of Computing, 261–270. ACM, New York, 2006.

[7] C. Borgs – J. T. Chayes – L. Lovász– V. T. Sós – K. Vesztergombi: *Counting graph homomorphisms*, Topics in Discrete Mathematics (ed. M. Klazar, J. Kratochvil, M. Loebl, J. Matousek, R. Thomas, P. Valtr), 315–371. Springer, New York, 2006.

[8] C. Borgs – J. T. Chayes – L. Lovász– V. T. Sós – K. Vesztergombi: *Convergent sequences of dense graphs I: Subgraph frequencies, metric properties and testing*, Preprint, 2007. http://arxiv.org/math.CO/0702004

[9] C. Borgs – J. T. Chayes – L. Lovász– V. T. Sós – K. Vesztergombi: *Convergent sequences of dense graphs II: Multiway cuts and statistical physics.*, Preprint, 2007. http://research.microsoft.com/ borgs/

[10] B. De Finetti: *Funzioni caratteristica di un fenomeno aleatorio*, Atti della R. Academia Nazionale dei Lincei Ser 6, Memorie Classe di Scienze Fisiche, Mathematiche e Naturali,**4** (1931), 251-299.

[11] B. De Finetti: *La prévision: ses lois logiques, ses sources subjectives*, Ann. Inst. Henri Poincaré, **7** (1937), 1-66.

[12] P. Diaconis – D. Freedman: *On the statistics of vision: The Julesz conjecture*, J. Math. Psychol. **24** (1981), 112-138.

[13] P. Diaconis – D. Freedman: *Partial exchangeability and sufficiency.*, Statistics: Applications and New Directions (eds. J. K. Jhosh and J. Roy), 205–236. Indian Statistical Institute, Calcutta, 1984.

[14] P. Diaconis – S. Holmes – S. Janson: *Threshold graph limits and random threshold graphs*, In preparation.

[15] M. Freedman – L. Lovász – A. Schrijver: *Reflection positivity, rank connectivity, and homomorphism of graphs*, J. Amer. Math. Soc. **20** (2007), 37-51.

[16] D. Hoover: *Relations on Probability Spaces and Arrays of Random Variables*, Preprint, Institute for Advanced Study, Princeton, NJ, 1979.

[17] O. Kallenberg: *Probabilistic Symmetries and Invariance Principles*, Springer, New York, 2005.

[18] L. Lovász: *The rank of connection matrices and the dimension of graph algebras.*, Eur. J. Comb. **27** (2006), 962–970.

[19] L. Lovász: *Connection matrices*, In: Combinatorics, Complexity, and Chance: A Tribute to Dominic Welsh (ed. G. Grimmet and C. McDiarmid), 179–190. Oxford University Press, Oxford, 2007.

[20] L. Lovász – V. T. Sós: *Generalized quasirandom graphs*, J. Comb. Theory B. **98** (2008), Issue 1, 146-163.

[21] L. Lovász – B. Szegedy: *Limits of dense graph sequences*, J. Comb. Theory B, **96** (2006), 933-957.

[22] L. Lovász – B. Szegedy: *Szemerédi's Lemma for the analyst*, J. Geom. Func. Anal. **17** (2007), 252-270.

[23] L. Lovász – B. Szegedy: *Contractors and connectors in graph algebras*, J. Comb. Theory B, to appear.

[24] L. Lovász – B. Szegedy: *Testing properties of graphs and functions*, Preprint, 2007.

Lavoro pervenuto alla redazione il 14 gennaio 2008
ed accettato per la pubblicazione il 18 gennaio 2008.
Bozze licenziate il 9 aprile 2008

Indirizzo degli Autori:

P. Diaconis – Department of Mathematics – Stanford University – Stanford California 94305, USA – Département de Mathématiques – Université de Nice - Sophia Antipolis – Parc Valrose – 06108 Nice Cedex 02, France

S. Janson – Department of Mathematics – Uppsala University – PO Box 480, SE-751 06 Uppsala, Sweden
E-mail: svante.janson@math.uu.se

Rendiconti di Matematica, Serie VII
Volume 28, Roma (2008), 63–81

Exchangeability and semigroups

PAUL RESSEL

Abstract: *Exchangeability of a "random object" is a strong symmetry condition, leading in general to a convex set of distributions not too far from a "simplex" - a set easily described by its extreme points, in this case distributions with very special properties as for example iid coin tossing sequences in de Finetti's original result. Although in most cases of interest the symmetry is defined via a non–commutative group acting on the underlying space, it very often can be described by a suitable factorization involving an abelian semigroup. The factorizing function typically turns out to be positive definite, and results from Harmonic Analysis on semigroups become applicable. In this way many known theorems on exchangeability can be given an alternative proof, more analytic/algebraic in a sense, but also new results become available.*

1 – Introduction

A sequence $X = (X_1, X_2, \dots)$ of random variables is called *exchangeable* if for any permutation π of $\mathbb{N}$ the sequence $(X_{\pi(1)}, X_{\pi(2)}, \dots)$ has the same distribution as X; of course it is enough to require this property for *finite* permutations π (in the sense that $\{i \in \mathbb{N} \mid \pi(i) \neq i\}$ is a finite set). This holds obviously for an iid–sequence and so also for a mixture (in distribution) of iid's, since exchangeable distributions form a convex set. As is well known, in 1930 Bruno de Finetti published the pathbreaking result that for $\{0, 1\}$–valued random variables the converse holds, too: exchangeable sequences are precisely the mixtures of iid Bernoulli sequences. A few years later de Finetti generalized this to real–valued random variables, and in 1955 Hewitt and Savage proved the

Key Words and Phrases: *Exchangeability – Semigroups – Positive definite functions – de Finetti's theorem – Random partition.*
A.M.S. Classification: 60E05, 43A35

corresponding result for arbitrary compact Hausdorff spaces, from which it is immediately seen to be true also for Borel subsets of compact spaces, for example for locally compact spaces. In ([6], Theorem 4) a further generalization to completely regular Hausdorff spaces was shown.

In contrast to this "global" point of view (i.e. considering *all* exchangeable distributions) a different kind of question seems natural: is it possible to characterize mixtures of iid–sequences of a particular type, say normal or Poisson distributed, or with a special form of their Fourier or Laplace transform, or even (for non–negative random variables) of their multivariate survival function?

An attempt for fairly general answers was given in [6] and subsequent papers. Here we present a certain overview, and as "Main Theorem" a new result which in a way is a de Finetti theorem for positive definite functions on abelian semigroups, which appear (perhaps surprising) as a natural tool in this connection.

After explaining some basic notions concerning positive definite and related functions on semigroups in Section 3, the main result will be presented in Section 4. The extended de Finetti–type theorem in Section 5 is given a new proof, based on the main theorem, and followed by a few typical examples. The above mentioned theorem of Hewitt and Savage is shown in Section 6 to be another "almost straightforward" consequence of the main theorem. Finally, the closing Section 7 presents a different point of view to the main theorem, followed by an application to exchangeable random partitions.

2 – Why semigroups?

They enter the scene naturally, as can be seen already in de Finetti's original result.

Let P be an exchangeable probability measure on the space of all (infinite) $0-1$ sequences, abbreviated $P \in M_+^{1,e}(\{0,1\}^\infty)$, the "$e$" referring to exchangeability. Then $P(x_1, \ldots, x_n)$ depends only on $x_1 + \ldots + x_n$, i.e.

$$P(x_1, \ldots, x_n) = \varphi_n\left(\sum_{i=1}^n x_i\right) = \varphi\left(\sum_{i=1}^n x_i, n\right) = \\ = \varphi\left(\sum_{i=1}^n (x_i, 1)\right)$$

with φ defined on the set

$$S := \left\{(k,n) \in \mathbb{N}_0^2 \mid k \leq n\right\}$$

which is a (sub-) semigroup inside $\mathbb{N}_0^2$. The crucial point will be that φ turns out to be a socalled *positive definite* function, therefore a (unique) mixture of socalled *characters*, taking here the form

$$\sigma : (k,n) \longmapsto \sigma(k,n) = p^k q^{n-k}, \quad p,q \in \mathbb{R}$$

and it is easy to see that only characters with $p, q \geq 0$ and $p+q=1$ play a rôle. Inserting this we get (slightly abusing the letter μ as a measure on the characters resp. on $[0,1]$)

$$\begin{aligned} P(x_1,\dots,x_n) &= \varphi\left(\sum_{i=1}^n (x_i,1)\right) = \\ &= \int \sigma\left(\sum_{i=1}^n (x_i,1)\right) d\mu(\sigma) = \\ &= \int \prod_{i=1}^n \sigma(x_i,1) d\mu(\sigma) = \\ &= \int_0^1 \prod_{i=1}^n p^{x_i}(1-p)^{1-x_i} d\mu(p) \end{aligned}$$

for some (unique) $\mu \in M_+^1([0,1])$, which is de Finetti's result, proved in 1930, cf. [3].

3 – Basic definitions and notations

We will make use in a crucial way of some notions and results about positive definite and related functions on semigroups, an introduction to which can be found in [1], Chapter 4.

Let S denote an abelian semigroup, written additively, with neutral element 0, and possibly with an involution, i.e. a mapping $s \longmapsto s^-$, with $(s+t)^- = s^- + t^-, 0^- = 0$ and $(s^-)^- = s$ which in many cases is just the identity.

$\sigma : S \longrightarrow \mathbb{C}$ is a *character* iff

$$\sigma(s+t) = \sigma(s)\cdot\sigma(t), \quad \sigma(s^-) = \overline{\sigma(s)}, \quad \sigma(0) = 1$$

$\varphi : S \longrightarrow \mathbb{C}$ is *positive definite* (abbrev. "p.d.") iff

$$\sum_{j,k=1}^n c_j \overline{c_k} \varphi(s_j + s_k^-) \geq 0 \quad \forall\, n \in \mathbb{N}, c_j \in \mathbb{C}, s_j \in S$$

$\varphi : S \longrightarrow \mathbb{C}$ is *completely positive definite* ("c.p.d.") iff $s \longmapsto \varphi(s+a)$ is positive definite $\forall a \in S$;

$\alpha : S \longrightarrow \mathbb{R}_+$ is an *absolute value* iff

$$\alpha(s+t) \leq \alpha(s)\cdot\alpha(t), \quad \alpha(s^-) = \alpha(s), \quad \alpha(0) = 1$$

$f : S \longrightarrow \mathbb{C}$ is α–bounded (with α a fixed absolute value) iff

$$|f(s)| \leq C\cdot\alpha(s) \quad \forall\, s \in S, \quad \text{for some} \quad C \geq 0, \quad \text{briefly:} \quad |f| \leq C\alpha$$

(if furthermore $f(0) = 1$ and f is p.d., then one can take $C = 1$);
f is *exponentially bounded* iff it is α–bounded with respect to some absolute value α;
$S^* :=$ set of all characters of S;
$\mathcal{P}(S) :=$ set of all positive definite functions on S;
$S^\alpha := \{\sigma \in S^* \mid \sigma \text{ is } \alpha\text{–bounded}\}$ then $S^\alpha = \{\sigma \in S^* \mid |\sigma| \leq \alpha\}$;
$\mathcal{P}^\alpha(S) := \{\varphi \in \mathcal{P}(S) \mid \varphi \text{ is } \alpha\text{–bounded}\}$;
$\hat{S} :=$ all bounded characters on S, then $\hat{S} = \{\sigma \in S^* \mid |\sigma| \leq 1\}$;
$\mathcal{P}^b(S) :=$ all bounded positive definite functions on S.

For any set B of complex functions on S, the symbols B_+ and B_1 denote respectively $B \cap \{f \mid f(x) \geq 0 \quad \forall s \in S\}$ and $B \cap \{f \mid f(0) = 1\}$.

It is easily seen that

$S^* \subseteq \mathcal{P}_1(S) := \{\varphi \in \mathcal{P}(S) \mid \varphi(0) = 1\}$;
$S^\alpha \subseteq \mathcal{P}_1^\alpha(S), \quad \varphi \in \mathcal{P}_1^\alpha(S) \Longrightarrow |\varphi| \leq \alpha$;
$\hat{S} \subseteq \mathcal{P}_1^b(S), \quad \varphi \in \mathcal{P}_1^b(S) \Longrightarrow |\varphi| \leq 1$

and each $\sigma \in S_+^*$ is even c.p.d.

4 – The main result

If K is a non–empty compact convex subset of some locally convex vector space, then K is by Krein-Milman's theorem the closed convex hull of $ex(K)$, the extreme points of K. If $ex(K)$ is closed, and if furthermore the representation of points in K as barycenters of (Radon) measures on $ex(K)$ is unique, K is called a *Bauer simplex*. In all subsequent applications K will be a subset of $\mathbb{C}^S$, the set of all complex–valued functions on S, equipped with the topology of pointwise convergence. Note that in this case for any given $\gamma : S \longrightarrow \mathbb{R}_+$ the set $K := \{f \in \mathbb{C}^S \mid |f| \leq \gamma\}$ is compact (and convex).

Let us recall first the basic result concerning exponentially bounded positive definite functions. We shall only consider abelian semigroups with a neutral element.

Theorem (Berg/Maserick, cf. [2] or [1], 4.2.6 and 4.2.7). *For a semigroup S and an absolute value α on S the set $\mathcal{P}_1^\alpha(S)$ is a Bauer simplex with S^α as its* *set of extreme points. In other words, for any* $\varphi \in \mathcal{P}_1^\alpha(S)$ *there exists a unique Radon probability measure* μ *on* S^α *such that*

$$\varphi(s) = \int \sigma(s) d\mu(\sigma) \quad \forall\, s \in S\,.$$

We'll also make use of the following

COROLLARY (cf. [6], Proposition 1). *If* $\varphi \in \mathcal{P}^\alpha(S)$ *is completely positive definite, then the unique measure representing* φ *is concentrated on* S_+^α.

From now on we will typically deal with two semigroups R, S, and a mapping $t : R \longrightarrow S$ with the properties $t(0) = 0$, $t(r^-) = (t(r))^-$ and such that $t(R)$ generates S as a semigroup. Furthermore, a function $\beta : R \longrightarrow \mathbb{C}\backslash\{0\}$ is given with $\beta(0) = 1$ and $\beta(r^-) = \overline{\beta(r)}$ for all r; in most of the examples we'll have $\beta \equiv 1$. The direct product

$$R^{(\infty)} := \{(r_1, r_2, \dots) \in R^\infty \mid r_i = 0 \;\text{ finally}\}$$

of countably many copies of R will play a particular rôle.

The following result is new in this generality.

MAIN THEOREM. *Let* R *and* S *be semigroups, and* $t : R \longrightarrow S$, $\beta : R \longrightarrow \mathbb{C} \setminus \{0\}$ *be functions as just described*:

(i) *if* $\Phi(r_1, r_2, \dots) := \prod \beta(r_i) \cdot \varphi(\sum t(r_i))$ *for some function* $\varphi : S \longrightarrow \mathbb{C}$, *and* Φ *is positive definite then so is* φ;

(ii) *if furthermore* $|\Phi(r_1, r_2, \dots)| \le C \cdot \prod \gamma(r_i)$ *for some function* $\gamma : R \longrightarrow \mathbb{R}_+, \gamma(0) = 1$, *and some* $C > 0$, *then*

$$\alpha(s) := \inf \left\{ \prod \frac{\gamma(r_i)}{|\beta(r_i)|} \mid \sum t(r_i) = s \right\}$$

is an absolute value on S, φ *is* α*–bounded, and the measure* μ *representing* φ *is concentrated on*

$$W := \{\sigma \in S^\alpha \mid \beta \cdot (\sigma \circ t) \quad \textit{is positive definite on} \quad R\}$$

(iii) *conversely, for* $\mu \in M_+(W)$ *and* $\varphi(s) := \int \sigma(s) d\mu(\sigma)$ *the function* Φ *as defined in* (i) *is positive definite and fulfills* (ii) *for some* $C > 0$ *and some function* γ;

(iv) *a corresponding result holds for completely positive definite functions, the measure in* (ii) *being then concentrated on* W_+.

For the proof the following lemma is crucial. Since the statement here differs slightly from earlier presentations, we present it with the (short) proof. Recall that for a non–empty set M a function $\psi : M\times M \longrightarrow \mathbb{C}$ is a *positive semidefinite kernel* iff for any finite subset $\{x_1,\dots,x_k\}\subseteq M$ the matrix $(\psi(x_i,x_j))_{i,j\le k}$ is positive semidefinite.

Approximation lemma. *Let $p\ge 2$ be an integer, M a non-empty set, $\psi : M\times M \longrightarrow \mathbb{C}$ a positive semidefinite kernel, $(a_{ij})\in\mathbb{C}^{p\times p}$ a given $p\times p$–matrix. Suppose that for each $n\in\mathbb{N}$ there exist $\{x^n_{jm} \mid j=1,\dots,p; m=1,\dots,n\}\subseteq M$ such that*

$$\psi\left(x^n_{ik},x^n_{jm}\right)=a_{ij}\quad \forall\,(i,k)\neq(j,m)$$

and

$$\sup_{j,m,n}\psi\left(x^n_{jm},x^n_{jm}\right)<\infty\,.$$

Then (a_{ij}) is positive semidefinite.

Proof. Let $c_1,\dots,c_p\in\mathbb{C}$ be given; with $\{x^n_{jm}\}$ as indicated put $d_{jm}:=c_j/n$. Then

$$\begin{aligned}
0 &\le \sum_{i,j=1}^{p}\sum_{k,m=1}^{n} d_{ik}\bar d_{jm}\psi\left(x^n_{ik},x^n_{jm}\right)=\\
&=\sum_{\substack{i,j=1\\ i\neq j}}^{p} c_i\bar c_j a_{ij}+\frac{n^2-n}{n^2}\sum_{j=1}^{p}|c_j|^2 a_{jj}+\frac{1}{n^2}\sum_{j=1}^{p}\sum_{m=1}^{n}|c_j|^2\psi(x^n_{jm},x^n_{jm})=\\
&=\sum_{i,j=1}^{n} c_i\bar c_j a_{ij}+R_n
\end{aligned}$$

where $R_n:=-\frac{1}{n}\sum_{j=1}^{p}|c_j|^2a_{jj}+\frac{1}{n^2}\sum_{j=1}^{p}\sum_{m=1}^{n}|c_j|^2\psi(x^n_{jm},x^n_{jm})$ and so $R_n\longrightarrow 0$ for $n\longrightarrow\infty$, showing positive semidefiniteness of (a_{ij}). ∎

Proof of the Main Theorem.

(i) Let $s_1,\dots,s_p\in S$ and $n\in\mathbb{N}$ be given. By assumption

$$s_j=\sum_{\ell=1}^{q_j} t(r_{j\ell})\quad\text{for suitable}\quad r_{j\ell}\in R\,.$$

Let $\{N_{jm}\mid j=1,\dots,p; m=1,\dots,n\}$ be disjoint subsets of $\mathbb{N}$ with cardinalities $|N_{jm}|=q_j\ \forall\, j,m$, and define $x_{jm}\in R^{(\infty)}$ (for $N_{jm}=\{\nu_1,\dots,\nu_{q_j}\}$) by

$$x_{jm}(\nu_\ell):=r_{j\ell},\quad x_{jm}(i):=0\quad\text{for}\quad i\notin N_{jm}\,.$$

Put $\xi_j := \prod_{\ell=1}^{q_j} \beta(r_{j\ell})$, $j = 1, \dots, p$. Then for $(i,k) \neq (j,m)$

$$\Phi(x_{ik} + x_{jm}^-) = \xi_i \bar{\xi}_j \varphi(s_i + s_j^-)\,,$$

and $\Phi(x_{jm} + x_{jm}^-) = \prod_{\ell=1}^{q_j} \beta(r_{j\ell} + r_{j\ell}^-)\varphi(\sum_{\ell=1}^{q_j} t(r_{j\ell} + r_{j\ell}^-))$, independent of m and n. Hence by the Approximation lemma $(\xi_i \bar{\xi}_j \varphi(s_i + s_j^-))_{i,j \le p}$ is positive definite, and so is also $(\varphi(s_i + s_j^-))_{i,j \le p}$.
Suppose now Φ to be c.p.d., and let an additional element $a \in S$ be given, $a = t(r_1) + \ldots + t(r_v)$. Choose in the preceding argument the $N_{jm} \subseteq \mathbb{N} \setminus \{1, \dots, v\}$, and define $y \in R^{(\infty)}$ by $y(1) := r_1, \dots, y(v) := r_v, y(i) := 0$ else. Then for $(i,k) \neq (j,m)$

$$\Phi(y + x_{ik} + x_{jm}^-) = \xi_i \bar{\xi}_j \varphi(a + s_i + s_j^-) \cdot \prod_{\ell=1}^{v} \beta(r_\ell)$$

and if the positive semidefinite matrix on the RHS is not identically zero, $\prod_{\ell=1}^{v} \beta(r_\ell) > 0$ and then $(\varphi(a + s_i + s_j^-))$ is positive semidefinite, i.e. φ is completely positive definite.

(ii) If $s = \sum t(r_j)$ we get from

$$\Phi(r_1, r_2, \dots) = \prod \beta(r_j) \varphi(\sum t(r_j))$$

that

$$|\varphi(s)| \le C \cdot \prod \frac{\gamma(r_j)}{|\beta(r_j)|}\,,$$

hence

$$\frac{1}{C}|\varphi(s)| \le \alpha(s) := \inf \left\{ \prod \frac{\gamma(r_j)}{|\beta(r_j)|} \mid \sum t(r_j) = s \right\}$$

and α is immediately seen to be an absolute value. The function φ being positive definite and α–bounded, has a unique representing measure μ supported by the compact set S^α in view of the Berg/Maserick theorem. Define $f : S^* \longrightarrow \mathbb{C}$ by

$$(*) \qquad f(\sigma) := \sum_{u,v=1}^{w} c_u \bar{c}_v \beta(a_u + a_v^-) \sigma(t(a_u + a_v^-))$$

for given $a_1, \dots, a_w \in R$ and $c_1, \dots, c_w \in \mathbb{C}$. Then f is continuous, and on the compact subset S^α the function f is bounded. We want to show that f is μ–a.e. nonnegative, or equivalently that the measure $\nu := f \cdot \mu$ is nonnegative. This will be shown if $\hat{\nu}(s) := \int \sigma(s) d\nu(\sigma)$ turns out to be positive definite, by the Berg/Maserick theorem.

Let again $s_1, \dots, s_p \in S$, $d_1, \dots, d_p \in \mathbb{C}$ and $n \in \mathbb{N}$ be given, with

$$s_j = \sum_{\ell=1}^{q_j} t(r_{j\ell})$$

as in the proof of (i).
Let now $\{N_{ujm} \mid u = 1, \dots, w, j = 1, \dots, p; m = 1, \dots, n\}$ be disjoint subsets of $\mathbb{N} \setminus \{1\}$ with $|N_{ujm}| = q_j \ \forall\, u, j, m$, say $N_{ujm} = \{\nu_1, \dots, \nu_{q_j}\}$, and define $x_{ujm} \in R^{(\infty)}$ by

$$\begin{aligned} x_{ujm}(1) &:= a_u \\ x_{ujm}(\nu_\ell) &:= r_{j\ell} \\ x_{ujm}(i) &:= 0 \ \text{ else.} \end{aligned}$$

Put $\xi_j := \prod_{\ell=1}^{q_j} \beta(r_{j\ell})$. Then for $(u, i, k)) \neq (v, j, \ell)$

$$\Phi(x_{uik} + x_{vj\ell}^-) = \xi_i \bar{\xi}_j \beta(a_u + a_v^-) \varphi(t(a_u + a_v^-) + s_i + s_j^-)$$

and $\Phi(x_{uik} + x_{uik}^-)$ is again bounded uniformly in u, i, k, n. So again the matrix (with index set $A := \{1, \dots, w\} \times \{1, \dots, p\}$)

$$(\beta(a_u + a_v^-) \cdot \varphi(t(a_u + a_v^-) + s_i + s_j^-))_{(u,i),(v,j) \in A}$$

is positive semidefinite, leading to

$$\begin{aligned} \sum_{i,j=1}^{p} d_i \bar{d}_j \hat{\nu}(s_i + s_j^-) &= \sum_{i,j=1}^{p} \sum_{u,v=1}^{w} c_u d_i \bar{c}_v \bar{d}_j \times \\ &\quad \times \int \beta(a_u + a_v^-) \sigma(t(a_u + a_v^-) + s_i + s_j^-) d\mu(\sigma) = \\ &= \sum_{i,j} \sum_{u,v} c_u d_i \bar{c}_v \bar{d}_j \beta(a_u + a_v^-) \cdot \varphi(t(a_u + a_v^-) + s_i + s_j^-) \geq 0. \end{aligned}$$

We have shown $f \cdot \mu$ to be a positive measure , i.e.

$$\mu(f^{-1}(\mathbb{C} \setminus \mathbb{R}_+)) = 0\,.$$

Now $f^{-1}(\mathbb{C} \setminus \mathbb{R}_+)$ is open, and μ is a Radon measure, hence

$$\mu\left(\bigcup f^{-1}(\mathbb{C} \setminus \mathbb{R}_+)\right) = 0$$

the union being taken over all functions f of the form $(*)$. Hence μ–almost surely $r \longmapsto \beta(r) \cdot \sigma(t(r))$ is positive definite on R.

(iii) W is obviously closed, hence compact, so that $C := \mu(W) < \infty$. For $\sigma \in W$ the function

$$\Phi_\sigma(r_1, r_2, \dots) := \prod \beta(r_i) \cdot \sigma(t(r_i))$$

is positive definite as a (tensor) product of such functions. Also

$$|\beta(r)\sigma(t(r))| \leq |\beta(r)| \cdot \alpha(t(r)) =: \gamma(r)\,.$$

Now $\Phi = \int \Phi_\sigma d\mu(\sigma)$ is positive definite as a mixture of positive definite functions, and

$$|\Phi(r_1, r_2, \dots)| \leq C \cdot \prod \gamma(r_i) \quad \forall\, r_1, r_2, \dots \in R\,.$$

(iv) See the end of the proof of (i) . ∎

One of the most direct corollaries is the following result, characterizing *spherically exchangeable* (or *symmetric*) sequences, i.e. sequences of real random variables whose finite dimensional distributions are invariant under rotations.

THEOREM (Schoenberg, 1938). *Every infinite spherically exchangeable random sequence is a unique variance mixture of centered iid normal sequences. Or formally*:

$$\begin{aligned} &P \in M_+^1(\mathbb{R}^\infty) \quad \textit{is spherically symmetric} \\ \Longleftrightarrow\ &P = \int_0^\infty N(0,c)^\infty d\mu(c) \quad \exists! \mu \in M_+^1(\mathbb{R}_+)\,. \end{aligned}$$

PROOF. Only one direction needs a proof. Given a spherically exchangeable P we let Φ be its characteristic function, i.e.

$$\begin{aligned} \Phi(r_1, r_2, \dots) &:= E\left[\exp\left(i \sum r_j X_j\right)\right], \quad (r_1, r_2, \dots) \in \mathbb{R}^{(\infty)} = \\ &= \varphi\left(\sum r_j^2\right) \end{aligned}$$

for some function $\varphi : \mathbb{R}_+ \longrightarrow \mathbb{C}$, by assumption. With $t(r) = r^2, \beta \equiv \gamma \equiv 1$, we get from the Main Theorem that φ is a bounded positive definite function with $\varphi(0) = 1$. Then, for example applying the Berg/Maserick theorem, φ has the unique integral representation

$$\varphi(s) = \int e^{-\lambda s} d\mu(\lambda), \quad \mu \in M_+^1([0,\infty])$$

(with $e^{-\lambda\infty} = 1_{\{0\}}(\lambda), \lambda \in \mathbb{R}_+$). Now φ is obviously continuous, leading to $\mu(\{\infty\}) = 0$, and then from

$$\Phi(r_1, r_2, \dots) = \int_0^\infty e^{-\lambda \sum r_j^2} d\mu(\lambda)$$

we read off the wanted result. ∎

Schoenberg (cf. [8]) proved this result in the totally different connection of the imbedding problem for quasi-metric spaces into a Hilbert space.

With only slightly more effort we get the following characterization of

MIXTURES OF THE FULL 2-PARAMETER NORMAL FAMILY. Let $X = (X_1, X_2, \dots)$ be any real random sequence with characteristic function Φ. Then

$$\Phi(r_1, r_2, \dots) = \varphi\left(\sum r_j, \sum r_j^2\right) \quad \text{for some} \quad \varphi$$

iff

$$P^X = \int_{\mathbb{R}\times\mathbb{R}_+} N(a,c)^\infty d\mu(a,c) \quad \text{for some} \quad \mu \in M_+^1(\mathbb{R}\times\mathbb{R}_+)\,.$$

For a proof, see [6], Example 6.

Different transforms may of course be used. An example with Laplace transforms is this:

Let $X = (X_1, X_2, \dots)$ *be non–negative random variables. Then*

$$E\left[\exp\left(-\sum r_j X_j\right)\right] = \varphi\left(\prod (1+r_j)\right)\,.$$

For some $\varphi : [1,\infty[\longrightarrow \mathbb{R}$ *iff*

$$P^X = \int_0^\infty \gamma_\lambda^\infty d\mu(\lambda)$$

where γ_λ *denotes the Gamma* $(\lambda, 1)$ *distribution, with* $\gamma_1 = e_1$*, the exponential distribution with parameter* 1; cf. Example 8 in [6].

The natural question if mixtures of exponential iid sequences can be characterized similarly, can be answered immediately:

For a non–negative sequence X *we have*

$$P^X = \int_0^\infty e_\lambda^\infty d\mu(\lambda)$$

iff

$$P(X_1 \geq a_1, X_2 \geq a_2, \dots) = \varphi\left(\sum a_j\right), \quad a \in \mathbb{R}_+^{(\infty)}$$

for some $\varphi : \mathbb{R}_+ \longrightarrow \mathbb{R}$; cf. Example 11 in [6].

5 – De Finetti's theorem in extended form

Theorem. *Let $\mathcal{X}$ be a finite or countable set, S a semigroup, $t : \mathcal{X} \longrightarrow S$ such that $t(\mathcal{X})$ generates $S \smallsetminus \{0\}$, $\beta : \mathcal{X} \longrightarrow \,]0,\infty[$, $\varphi : S \longrightarrow \mathbb{R}_+$. Then $P \in M_+^1(\mathcal{X}^\infty)$ fulfills*

$$P(x_1,\dots,x_n) = \prod_{i=1}^n \beta(x_i)\cdot\varphi\left(\sum_{i=1}^n t(x_i)\right) \quad \forall\, n, x_i$$

iff

$$P = \int \kappa_\sigma^\infty \, d\mu(\sigma)$$

where $\mu \in M_+^1(S_+^*)$ is concentrated on

$$W := \left\{\sigma \in S_+^* \mid \kappa_\sigma := \beta\cdot(\sigma\circ t) \in M_+^1(\mathcal{X})\right\}$$

(*cf. Theorem* 4 *in* [6]*; we'll derive it here as a consequence of the Main Theorem*).

Proof. Let $R := \{1_{\{x\}} \mid x \in \mathcal{X}\} \cup \{0,1\}$ with pointwise multiplication, considered as a subsemigroup of $\mathbb{R}^{\mathcal{X}}$, add an absorbing element ζ to $S, S' := S \cup \{\zeta\}$, and define $t' : R \longrightarrow S'$ by $t'(1_{\{x\}}) := t(x), t'(1) := 0, t'(0) := \zeta$. Put $\beta'(1_{\{x\}}) := \beta(x), \beta'(1) := 1, \beta'(0) := 2$ (or any number > 1), $\varphi(\zeta) := 0$, and let $X_1, X_2, \dots$ be the natural projections $\mathcal{X}^\infty \longrightarrow \mathcal{X}$. Then

$$\begin{aligned}\Phi(r_1, r_2, \dots) :&= E[r_1(X_1)\cdot r_2(X_2)\cdot\ldots] = \\ &= \prod \beta'(r_j)\cdot\varphi\left(\sum t'(r_j)\right)\end{aligned}$$

for all $(r_1, r_2, \dots) \in R^{(\infty)}$.

Denoting the semigroup operation in $R^{(\infty)}$ by "$\oplus$" we get for $r^{(1)}, \dots, r^{(n)} \in R^{(\infty)}$ and $c_1, \dots, c_n \in \mathbb{R}$

$$\sum_{i,j=1}^n c_i c_j \Phi(r^{(i)} \oplus r^{(j)}) = E\left\{\left[\sum_{i=1}^n c_i r^{(i)}(X)\right]^2\right\} \geq 0$$

where $r(X) := r_1(X_1)\cdot r_2(X_2)\cdot\ldots$ for $r \in R^{(\infty)}$, showing Φ to be positive definite.

By the Main Theorem (i) φ is positive definite, and (ii) being fulfilled with $C = 1, \gamma \equiv 1, \varphi$ is α–bounded with

$$\alpha(s) = \inf\left\{\left(\prod\beta(x_i)\right)^{-1} \mid \sum t(x_i) = s\right\} \quad \text{for} \quad s \in S$$

and

$$\alpha(\zeta) = 0 \quad (\text{since } \ \beta'(0) > 1)\,.$$

Also, the measure μ' representing φ (on S') concentrates on $\{\sigma' \in (S')^\alpha \mid \beta' \cdot (\sigma' \circ t')$ is positive definite on $R\} =: V'$ and each $\sigma' \in V'$ is non-negative since R is idempotent, so $\sigma := \sigma' \mid_S \geq 0$. Let μ be the image of μ' under $\sigma' \longmapsto \sigma' \mid_S$, and $V := \{\sigma' \mid_S \mid \sigma' \in V'\}$. Then

$$\varphi(s) = \int_{V'} \sigma'(s) d\mu'(\sigma') = \int_V \sigma(s) d\mu(\sigma)$$

for $s \in S$, and

$$\begin{aligned} P(x_1, \dots, x_n) &= \Phi\left(1_{\{x_1\}}, 1_{\{x_2\}}, \dots, 1_{\{x_n\}}, 1, 1, \dots\right) = \\ &= \prod_{i=1}^n \beta(x_i) \cdot \varphi\left(\sum_{i=1}^n t(x_i)\right) = \\ &= \int_V \prod_{i=1}^n \beta(x_i) \sigma(t(x_i)) d\mu(\sigma)\,, \end{aligned}$$

leading to

$$1 = \sum_{x_1, \dots, x_n \in \mathcal{X}} P(x_1, \dots, x_n) = \int \left[\sum_{x \in \mathcal{X}} \beta(x) \sigma(t(x))\right]^n d\mu(\sigma)$$

for all $n \in \mathbb{N}$, which shows that μ is in fact concentrated on W. □

The technicalities in the above proof were perhaps slightly more complicated than expected, but then calculations with (Fourier, Laplace) transforms are often easier than those with the distributions themselves ... The following examples will show, however, that the result is easy to apply.

Example 4.1. The original De Finetti theorem: here $\mathcal{X} = \{0, 1\}, S = \{(k, n) \in \mathbb{N}_0^2 \mid k \leq n\}$ (cf. Section 1), $t(x) = (x, 1), \beta \equiv 1$. A general non–negative character on S has the form $\sigma(k, n) = p^k q^{n-k}$ with $p, q \geq 0$. The condition $\sigma \circ t \in M_+^1(\mathcal{X})$ translates into

$$\sigma(t(0)) + \sigma(t(1)) = \sigma(0, 1) + \sigma(1, 1) = q + p = 1$$

which gives the result.

Example 4.2. A slight extension of 4.1. We consider $\mathcal{X} = \{0, 1, 2, \dots, k\}$, where $k \in \mathbb{N}$. Let again $P \in M_+^1(\mathcal{X}^\infty)$ fulfill

$$P(x_1, \dots, x_n) = \varphi_n\left(\sum_{i=1}^n x_i\right) = \varphi\left(\sum_{i=1}^n (x_i, 1)\right)$$

as before. Then

$$P = \int_0^1 \kappa_p^\infty \, d\mu(p)$$

with

$$\kappa_p(\{j\}) = p^j q^{k-j}, \quad q = q(p) \quad \text{from} \quad p^k + p^{k-1}q + \ldots + pq^{k-1} + q^k = 1\,.$$

Example 4.3. A further "extension": $\mathcal{X} = \mathbb{N}_0, P$ as before. Then

$$P = \int_{]0,1]} \gamma_a^\infty \, d\mu(a)$$

γ_a denoting the geometric distribution with parameter a, i.e. $\gamma_a(\{k\}) = a(1-a)^k$.

Example 4.4. $\mathcal{X} = \mathbb{N}_0$ as before, $P \in M_+^1(\mathcal{X}^\infty)$. Then

$$P(x_1, \ldots, x_n) = \frac{1}{\prod_{i=1}^n x_i!} \cdot \varphi_n\left(\sum_{i=1}^n x_i\right)$$

iff

$$P = \int_0^\infty \pi_\lambda^\infty \, d\mu(\lambda)$$

where π_λ denotes the Poisson distribution with parameter λ. Here we have for the first time the non–trivial function $\beta(x) = 1/x!$. The choice $\beta(x) = 1/(x+1)$ leads instead to mixtures of

$$\kappa_u(\{x\}) := \frac{1}{-\log(1-u)} \cdot \frac{u^{x+1}}{x+1} \quad (0 < u < 1) \quad \text{and} \quad \kappa_0 = \varepsilon_0\,,$$

and $\beta(x) = \binom{x+r-1}{r-1}$ would lead to negative binomials.

6 – Abstract results

De Finetti's original result dealt with $\{0,1\}$–valued random variables, and was generalized by him a few years later to the real-valued case. In 1955 a considerable further extension to arbitrary compact Hausdorff spaces was presented:

Theorem (Hewitt–Savage). *Let $\mathcal{X}$ be a compact Hausdorff space. Then $P \in M_+^1(\mathcal{X}^\infty)$ is exchangeable iff*

$$P = \int \kappa^\infty \, d\mu(\kappa)$$

for some $\mu \in M_+^1(M_+^1(\mathcal{X}))$.

[Here $M_+^1(\mathcal{X})$ is by definition the set of all Radon probability measures on $\mathcal{X}$; and $M_+^1(\mathcal{X})$ is given the usual weak topology in which it is again compact.]

In the following proof we'll use the notion of the *free abelian semigroup* (without involution) over a set A, denoted $\mathbb{N}_0^{(A)}$, and consisting of all functions $s : A \longrightarrow \mathbb{N}_0$ such that $\{s > 0\}$ is finite, with the usual addition.

A character σ on $\mathbb{N}_0^{(A)}$ can be identified with a function $\tau : A \longrightarrow \mathbb{R}$ via $\tau(a) = \sigma(\delta_a)$, where $\delta_a := 1_{\{a\}} \in \mathbb{N}_0^{(A)}$.

Proof. Let $A := \{f : \mathcal{X} \longrightarrow [0,1] \mid f \text{ is continuous}\}$. Then for any finite collection of $f_1, f_2, \dots \in A$

$$E\left[\prod f_j(X_j)\right] = \varphi\left(\sum \delta_{f_j}\right)$$

with φ being defined on $\mathbb{N}_0^{(A)}$. By the Main Theorem, φ is c.p.d. (and bounded), so

$$\varphi\left(\sum \delta_{f_j}\right) = \int \prod \tau(f_j) d\mu(\tau)$$

for some $\mu \in M_+^1([0,1]^A)$.

An easy argument shows μ to be concentrated on

$$T := \{\tau : A \longrightarrow [0,1] \mid \tau(1) = 1, \tau \quad \text{finitely additive}\}$$

(cf. [6], Theorem 2) , and each $\tau \in T$ extends uniquely to a positive linear functional on $C(\mathcal{X})$, i.e. τ can be identified with a Radon probability measure on $\mathcal{X}$. Inserting this above gives the desired result. $\square$

Remark 1. If $\mathcal{X}$ was just a measurable space then with $A := \{f : \mathcal{X} \longrightarrow [0,1] \mid f \text{ measurable}\}$ one obtains

$$E\left[\prod f_j(X_j)\right] = \int \prod \tau(f_j) d\mu(\tau),\ f_j \in A\,,$$

with $\mu \in M_+^1(T)$ and

$$T := \{\tau : A \longrightarrow [0,1] \mid \tau(1) = 1, \tau \quad \text{additive}\}$$

which is a "weak" form of a general De Finetti type result.

Remark 2. As noted above, the Berg/Maserick theorem is an essential ingredient in the proof of the Main Theorem. It can however also be deduced from it: if $\varphi : S \longrightarrow \mathbb{C}$ is p.d. and α–bounded then $\Phi(s_1, s_2, \dots) := \varphi(\sum s_j)$ is p.d., and

$$\mid \Phi(s_1, s_2, \dots) \mid \leq C \cdot \prod \alpha(s_j)\,.$$

With $R = S$, $t = id_S$ and $\beta \equiv 1$ the set W in the Main Theorem reduces to S^α.

Remark 3. The Main Theorem can be looked at as a result on *exchangeable p.d. functions* (here for simplicity we assume S without involution): let $\Phi : R^{(\infty)} \longrightarrow \mathbb{R}$ be p.d. and exchangeable. This leads to a factorization of the form

$$\Phi(r_1, r_2, \dots) = \varphi\left(\sum \delta_{r_j}\right)$$

with $\varphi : \mathbb{N}_0^{(R \setminus \{0\})} \longrightarrow \mathbb{R}$ (and $\delta_0 := 0$). Then φ is p.d., and if $|\Phi(r_1, r_2, \dots)| \leq C \cdot \prod \gamma(r_j)$ for some function $\gamma : R \longrightarrow \mathbb{R}_+, \gamma(0) = 1$, and some $C > 0$, the function φ is α-bounded with $\alpha(\sum \delta_{r_j}) := \prod \gamma(r_j)$ – so α is even a character. We get

$$\varphi\left(\sum \delta_{r_j}\right) = \int \sigma\left(\sum \delta_{r_j}\right) d\mu(\sigma)$$

where μ is a Radon measure on all characters σ of $\mathbb{N}_0^{(R \setminus \{0\})}$ with $|\sigma| \leq \alpha$. Since such a σ can be identified with the function τ on $R \setminus \{0\}$ given by $\tau(r) := \sigma(\delta_r)$, completed by $\tau(0) := 1$, we see that μ can be considered as a measure on

$$W := \{\tau \in \mathcal{P}_1(R) \mid |\tau| \leq \gamma\},$$

leading to

$$\Phi(r_1, r_2, \dots) = \int_W \prod \tau(r_j) d\mu(\tau),$$

a mixture of tensor powers of p.d. functions on R.

Note that in the special case where $\Phi(r_1, r_2, \dots) = \varphi(\sum r_j)$ depends on the sum of the entries, the function $\varphi : R \longrightarrow \mathbb{R}$ is automatically p.d., so that if furthermore φ is a *moment function* (i.e. a mixture of characters) Φ would be the corresponding mixture of infinite tensor powers of characters.

7 – A different point of view

Let's take another look at the Main Theorem (with $\beta \equiv 1$) :

$$\Phi(r_1, r_2, \dots) = \varphi\left(\sum t(r_j)\right)$$

and the conclusion Φ p.d. $\Longrightarrow \varphi$ p.d.

Put $U := R^{(\infty)}, \psi(r_1, r_2, \dots) := \sum t(r_j)$, then $\psi : U \longrightarrow S$ is onto and the theorem says: $\boxed{\varphi \circ \psi \text{ p.d.} \implies \varphi \text{ p.d.}}$

What is the crucial property of ψ enabling this conclusion?

The answer looks complicated:

$$\left|\begin{array}{l} \forall \text{ finite subsets } \{s_1, \dots, s_n\} \subseteq S \text{ and } \{u_1, \dots, u_m\} \subseteq U \text{ and} \\ \forall\, N \in \mathbb{N} \quad \exists\, \{u_{jp\alpha} \mid j \leq n, p \leq m, \alpha \leq N\} \subseteq U \text{ such that} \\ \psi(u_{jp\alpha} + u_{kq\beta}^-) = s_j + s_k^- + \psi(u_p + u_q^-) \text{ for } (j,p,\alpha) \neq (k,q,\beta) \end{array}\right.$$

If this is fulfilled, and $\psi(0) = 0$, we call ψ *strongly almost additive.*

This holds for example if ψ is a homomorphism and onto, but this case is not too interesting.

In this more general framework we shall deal only with bounded functions, being no restriction for the applications we have in mind.

THEOREM. *Let U, S be two semigroups, $\psi : U \longrightarrow S$ be strongly almost additive, and $\varphi : S \longrightarrow \mathbb{C}$ bounded. Then*

$$\varphi \circ \psi \ p.d. \ \Longrightarrow \varphi \ p.d.$$

and φ is in fact a mixture of characters in

$$\hat{S}_\psi := \{\sigma \in \hat{S} \mid \sigma \circ \psi \ p.d.\}$$

(n.b.: a compact subsemigroup of $\hat{S}$).

Furthermore:

$$\{\varphi : S \longrightarrow \mathbb{C} \mid \varphi \text{ bounded}, \varphi(0) = 1, \varphi \circ \psi \ p.d.\}$$

is a Bauer simplex with $\hat{S}_\psi$ as extreme points.

Cf. [7], Theorem 1, where a slightly more general result is shown.

We shall apply this theorem to socalled exchangeable random partitions of the positive integers.

$V = \{v_1, v_2, \ldots\}$ is a *partition* of $\mathbb{N} :\Longleftrightarrow v_j \neq \emptyset, v_j \cap v_k = \emptyset$ for $j \neq k$, and $\bigcup_j v_j = \mathbb{N}$.

Examples are $\{\{i\} \mid i \in \mathbb{N}\}$ or $\{\mathbb{N}\}$, the two "extreme" partitions of $\mathbb{N}$.

Let $\mathcal{P}$ denote the set of all partitions of $\mathbb{N}$. Any $V \in \mathcal{P}$ can be identified with the equivalence relation $E(V) := \bigcup_{v \in V} v \times v \subseteq \mathbb{N}^2$ or with $1_{E(V)} \in \{0,1\}^{\mathbb{N}^2}$, this last identification defining the (natural) topology on $\mathcal{P}$, turning it into a compact metric space.

For $A \subseteq \mathbb{N}$ and $V \in \mathcal{P}$ we write

$$A \sqsubseteq V :\Longleftrightarrow \exists\, v \in V \quad \text{with} \quad A \subseteq v$$

(that is: A is not separated by the classes of V).

For $U, V \in \mathcal{P}$ we define

$$U \leq V :\Longleftrightarrow u \sqsubseteq V \quad \forall\, u \in U \quad [\Longleftrightarrow E(U) \subseteq E(V)] \ .$$

Every subset of $\mathcal{P}$ has a unique minimal element w.r. to "$\leq$ ", and for a family $\mathcal{A}$ of subsets of $\mathbb{N}$ there is a smallest $W \in \mathcal{P}$ such that $A \sqsubseteq W$ for each $A \in \mathcal{A}$. In the particular case of $\mathcal{A} = U \cup V$ for $U, V \in \mathcal{P}$ we write $U \vee V$ for this minimum, and call it (of course) their *maximum*.

The order intervals $P_U := \{W \in \mathcal{P} \mid U \leq W\}$ fulfill $P_U \cap P_V = P_{U \vee V}$. For $U \in \mathcal{P}$ the classes $u \in U$ with $|u| \geq 2$ are called *non-trivial*, their union $\langle U \rangle$ is called the *support* of U. Obviously $\langle U \vee V \rangle \subseteq \langle U \rangle \cup \langle V \rangle$, so that

$$\mathcal{U} := \{U \in \mathcal{P} \mid \langle U \rangle \text{ is finite}\}$$

is a subsemigroup w.r. to "$\vee$ ", with neutral element $U_0 = \{\{j\} \mid j \in \mathbb{N}\}$. The order intervals P_U for $U \in \mathcal{U}$ are clopen and generate the Borel sets of $\mathcal{P}$. Probability measures on $\mathcal{P}$ will be called *random partitions*.

THEOREM. *$\varphi : \mathcal{U} \longrightarrow \mathbb{R}$ is p.d. and normalized (i.e. $\varphi(U_0) = 1$) $\Longleftrightarrow \exists$ (unique) random partition $\mu \in M_+^1(\mathcal{P})$ with*

$$\varphi(U) = \mu(P_U) \quad \forall U \in \mathcal{U}$$

[*cf.* [4], *Theorem* 1].

Note that the easy direction "$\Longleftarrow$" follows immediately from

$$\sum_{j,k=1}^{n} c_j c_k \varphi(U_j \vee U_k) = \int \left(\sum_{j=1}^{n} c_j 1_{P_{U_j}} \right)^2 d\mu \ \geq 0\,.$$

A permutation π of $\mathbb{N}$ induces $\overline{\pi} : \mathcal{P} \longrightarrow \mathcal{P}$, $\overline{\pi}(V) := \{\pi(v) \mid v \in V\}$, and $\overline{\pi}$ is continuous. π is *finite* iff $\{i \in \mathbb{N} \mid \pi(i) \neq i\}$ is finite.

DEFINITION. $\mu \in M_+^1(\mathcal{P})$ is *exchangeable* $:\Longleftrightarrow \mu^{\overline{\pi}} = \mu \ \forall$ finite π.

Now $\mu^{\overline{\pi}}(P_U) = \mu(P_{\overline{\pi^{-1}(U)}})$, so μ is exchangeable iff

$$\mu(P_U) = \mu(P_V)$$

$\forall U, V \in \mathcal{U}$ with $|\{u \in U \mid |u| = k\}) = |\{v \in V | |v| = k\}|$ for $k = 2, 3, \dots$ iff $\mu(P_U) = \varphi \circ g(U)$ for some φ defined on the semigroup

$$S := \mathbb{N}_0^{(\{2,3,\dots\})}\,,$$

with $g(U) := \sum_{\substack{u \in U \\ |u| \geq 2}} \delta_{|u|}$.

This function $g : \mathcal{U} \longrightarrow S$ is in fact strongly almost additive, cf. [4], Lemma 5.

THEOREM (Kingman). *$M_+^{1,e}(\mathcal{P}) := \{\mu \in M_+^1(\mathcal{P}) | \mu$ exchangeable$\}$ is a Bauer simplex whose extreme points are precisely those μ for which*

$$\mu(P_U) = \sigma(g(U)), U \in \mathcal{U} \quad \textit{with} \quad \sigma \in \hat{S}_+\,.$$

Such a character σ is given by a sequence $(t_2, t_3, \dots)$ in $[0,1]$, and we see that

$$t_n = \mu(P_{\{\{1,\dots,n\},\{n+1\},\{n+2\},\dots\}}), \quad n \geq 2$$

is the μ–probability for $\{1, \dots, n\}$ not getting separated. For general $U \in \mathcal{U}$ the multiplicativity of σ is reflected in a certain pattern of independence:

$$\mu(P_U) = \prod_{\substack{u \in U \\ |u| \geq 2}} t_{|u|}\,.$$

Kingman ([5]) showed that there exists a sequence $x = (x_1, x_2, \dots)$ with $x_i \geq 0, \sum x_i \leq 1$, such that

$$t_n = \sum_{i=1}^{\infty} x_i^n \quad \text{for} \quad n = 2, 3, \dots$$

is the associated sequence of power sums.

Using x, there is in fact a natural way to describe this distribution μ : put $x_0 := 1 - \sum_{i=1}^{\infty} x_i$ and let $X_1, X_2, \dots$ be iid with $P(X_1 = i) = x_i, i \geq 0$. Then

$$G := \{\{j \in \mathbb{N} \mid X_j = c\} \mid c \in \mathbb{N}\} \cup \{\{i\} \mid X_i = 0\} \setminus \{\emptyset\}$$

is $\mathcal{P}$–valued and has μ as its distribution.

Acknowledgements

The author is very grateful to the organizing committee of the *Bruno de Finetti Centenary Conference* (Rome, November 2006), in particular to Professor Fabio Spizzichino, for the invitation to give at this highly interesting congress, taking place in the fascinating "Eternal City", a talk of which the present paper is an elaborate version.

Many thanks are due also to the referee(s) for a particularly thorough revision of my first draft, resulting in many improvements.

REFERENCES

[1] C. Berg – J. P. R. Christensen – P. Ressel: *Harmonic Analysis on Semigroups*, Springer–Verlag, Berlin/Heidelberg/New York, 1984.

[2] C. Berg – P. H. Maserick: *Exponentially bounded positive definite functions*, Illinois J. Math., **28** (1984), 162–179.

[3] B. De Finetti: *Funzione caratteristica di un fenomeno aleatorio*, Memorie della R. Accademia Nazionale dei Lincei, **V** (IV) (1930), 86–133.

[4] U. Hirth – P. Ressel: *Random partitions by semigroup methods*, Semigroup Forum, **59** (1999), 126–140.

[5] J. F. C. Kingman: *The representation of partition structures*, J. London Math. Soc., **18** (1978), 374–380.

[6] P. Ressel: *De Finetti–type theorems: an analytical approach*, Ann. Prob., **13** (1985), 898–922.

[7] P. Ressel: *Integral representations for distributions of symmetric stochastic processes*, Prob. Th. Rel. Fields, **79** (1988), 451–467.

[8] I. J. Schoenberg: *Metric spaces and positive definite functions*, Trans. Amer. Math. Soc., **44** (1938), 522–536.

Lavoro pervenuto alla redazione il 8 ottobre 2007
ed accettato per la pubblicazione il 12 ottobre 2007.
Bozze licenziate il 9 aprile 2008

Indirizzo dell'Autore:

Paul Ressel – Universität Eichstätt – Osten str. 26-28 D 85071 Eichstatt, Germany
E-mail: paul.ressel@ku-eichstaett.de

Rendiconti di Matematica, Serie VII
Volume 28, Roma (2008), 83–95

Exchangeable Rasch matrices

STEFFEN L. LAURITZEN

Abstract: *This article is concerned with binary random matrices that are exchangeable with the probability of any finite submatrix only depending on its row- and column sums. We describe basic representations of such matrices both in the case of full row- and column exchangeability and the case of weak exchangeability. Finally the results are interpreted in terms of random graphs with exchangeable labels and with a view towards their potential application to social network analysis.*

1 – Introduction

Let us initially consider the sequential case and recall that a sequence of random variables $X_1, \ldots, X_n, \ldots$ is said to be *exchangeable* if for all n and $\pi \in S(n)$ it holds that

$$X_1, \ldots, X_n \stackrel{\mathcal{D}}{=} X_{\pi(1)}, \ldots, X_{\pi(n)}$$

where $S(n)$ is the group of permutations of $\{1, \ldots, n\}$ and $\stackrel{\mathcal{D}}{=}$ denotes equality in distribution; i.e. the distribution of an exchangeable sequence of random variables is unchanged whenever the order of any finite number of them is rearranged.

de Finetti's theorem for binary sequences [1] then says that *a binary sequence $X_1, \ldots, X_n, \ldots$ is exchangeable if and only if there exists a distribution function*

Key Words and Phrases: *Partial exchangeability – Random graphs – Social network analysis – Sufficiency.*
A.M.S. Classification: 62B99, 60C05
This paper was presented at the Bruno de Finetti Centenary Conference in Rome, November 2006.

F on $[0,1]$ such that for all n

$$p(x_1,\dots,x_n)=\int_0^1 \theta^{t_n}(1-\theta)^{n-t_n}\,dF(\theta)\,,$$

where $t_n=\sum_{i=1}^n x_i$. Further, the limiting frequency $Y=\lim_{n\to\infty}\sum_{i=1}^n X_i/n$ exists and F is the distribution function of this limit. Conditionally on $Y=\theta$, the sequence $X_1,\dots,X_n,\dots$ are independent and identically distributed with expectation θ.

An alternative formulation of exchangeability focuses on its relationship to sufficiency [2], [3], [4]. A statistic $t(x)$ is *summarizing* for p [5] if for some ϕ

$$p(x)=\phi(t(x))\,.$$

Note that if t is summarizing for all p in a family $\mathcal{P}$ of distributions, it is sufficient for $\mathcal{P}$.

For binary variables, $X_1,\dots,X_n,\dots$ is exchangeable if and only if for all n

$$P(X_1=x_1,\dots,X_n=x_n)=\phi_n\left(\sum_{i=1}^n x_i\right)$$

i.e., if and only if $t_n=\sum_{i=1}^n x_i$ is summarizing for its distribution. This is due to the fact that $t_n=\sum_{i=1}^n x_i$ is the maximal invariant for the action of the permutation group $S(n)$ on the binary sequences of length n and $S(n)$ acts transitively on the sequences with given value of t_n, so that any two such sequences are permutations of each other.

The present paper is investigating the interplay between these ideas in the case of random binary matrices where the situation is somewhat more complex.

The next section is initially giving an overview of results in [6] and the reader is referred to this paper for details not described here. Some of the considerations of [6] are further extended to the case of weakly exchangeable arrays. The last section touches upon the relation of these to random graphs and social network analysis.

2 – Exchangeable binary matrices

2.1 – Random Rasch matrices

The Rasch model [7] was originally developed to analyse data from intelligence tests where X_{ij} indicates a binary outcome when problem i was attempted by person j, $X_{ij}=1$ denoting success and $X_{ij}=0$ denoting failure. The model is described by 'easinesses' $\alpha = (\alpha_i)_{i=1,\dots}$ and 'abilities' $\beta = (\beta_j)_{j=1,\dots}$ so that binary responses X_{ij} are conditionally independent given (α, β) and

$$P(X_{ij} = 1 \mid \alpha, \beta) = 1 - P(X_{ij} = 0 \mid \alpha, \beta) = \frac{\alpha_i \beta_j}{1 + \alpha_i \beta_j}\,.$$

The model is a potential model for a large variety of phenomena such as, for example, a batter i getting a hit against a pitcher j in baseball matches [8] or the occurrence of species i on island j [9]; see for example [10] for a survey.

A *random Rasch matrix* has (α_i) i.i.d. with distribution A, (β_j) i.i.d. with distribution B, the entire sequences α and β also being independent of each other. Such a model would be relevant if each of batters and pitchers were *a priori* exchangeable. An example of a random Rasch matrix is displayed in fig. 1.

Fig. 1: Two random matrices, each of dimension 100×100. The matrix to the left is a random Rasch matrix and thus RCES whereas the matrix to the right is RCE but not RCS. The two matrices have the same overall mean equal to 0.5.

2.2 – Exchangeable and summarized matrices

Recall that a doubly infinite matrix $X = \{X_{ij}\}_{1,1}^{\infty,\infty}$ is *row–column exchangeable* [11] (an RCE-matrix) if for all m, n, $\pi \in S(m)$, $\rho \in S(n)$

$$\{X_{ij}\}_{1,1}^{m,n} \stackrel{\mathcal{D}}{=} \{X_{\pi(i)\rho(j)}\}_{1,1}^{m,n}\,,$$

i.e. if the distribution is unchanged when rows or columns are permuted. A random Rasch matrix is clearly RCE.

A doubly infinite (binary) matrix $X = \{X_{ij}\}_{1,1}^{\infty;\infty}$ is said to be *row-column summarized* (RCS-matrix) if for all m, n

$$p(\{x_{ij}\}_{1,1}^{m,n}) = \phi_{m,n}\{r_1, \dots, r_m; c_1, \dots, c_n\},$$

where $r_i = \sum_j x_{ij}$ and $c_j = \sum_j x_{ij}$ are the row- and column sums.

In contrast to the sequential case, there is no simple summarizing statistic for an RCE matrix. RCE-matrices are generally not RCS-matrices and *vice versa* because the group G_{RC} of row- and column permutations does not act transitively on matrices with fixed row- and column sums. To see the latter, consider

$$M_1 = \begin{Bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 1 \end{Bmatrix}, \quad M_2 = \begin{Bmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{Bmatrix}.$$

The matrices M_1 and M_2 have identical row- and column sums. However, their determinants are different: $|\det M_1| = 1$ whereas $|\det M_2| = 0$. Since the absolute value of the determinant is invariant under row- and column permutations, it follows that no combination of such permutations will ever modify M_1 to become M_2.

If a matrix is both RCE and RCS, we say that it is an *RCES-matrix.* Random Rasch matrices are RCES matrices since, conditionally on (α, β), we have

$$p(\{x_{ij}\}_{1,1}^{m,n} \mid \alpha, \beta) = \prod_{i=1}^{m} \prod_{j=1}^{n} \frac{(\alpha_i \beta_j)^{x_{ij}}}{1 + \alpha_i \beta_j} = \frac{\prod_{i=1}^{m} \alpha_i^{r_i} \prod_{j=1}^{n} \beta_j^{c_j}}{\prod_{i=1}^{m} \prod_{j=1}^{n} (1 + \alpha_i \beta_j)}, \tag{1}$$

which only depends on row- and column sums, implying that this also holds after taking expectation w.r.t. (α, β).

The difference between an RCE and RCS matrix is not always immediately visible. Figure 1 displays a random Rasch matrix next to an RCE matrix which is not RCS, yet they are not easily distinguishable.

2.3 – de Finetti's theorem for RCE matrices

The set of distributions $\mathcal{P}_{\mathrm{RCE}}$ of binary RCE matrices is a convex simplex. In particular, *every* $P \in \mathcal{P}_{\mathrm{RCE}}$ has a unique representation as a mixture of extreme points $\mathcal{E}_{\mathrm{RCE}}$ of $\mathcal{P}_{\mathrm{RCE}}$, i.e.

$$P(A) = \int_{\mathcal{E}} Q(A) \mu_P(Q).$$

The same holds if RCE is replaced by RCS or RCES. In addition, it can be shown that

$$\mathcal{E}_{\mathrm{RCES}} = \mathcal{E}_{\mathrm{RCE}} \cap \mathcal{P}_{\mathrm{RCS}}.$$

The extreme measures are particularly simple. Aldous [11] shows that *for any* $P \in \mathcal{P}_{\text{RCE}}$ *the following are equivalent*:

- $P \in \mathcal{E}_{\text{RCE}}$;
- *the tail* σ*-field* $\mathcal{T}$ *is trivial*;
- *the corresponding RCE-matrix* X *is dissociated.*

Here the *tail* $\mathcal{T}$ is

$$\mathcal{T} = \bigcap_{n=1}^{\infty} \sigma\{X_{ij}, \min(i,j) \geq n\}$$

and a matrix is *dissociated* if for all A_1, A_2, B_1, B_2 with $A_1 \cap A_2 = B_1 \cap B_2 = \emptyset$

$$\{X_{ij}\}_{i \in A_1, j \in B_1} \perp\!\!\!\perp \{X_{ij}\}_{i \in A_2, j \in B_2} .$$

Following [12], a binary doubly infinite random matrix X is a ϕ*-matrix* if X_{ij} are independent given $U = (U_i)_{i=1,\dots}$ and $V = (V_j)_{j=1,\dots}$ where U_i and V_j are independent and uniform on $(0,1)$ and

$$P(X_{ij} = 1 \,|\, U = u, V = v) = \phi(u_i, v_j) .$$

In [11], [12], [13] it is shown that *distributions of* ϕ*-matrices are the extreme points of* $\mathcal{P}_{\text{RCE}}$*,* i.e. binary RCE matrices are mixtures of ϕ-matrices. Different ϕ may in general have identical distributions of their ϕ-matrix. Clearly, if (g,h) is a pair of measure-preserving transformations of the unit interval into itself, $\tilde{\phi}(u,v) = \phi(g(u), h(v))$ yields the same distribution of X as ϕ. In fact, ϕ is exactly determined up to such a pair of measure-preserving transformations [14].

2.4 – Rasch type ϕ-matrices

As shown in [6], if a ϕ-matrix is also RCS, then

$$P\left(\begin{Bmatrix} 1 & 0 \\ 0 & 1 \end{Bmatrix} \,\Big|\, U = u, V = v\right) = P\left(\begin{Bmatrix} 0 & 1 \\ 1 & 0 \end{Bmatrix} \,\Big|\, U = u, V = v\right)$$

which holds if and only if ϕ is of *Rasch type*, i.e. if for all u, v, u^*, v^*:

$$\phi(u,v)\bar{\phi}(u,v^*)\bar{\phi}(u^*,v)\phi(u^*,v^*) = \bar{\phi}(u,v)\phi(u,v^*)\phi(u^*,v)\bar{\phi}(u^*,v^*) , \tag{2}$$

where we have let $\bar{\phi} = 1 - \phi$. This is the *Rasch functional equation* [15].

Although RCE matrices have no simple summarizing statistics, RCES-matrices do: they are summarized by the empirical distributions of row- and column sums:

$$t_{mn} = \left(\sum_{i=1}^{m} \delta_{r_i}, \sum_{j=1}^{n} \delta_{c_j}\right) ,$$

where δ_s is the measure with unit mass in s. This is a semigroup statistic, and RCES matrices can also be represented via mixtures of characters on the image semigroup (Ressel 2002, personal communication). General solutions of the Rasch functional equation thus represent characters of the image semigroup of the empirical row- and column sum measures.

Lauritzen (2003) [6] shows that *any RCES matrix is a mixture of Rasch type ϕ-matrices* and also that *any regular RCES matrix is a mixture of random Rasch matrices.* Here, a random binary matrix is said to be *regular* if

$$0 < P(X_{ij} = 1 \,|\, \mathcal{S}) < 1 \text{ for all } i, j\,,$$

where the *shell σ-algebra* $\mathcal{S}$ is

$$\mathcal{S} = \bigcap_{n=1}^{\infty} \sigma\{X_{ij}, \max(i,j) \geq n\}\,.$$

Regular solutions ($0 < \phi < 1$) to the Rasch functional equation are all of the form

$$\phi(u,v) = \frac{a(u)b(v)}{1 + a(u)b(v)}\,,$$

where a and b are positive real-valued functions on the unit interval, leading to random Rasch models.

The matrix to the right in fig. 1 is a ϕ-matrix with $\phi(u,v) = (u+v)/2$. Since this does not satisfy Rasch's functional equation it is not RCS. The matrix to the left is similarly a Rasch matrix with $\phi = 6.49186uv/(1 + 6.49186uv)$. The two matrices have the same overall mean equal to 0.5.

There are interesting non-regular solutions to the Rasch equation, for example

$$\phi(u,v) = \chi_{\{u \leq v\}} = \begin{cases} 1 & \text{if } u \leq v \\ 0 & \text{otherwise.} \end{cases}$$

A corresponding ϕ-matrix is displayed in fig. 2. But there are also solutions such as, for example,

$$\phi(u,v) = \begin{cases} \dfrac{a(u)b(v)}{1 + a(u)b(v)} & \text{if } 1/3 < u, v < 2/3 \\ \chi_{\{u \leq v\}} & \text{otherwise.} \end{cases}$$

Both of these non-regular solutions imply the existence of incomparable groups, so that some questions are always answered correctly for a subgroup of the persons and some questions never answered by some. More complex variants of the latter example lead to Cantor–Rasch matrices, see [6] for further details.

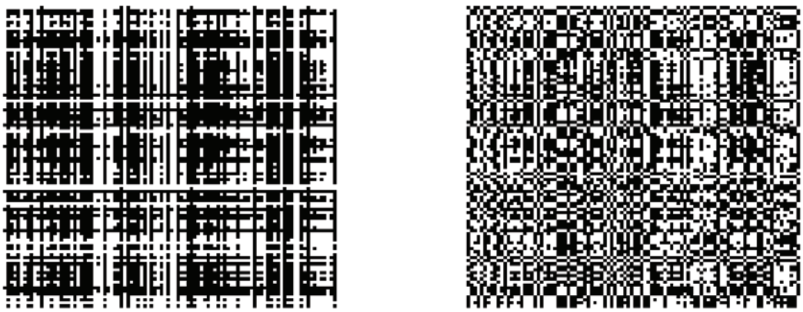


Fig. 2: The left-hand matrix is a non-regular RCES ϕ-matrix with columns $\phi(u,v) = \chi_{\{u\le v\}}$. The matrix on the right-hand side is a ϕ-matrix with $\phi(u,v) = \chi_{\{|u-v|\le 1-1/\sqrt{2}\}}$. It is RCE but not RCS. The two matrices have the same overall mean equal to 0.5.

The difference between RCE and RCES can be striking if the corresponding matrix is manipulated by sorting the rows and columns by their row- and column sums, as shown in fig. 3, where both diagrams have been obtained from fig. 2 in this way.

Fig. 3: The matrices are obtained from those displayed in fig. 2 by sorting rows and columns according to their sum. The left-hand matrix was a non-regular RCES matrix. The right-hand matrix was RCE but not RCS.

2.5 – WE matrices

A random matrix is said to be *weakly exchangeable* [16], [17] (a WE-matrix) if for all n and $\pi \in S(n)$

$$\{X_{ij}\}_{1,1}^{n,n} \stackrel{\mathcal{D}}{=} \{X_{\pi(i)\pi(j)}\}_{1,1}^{n,n},$$

i.e. if the distribution of X is unchanged when rows and columns are permuted *using the same* permutation for rows and columns. Similarly we say that a doubly infinite (binary) matrix $X = \{X_{ij}\}_{1,1}^{\infty,\infty}$ is *weakly summarized* (WS-matrix) if for all n

$$p(\{x_{ij}\}_{1,1}^{n,n}) = \phi_n\{r_1 + c_1, \dots, r_n + c_n\},$$

where $r_i = \sum_j x_{ij}$ and $c_j = \sum_j x_{ij}$ are the row- and column sums as before.

Again *WE-matrices are generally not WS-matrices and vice versa.* No joint permutation of rows and columns take M_3 into M_4, where

$$M_3 = \begin{Bmatrix} 0 & 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 0 \end{Bmatrix}, \quad M_4 = \begin{Bmatrix} 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 & 0 \end{Bmatrix}$$

have identical row- and column sums, since $\det M_3 = 4$ whereas $\det M_4 = -4$ and simultaneous permutation of rows and columns also preserves the sign of the determinant.

If a matrix is both WE and WS, it is a *WES-matrix.* If in addition, $\{X_{ij} = X_{ji}\}$, i.e. the matrix is *symmetric*, we may consider SWE, SWS, SWES matrices, etc. Note that one could also consider the weaker *distributional symmetry* by assuming $X \stackrel{\mathcal{D}}{=} X^\top$, i.e. that transposition of X does not alter the distribution of X. In the following we shall mostly restrict attention to the fully symmetric case and write $X_{\{ij\}} = X_{ij} = X_{ji}$.

2.6 – de Finetti's theorem for SWE matrices

A symmetric binary doubly infinite random matrix X is a *ψ-matrix* if $X_{\{ij\}}$ are all independent given $U = (U_i)_{i=1,\dots}$ where U_i are mutually independent and uniform on $(0, 1)$ and

$$P(X_{\{ij\}} = 1 \mid U = u) = \psi(u_i, u_j).$$

Reformulating results in [11] yields that *binary SWE matrices are mixtures of ψ-matrices.* Exactly as in the case of RCES matrices, it is easy to show that

$$\mathcal{E}_{\mathrm{SWES}} = \mathcal{E}_{\mathrm{SWE}} \cap \mathcal{P}_{\mathrm{SWS}}.$$

implying that SWES matrices are mixtures of ψ-matrices where ψ satisfies the Rasch functional equation. The latter follows from the fact that we then must have for all $y, z \in \{0,1\}$ that

$$P\left(\left\{\begin{matrix} 0 & y & 0 & 1 \\ y & 0 & 1 & 0 \\ 0 & 1 & 0 & z \\ 1 & 0 & z & 0 \end{matrix}\right\} \,\Big|\, U = u\right) = P\left(\left\{\begin{matrix} 0 & y & 1 & 0 \\ y & 0 & 0 & 1 \\ 1 & 0 & 0 & z \\ 0 & 1 & z & 0 \end{matrix}\right\} \,\Big|\, U = u\right).$$

Hence *regular SWES ψ-matrices have the form*

$$\psi(u,v) = \frac{a(u)a(v)}{1 + a(u)a(v)}.$$

There are also non-regular solutions of interest in the symmetric case. For example

$$\psi(u,v) = \begin{cases} 0 & \text{if } u < 1/3 \text{ or } v < 1/3 \\ \dfrac{a(u)a(v)}{1 + a(u)a(v)} & \text{if } 1/3 < u, v < 2/3 \\ 1 & \text{otherwise.} \end{cases} \tag{3}$$

It seems complex to give a complete description of all symmetric solutions to the Rasch functional equation.

3 – Random graphs

3.1 – Exchangeable matrices as random graphs

The results described in the previous sections become particularly relevant when the binary matrix X is considered to represent a random graph. This representation can be made in a number of ways. If we consider the rows and colums as labels of two different sets of vertices, an undirected random bipartite graph can be defined from X by ignoring the diagonal and placing an edge between i and j if and only if $X_{ij} = 1$.

In this interpretation, an RCE-matrix corresponds to a random bipartite graph with exchangeable labels within each group of graph vertices. Similarly, an RCS-matrix is one where any two bipartite graphs with the same vertex degree for every vertex are equally likely. An RCES matrix represents one where the two distributions of vertex degrees determine the probability of the graph.

If we consider the row-and column numbers to label the *same vertex set*, the matrix X represents a random (directed) graph by placing a directed edge from i to j if and only if $X_{ij} = 1$. A WE-matrix then represents a random graph with exchangeable labels.

If the matrix X is symmetric it naturally represents a random undirected graph, again by placing an edge between i and j if and only if $X_{ij} = 1$. An SWE-matrix then represents an undirected random graph with exchangeable labels, an SWS-matrix represents a random graph with probability only depending on its vertex degrees, and an SWES matrix one with probability only depending on the distribution of vertex degrees.

Examples of WES and SWES graphs with non-regular ψ-matrices are displayed in fig. 4.

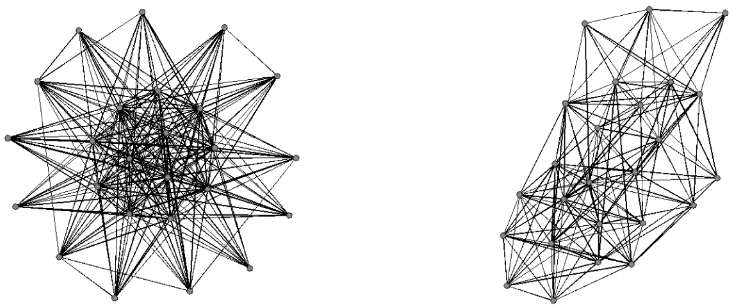


Fig. 4: The graph on the left-hand side is a non-regular SWES graph with ψ-matrix given by (3). The graph on the right-hand side has a ψ-matrix with $\phi(u,v) = \chi_{\{|u-v|\leq 1-1/\sqrt{2}\}}$. It is SWE but not SWES. Both graphs have 25 vertices.

3.2 – Social network analysis

Random graphs with exchangeability properties form natural models for *social networks* [18]. Frank and Strauss [19] consider *Markov graphs* which are random graphs with

$$X_{\{i,j\}} \perp\!\!\!\perp X_{\{k,l\}} \,|\, X_{E\setminus\{\{i,j\},\{k,l\}\}} \tag{4}$$

whenever all indices i, j, k, l are different. Here E denotes the edges in the complete graph on $\{1, \ldots, n\}$. They show that *weakly exchangeable Markov graphs on n vertices* all have the form

$$p(\{x_{ij}\}_{1,1}^{n,n}) \propto \exp\left\{\tau_n t(x) + \sum_{k=1}^{n-1} \eta_{nk}\nu_k(x)\right\}$$

where τ_n and η_{nk} are arbitrary real constants, $x = \{x_{ij}\}_{1,1}^{n,n}$, $t(x)$ is the number of triangles in x, and $\nu_k(x)$ is the number of vertices in x of degree k. Such

Markov graphs are also SWE, but not extendable as such unless $\tau_n = 0$ in which case they are also SWES. If $\tau_n \neq 0$ and $n > 5$, they are not SWES.

Typically, exchangeable graphs generated by ψ-matrices differ from Markov graphs in that they are *dissociated*, hence marginally rather than conditionally independent:

$$X_{\{i,j\}} \perp\!\!\!\perp X_{\{k,l\}} \tag{5}$$

whenever all indices i, j, k, l are different. In fact *infinite weakly exchangeable Markov graphs are Bernoulli graphs* because the conjunction of (4) and (5) implies complete independence.

It seems unfortunate that random induced subgraph of Markov graphs are not Markov themselves and it could be of interest to develop alternative models for social networks that preserve their structure when sampling subgraphs based on ψ-matrix models. This holds, for example, for the *latent space models* [20] and *latent position cluster models* [21], both of which are instances of ψ-matrix models.

For example, one could consider exchangeable random graphs which for every n also are summarized by the number of triangles and the empirical distribution of vertex degrees

$$p(\{x_{ij}\}_{1,1}^{n,n}) = f_n \left\{ t(x), \sum_{k=1}^{n} \delta_{r_k(x)} \right\}, \tag{6}$$

or similar graphs with summarizing statistics being counts of specific types of subgraph.

Characterizing exchangeable solutions to (6) or functions involving similar statistics in general use in social network analysis, could establish an interesting class of alternatives to the generalizations of Markov graphs known as exponential random graph models [22], [23], [24].

REFERENCES

[1] B. de Finetti: *Funzione caratteristica di un fenomeno aleatorio*, Atti della R. Academia Nazionale dei Lincei, Serie 6. Memorie, Classe di Scienze Fisiche, Mathematice e Naturale, **4** (1931), 251–299.

[2] S. L. Lauritzen: *Extreme point models in statistics (with discussion)*, Scandinavian Journal of Statistics, **11** (1984), 65–91.

[3] P. Diaconis – D. Freedman: *Partial exchangeability and sufficiency*, in "Statistics, Applications and New Directions; Proceedings of the Indian Statistical Institute Golden Jubilee International Conference", J. K. Ghosh and J. Roy, (eds), Sankhya A, Indian Statistical Institute, 1984, 205–236.

[4] S. L. Lauritzen: *Extremal Families and Systems of Sufficient Statistics*, Lecture Notes in Statistics, **49**, Springer-Verlag, Heidelberg, 1988.

[5] D. Freedman: *Invariants under mixing which generalize de Finetti's theorem*, Annals of Mathematical Statistics, **33** (1962), 916–923.

[6] S. L. Lauritzen: *Rasch models with exchangeable rows and columns*, in "Bayesian Statistics 7", J. M. Bernardo, M. J. Bayarri, J. O. Berger, A. P. Dawid, D. Heckerman, A. F. M. Smith, and M. West, (eds), Oxford, 2003, Clarendon Press, 215–232.

[7] G. Rasch: *Probabilistic Models for Some Intelligence and Attainment Tests*, Studies in Mathematical Psychology, **1**, Danmarks Pædagogiske Institut, Copenhagen, 1960.

[8] S. Gutmann – J. H. B. Kemperman – J. A. Reeds – L. A. Shepp: *Existence of probability measures with given marginals*, Annals of Probability, **19** (1991), 1781–1797.

[9] J. B. Wilson: *Methods for detecting non-randomness in species co-occurrences: A contribution*, Oecologica, **73** (1987), 579–582.

[10] G. H. Fischer – I. W. Molenaar: *Rasch Models. Foundations, Recent Developments, and Applications*, Springer-Verlag, New York, 1995.

[11] D. Aldous: *Representations for partially exchangeable random variables*, Journal of Multivariate Analysis, **11** (1981), 581–598.

[12] P. Diaconis – D. Freedman: *On the statistics of vision: the Julesz conjecture*, Journal of Mathematical Psychology, **24** (1981), 112–138.

[13] D. N. Hoover: *Row–column exchangeability and a general model for exchangeability*, in "Exchangeability in Probability and Statistics", G. Koch and F. Spizzichino (eds), North-Holland, Amsterdam, 1982, 281–291.

[14] O. Kallenberg: *On the representation theorem for exchangeable arrays*, Journal of Multivariate Analysis, **30** (1989), 137–154.

[15] G. Rasch: *Proof that the necessary condition for the validity of the multiplicative dichotomic model is also sufficient*, unpublished manuscript, 1971.

[16] B. Silverman: *Limit theorems for dissociated random variables*, Advances of Applied Probability, **8** (1976), 806–819.

[17] G. K. Eagleson – N. C. Weber: *Limit theorems for weakly exchangeable random variables*, Proceedings of the Cambridge Philosophical Society, **84** (1979), 73–80.

[18] S. Wasserman – K. Faust: *Social Network Analysis*, Cambridge University Press, Cambridge, UK, 1995.

[19] O. Frank – D. Strauss: *Markov graphs*, Journal of the American Statistical Association, **81** (1986), 832–842.

[20] P. D. Hoff – A. E. Raftery – M. S. Handcock: *Latent space approaches to social network analysis*, Journal of the American Statistical Association, **97** (2002), 1090–1098.

[21] M. S. Handcock – A. E. Raftery – J. M. Tantrum: *Model-based clustering for social networks (with discussion)*, Journal of the Royal Statistical Society, Series A, **170** (2007), 301–354.

[22] O. Frank: *Statistical analysis of change in networks*, Statistica Neerlandica, **45** (1991), 283–293.

[23] S. Wasserman – P. E. Pattison: *Logit models and logistic regression models for social networks I: An introduction to Markov graphs and* p^*, Psychometrika, **61** (1996), 401–425.

[24] T. A. B. Snijders – P. E. Pattison – G. L. Robins – M. S. Handcock: *New specifications for exponential random graph models*, Sociological Methodology, **36** (2006), 99–153.

Lavoro pervenuto alla redazione il 12 ottobre 2007
ed accettato per la pubblicazione il 18 ottobre 2007.
Bozze licenziate il 9 aprile 2008

Indirizzo dell'Autore:

Steffen L. Lauritzen – Department of Statistics – University of Oxford – 1 South Parks Road – Oxford OX1 3TG, United Kingdom
E-mail: steffen@stats.ox.ac.uk.

Rendiconti di Matematica, Serie VII
Volume 28, Roma (2008), 97–115

The dependence structure of log-fractional stable noise with analogy to fractional Gaussian noise

MURAD S. TAQQU – JOSHUA B. LEVY

Abstract: *We examine the process log-fractional stable motion (log-FSM), which is an* α*-stable process with* $\alpha \in (1,2)$*. Its tail probabilities decay like* $x^{-\alpha}$ *as* $x \to \infty$*, and hence it has a finite mean, but its variance is infinite. As a result, its dependence structure cannot be described by using correlations. Its increments, log-fractional noise (log-FSN), are stationary and so the dependence between any two points in time can be determined by a function of only the distance (lag) between them. Since log-FSN is a moving average and hence "mixing," the dependence between the two time points decreases to zero as the lag tends to infinity. Using measures such as the codifference and the covariation, which can replace the covariance when the variance is infinite, we show that the decay is so slow that log-FSN (or, conventionally, log-FSM) displays long-range dependence. This is compared to the asymptotic dependence structure of fractional Gaussian noise (FGN), a befitting circumstance since log-FSN and FGN share a number of features.*

1 – Introduction

The classical Central Limit Theorem deals with the convergence of normalized sums of independent and identically distributed random variables, and states that if these random variables have finite variance then the limit is Gaussian. The cases of infinite variance and triangular arrays are more involved. The limits are then infinitely divisible. Bruno de Finetti was one of the first to consider infinitely divisible distributions (see [3] as well as [2] and [10]). Since then the

Key Words and Phrases: *α-stable distribution – Codifference – Covariation – Fractional Brownian motion – Log-fractional stable noise – Self-similar process.*
A.M.S. Classification: 60G15, 60G52, 60E07, 60G18

subject has developed in many directions. One of them concerns dependence. The dependence of infinitely divisible random variables is most conveniently described when they have an α-stable distribution, because linear combinations of α-stable random variables remain α-stable. Sequences and random processes with α-stable distributions then can be readily defined and their dependence structure investigated.

An α-stable process with index of stability $0 < \alpha < 2$ has tail probabilities that decrease to zero hypergeometrically, that is, like the power function $x^{-\alpha}$, as $x \to \infty$. These are the proverbial "heavy" tails since the rate of decrease can be very slow. Moments of the process that have order $p < \alpha$ are necessarily finite, but they are infinite if $p \geq \alpha$. An important case is when $\alpha \in (1, 2)$, so that the mean is finite but the variance is infinite. This contrasts markedly with the more familiar Gaussian process (by convention, the case $\alpha = 2$), which has exponentially "light" tails of order $c_1 x^{-1} e^{-c_2 x^2} (c_i > 0)$, and hence has all moments finite. Unlike the Gaussian distribution, which is symmetric about its mean, a non-Gaussian stable distribution can be also skewed either to the left or to the right of its mean. We will concentrate, though, on *symmetric* α-stable processes for which the distribution is symmetric around the origin. Processes that are α-stable ($0 < \alpha < 2$) can be used to model high variability, namely, phenomena exhibiting "acute spikes" and "eruptions," a behavior that is also often described as *burstiness*.

A random process is *self-similar* if it has finite-dimensional distributions that scale. Specifically, $\{X_t\}$, $t \in \mathbb{R}$, is H-self similar (H-ss), $H > 0$, if

$$X_{ct} \stackrel{d}{=} c^H X_t$$

for any $c > 0$ and $t \in \mathbb{R}$. The notation $\stackrel{d}{=}$ signifies equality of the finite-dimensional distributions, that is, for any finite set of times $t_1, \dots, t_n$

$$\mathbb{P}(X_{ct_1} \leq x_1, \dots, X_{ct_n} \leq x_n) = \mathbb{P}(c^H X_{t_1} \leq x_1, \dots, c^H X_{t_n} \leq x_n).$$

H is called the self-similarity index for $\{X_t\}$. Thus, the finite-dimensional distributions maintain an invariance through a simple scaling of time and space. (Refer to the excellent monograph by Embrechts and Maejima [5] and to the review paper [13] for details.) The process $\{X_t\}$ has stationary increments (si) if $X_{t+s} - X_s \stackrel{d}{=} X_t - X_0$ for all $t, s \in \mathbb{R}$. Processes that are both H-self-similar and have stationary increments (indicated by H-sssi) are helpful for describing natural events that display *long-range dependence*. Long-range dependence occurs, for example, in economic time series and internet communication. Processes that are both α-stable with $\alpha < 2$ and H-sssi are effective models with which to investigate both burstiness and long-range dependence in (but are not limited to) network traffic, hydrology, and financial data. Besides articles in the literature about α-stable, H-sssi, or α-stable H-sssi processes and their applications, see also the texts [4], [11], and [6].

In the case of Gaussian or any finite variance process, the dependence structure in a weaker form can be studied readily through the correlations. For example, zero correlation between the components of a Gaussian process is equivalent to their independence. If the components are stationary, then one can examine their dependence over time durations, or lags. If as the lags get larger, the correlations converge rapidly to zero, then the dependence is "weak." On the other hand, the dependence is "strong" if the convergence is so slow that the sum of the correlations diverge. Such divergence intrinsically characterizes the random cycles of abnormality and regularity exhibited by long-range dependence.

This paper focuses on the *symmetric* α-stable ($S\alpha S$) H-sssi process *log-fractional stable motion* (log-FSM), which is defined for $1 < \alpha < 2$ and has the self-similarity index $H = 1/\alpha$. Log-FSM has zero mean but infinite variance. In particular, its dependence cannot be measured by correlations. There are, however, "stable" alternatives that replace the covariance.

Two of them are the *codifference* and the *covariation*. Both can be applied to the stationary increments of log-FSM, which is the process known as *log-fractional stable noise*, log-FSN. ("Motion" refers to a process with stationary increments and "noise" to a stationary process.) The behavior of these dependence measures for log-FSN in turn gives an indication about the dependence structure for log-FSM. The codifference, in fact, is defined for any stationary process. The covariation is restricted to $S\alpha S$ processes, albeit not necessarily stationary, for which $1 < \alpha < 2$.

The rest of the paper is carried out as follows. Section 2 briefly reviews $S\alpha S$ processes and their representation as integrals with respect to $S\alpha S$ random measures. Log-fractional stable motion and its increment process log-fractional stable noise (log-FSN) are reviewed in Section 3. The measures of dependence, the codifference and the covariation, are presented in Section 4. Section 5 contains the main results, namely, the asymptotic behavior of the measures when applied to log-FSN. Section 6 makes an analogy to fractional Brownian motion and its increment process fractional Gaussian noise. Some extensions of this work and potential research are mentioned in the concluding Section 7.

2 – A brief approach to symmetric α-stable laws and processes

Aside from the applications described in the introduction, one may ask: why consider stable distributions? The usual answer arises from the central limit theorem, which obtains that they are the unique limits of properly rescaled sums of independent and identically distributed (i.i.d.) random variables. The Gaussian (normal) distribution is the limit if the sequence has a finite variance. If the variance is infinite but the tails of the sequence demonstrate hypergeometric decay for $\alpha < 2$, then the limit turns out to be stable and has the same index α.

More explicitly, let $\{X_j\}_{j=1}^{\infty}$ be an i.i.d. sequence. In the case $\alpha = 2$, suppose X_j has mean μ and variance σ_0^2. Then

$$\frac{1}{n^{1/2}} \sum_{j=1}^{n} (X_j - \mu) \xrightarrow{L} Z \sim N(0, \sigma_0^2)$$

where $\xrightarrow{L}$ stands for convergence in law, that is, in distribution. The limit is a random variable having characteristic function $\mathbb{E}e^{i\theta Z} = e^{-\frac{1}{2}\sigma_0^2|\theta|^2}$ and, consequently, must be Gaussian. In particular, $\mathbb{E}|Z|^p < \infty$ for all $p > 0$. By contrast, if $0 < \alpha < 2$, then the tail probabilities of X_j are "heavy": $\mathbb{P}(|X_j| \geq x) \sim cx^{-\alpha}$ as $x \to \infty$ with $\sigma_0^2 = \infty$. Assume, in addition, X_j is symmetric ($X_j \stackrel{L}{=} -X_j$) if $\alpha = 1$ and has mean μ if $1 < \alpha < 2$. Then

$$\frac{1}{n^{1/\alpha}} \sum_{j=1}^{n} X_j \xrightarrow{L} Z_\alpha$$

if $\alpha \leq 1$ and

$$\frac{1}{n^{1/\alpha}} \sum_{j=1}^{n} (X_j - \mu) \xrightarrow{L} Z_\alpha.$$

if $1 < \alpha < 2$. The limit is a symmetric non-Gaussian α-stable random variable Z_α. We indicate this by writing $Z_\alpha \sim S\alpha S$. The limit Z_α recovers the tail behavior of X_j since also $\mathbb{P}(|Z_\alpha| \geq x) \sim cx^{-\alpha}$ as $x \to \infty$, perhaps with a different c. Thus, $\mathbb{E}|Z_\alpha|^p < \infty$ if and only if $p < \alpha$; $\mathbb{E}|Z_\alpha|^2 = \infty$, $\mathbb{E}Z_\alpha = 0$ for $1 < \alpha < 2$, and $\mathbb{E}|Z_\alpha| = \infty$ iff $\alpha \leq 1$. Its characteristic function satisfies

$$\mathbb{E}e^{i\theta Z_\alpha} = e^{-\sigma^\alpha|\theta|^\alpha}, \quad \theta \in \mathbb{R} := (-\infty, \infty), \tag{2.1}$$

where the scale parameter σ depends on α and c. (When $\alpha = 2, \sigma = \sqrt{\sigma_0^2/2}$.)

Relation (2.1) identifies the specific random variable arising in the stable central limit theorem. Any random variable X is, by definition, symmetric α-stable ($S\alpha S$) if it satisfies (2.1). Its scale parameter σ is denoted by $\|X\|_\alpha$. If for instance X is measured in meters, then so is $\|X\|_\alpha$.

Remark. Several easy facts about $X \sim S\alpha S$ ($0 < \alpha \leq 2$) are worth noting (see also [12, ch. 1.2]).

- $a \in \mathbb{R}, a \neq 0$ implies aX is $S\alpha S$ with $\|aX\|_\alpha = |a|\|X\|_\alpha$.
- If $\alpha = 2$ then $X \sim N(0, 2\sigma^2)$.
- $\mathbb{E}|X|^p < \infty$ only for $p < \alpha$.

- Any linear combination of independent $S\alpha S$ random variables is $S\alpha S$: if $\epsilon_j \sim S\alpha S$ are independent and $a_j \in \mathbb{R}$, $1 \le j \le n$, then $X = \sum_{j=1}^n a_j\epsilon_j \sim S\alpha S$ with
$$\|X\|_\alpha^\alpha = \|\sum_{j=1}^n a_j\epsilon_j\|_\alpha^\alpha = \sum_{j=1}^n |a_j|^\alpha \|\epsilon_j\|_\alpha^\alpha.$$
It is instructive to see why this last relation holds. The characteristic function of X is, for $\theta \in \mathbb{R}$,
$$\begin{aligned}\phi_X(\theta) = \mathbb{E}e^{i\theta X} &= \mathbb{E}\exp\left\{i\theta\sum_{j=1}^n a_j\epsilon_j\right\}\\ &= \prod_{j=1}^n \mathbb{E}\exp\{i\theta a_j\epsilon_j\} = \prod_{j=1}^n e^{-|\theta|^\alpha\,|a_j|^\alpha\,\|\epsilon_j\|_\alpha^\alpha}\\ &= \exp\left\{-|\theta|^\alpha\sum_{j=1}^n |a_j|^\alpha\,\|\epsilon_j\|_\alpha^\alpha\right\} = \mathbb{E}e^{i\theta X},\end{aligned}$$
on using the independence of the ϵ_j and the fact that they are $S\alpha S$.

The vector $\mathbf{X} = (X_1, \dots, X_d)$ in $\mathbb{R}^d$ is Gaussian if and only if the random variables $\{X_1 \dots, X_d\}$ are *jointly* Gaussian, that is, any linear combination of them is Gaussian. Similarly, for $0 < \alpha < 2$, $\mathbf{X} = (X_1, \dots, X_d)$ in $\mathbb{R}^d$ is a $S\alpha S$ vector if and only if $\{X_1 \dots, X_d\}$ are jointly $S\alpha S$, that is, if linear combinations $\sum_{j=1}^n a_jX_j$ are $S\alpha S$ random variables.

By "going to the limit" in the sum $\sum_{j=1}^n a_j\epsilon_j$ one can define a $S\alpha S$ random variable as an integral,

$$X = \int_{\mathbb{R}} f(x)M_\alpha(\mathrm{d}x), \tag{2.2}$$

where f is a deterministic function and M_α is a symmetric α-stable *random measure* (see [12, ch. 3.3]). The scale parameter for X satisfies

$$\|X\|_\alpha^\alpha = \int_{\mathbb{R}} |f(x)|^\alpha \mathrm{d}x < \infty$$

with $\mathrm{d}x$ denoting the Lebesgue measure on $\mathbb{R}$. Formally, the function $f(x)$ plays the role of the a_j's and $M_\alpha(\mathrm{d}x)$ plays the role of the ϵ_j's with $\|M_\alpha(\mathrm{d}x)\|_\alpha^\alpha = \mathrm{d}x$. M_α is defined on $(\mathbb{R}, \mathcal{B}, |\cdot|)$, where $\mathcal{B}$ is the Borel σ-algebra on $\mathbb{R}$ and $|\cdot|$ is Lebesgue measure. Here $|\cdot|$ is the *control* measure and $(\mathbb{R}, \mathcal{B}, |\cdot|)$ is called the *control* space for M_α. This means that if $B \in \mathcal{B}$ with finite Lebesgue measure $|B|$, then $M_\alpha(B)$ is a $S\alpha S$ random variable for which

$$\mathbb{E}e^{i\theta M_\alpha(B)} = e^{-|\theta|^\alpha|B|}.$$

Furthermore, suppose $\{B_n\}_{n=1}^{\infty}$ is a pairwise disjoint sequence of sets in $\mathcal{B}$ with $|B_n| < \infty$. Then any finite subcollection $\{M_\alpha(B_n)\}_{n=1}^{\infty}$ are independent random variables (M_α is said to be *independently scattered*), and if $|\bigcup_{n=1}^{\infty} B_n| < \infty$, $M_\alpha(\bigcup_{n=1}^{\infty} B_n) = \sum_{n=1}^{\infty} M_\alpha(B_n)$ almost surely (a.s.) (M_α is σ-*additive*).

Thus, M_α plays the dual role of being a measure and being random. As suggested above, the $M_\alpha(\mathrm{d}x), x \in \mathbb{R}$ play the role of i.i.d. infinitesimal ϵ_j, with the continuous x replacing the discrete label j, and the infinitesimal $\mathrm{d}x$ replacing the common value $\|\epsilon_j\|_\alpha^\alpha$, namely, the scale parameter raised to the power α.

In the Gaussian case $\alpha = 2$, one usually takes $(\mathbb{R}, \mathcal{B}, |\cdot|/2)$ as the control space, and in this case, $M_2(B)$ has characteristic function $\mathbb{E}e^{i\theta M_2(B)}$; hence, $M_2(B)$ is a normal random variable with mean 0 and variance $|B|$. One can view $M_2(\mathrm{d}x)$ heuristically as a normal random variable having mean zero and infinitesimal measure $\mathrm{d}x$, with $M_2(\mathrm{d}x)$ and $M_2(\mathrm{d}x')$ being independent if the infinitesimal intervals $\mathrm{d}x$ and $\mathrm{d}x'$ are disjoint. The same intuition prevails in the $S\alpha S$ case with $\alpha < 2$. The normal distribution is replaced by the stable distribution and the variance is replaced by the scale parameter raised to the power α.

Suppose $B \in \mathcal{B}$ and $a > 0$. Then

$$M_\alpha(aB) \stackrel{d}{=} a^{1/\alpha} M_\alpha(B), \tag{2.3}$$

where aB is the set B scaled by a, and $\stackrel{d}{=}$ means equality of the finite-dimensional distributions. Indeed,

$$\begin{aligned}\mathbb{E}e^{i\theta M_\alpha(aB)} &= \exp\{-|aB|\,|\theta|^\alpha|\}\\ &= \exp\left\{-|B|^\alpha \left|a^{1/\alpha}\theta\right|^\alpha\right\}\\ &= \mathbb{E}e^{i\theta a^{1/\alpha} M_\alpha(B)}.\end{aligned}$$

Relation (2.3) can be denoted informally by

$$M_\alpha(a\mathrm{d}x) \stackrel{d}{=} a^{1/\alpha} M_\alpha(\mathrm{d}x).$$

Thus, M_α can be regarded as being "self-similar" with "index" $H = 1/\alpha$.

The definition (2.2) can be extended to a random process. Consider the set T to be either $\mathbb{R}, \mathbb{R}_+ = \{t : t \geq 0\}$, or $\{t : t > 0\}$. Let $f_t : \mathbb{R} \to \mathbb{R}$ be measurable and satisfy for each $t \in T$

$$\int_{\mathbb{R}} |f_t(x)|^\alpha \mathrm{d}x < \infty \tag{2.4}$$

and also, if $\alpha = 1$, $\int_{\mathbb{R}} |\,|f_t(x) \ln |f_t(x)|\,|\,\mathrm{d}x < \infty$. (In fact, the condition (2.4) alone suffices to ensure the existence of the subsequent random process in the $S\alpha S$ case when $\alpha = 1$.) Then $\{X_t, t \in T\}$ defined by

$$X_t = \int_{\mathbb{R}} f_t(x) M_\alpha(\mathrm{d}x) \tag{2.5}$$

is a $S\alpha S$ *process* with

$$\|X_t\|_\alpha^\alpha = \int_{\mathbb{R}} |f_t(x)|^\alpha \mathrm{d}x. \tag{2.6}$$

The integral (2.5) is a "representation" of the process $\{X_t\}$. It says, intuitively, that $\{X_t\}$ is obtained by starting with i.i.d. infinitesimal random variables $M_\alpha(\mathrm{d}x)$, weighting them by $f_t(x)$, and integrating. The weights change in general with x and can also change as the time t evolves. Let $-\infty < t_1 \le \cdots \le t_d < \infty$. The joint characteristic function of a typical vector $(X_{t_1}, \ldots, X_{t_d})$ of the process is given by

$$\mathbb{E}\exp\left\{i\sum_{j=1}^d \theta_j X_{t_j}\right\} = \exp\left\{-\int_{\mathbb{R}}\left|\sum_{j=1}^d \theta_j f_{t_j}(x)\right|^\alpha \mathrm{d}x\right\}$$

for arbitrary $\theta_1, \ldots, \theta_d \in \mathbb{R}$. In fact, most $S\alpha S$ processes can be represented in the form (2.5). (For details refer to [12, ch. 3 and ch. 13.2].)

One can also define an integrated process with respect to an asymmetric α-stable random measure having arbitrary control measure that is asymmetric, or *skewed*. If the integrand is as above, then the resulting process has also asymmetric distributions. Our concern in this paper only involves processes that are defined by (2.5) based on $S\alpha S$ random measures having Lebesgue control space.

Now recall the definitions of self-similarity and stationarity of the increments given in the introduction. A process $\{X_t, t \in T\}$ is *H-self-similar* (H-ss) with $H > 0$ if

$$X_{ct} \overset{d}{=} c^H X_t \tag{2.7}$$

for all $c > 0$ and t, that is, $(X_{ct_1}, \ldots, X_{ct_d})$ and $c^H(X_{t_1}, \ldots, X_{t_d})$ are identically distributed. Note that $c^H X_0 \overset{d}{=} X_{c0} = X_0$, hence, letting $c \to \infty$ necessitates $X_0 = 0$ *a.s.* A process is said to have *stationary increments* if the finite-dimensional distributions of $\{X_{t+s} - X_s\}$ do not depend on s:

$$\{X_{s+t} - X_s, t \in T\} \overset{d}{=} \{X_t - X_0, t \in T\} \quad \text{for all} \quad s \in T. \tag{2.8}$$

Suppose now that the process $\{X_t, t \in \mathbb{R}\}$

- is H self-similar,
- has stationary increments, and
- is symmetric α-stable;

to wit, it is H-sssi and $S\alpha S$. Let $c > 0$ and $s, \theta_1, \dots, \theta_d, -\infty < t_1 \le \cdots \le t_d < \infty \in \mathbb{R}$. It follows that $\| \sum_{j=1}^d \theta_j (X_{ct_j+s} - X_s)\|_\alpha^\alpha = c^{\alpha H} \| \sum_{j=1}^d \theta_j X_{t_j} \|_\alpha^\alpha$ does not depend on s, since by (2.5)-(2.8),

$$\int_{\mathbb{R}} \left| \sum_{j=1}^d \theta_j \left(f_{ct_j+s}(x) - f_s(x) \right) \right|^\alpha \mathrm{d}x = c^{\alpha H} \int_{\mathbb{R}} \left| \sum_{j=1}^d \theta_j f_{t_j}(x) \right|^\alpha \mathrm{d}x. \tag{2.9}$$

If $\alpha = 2$, M_2 is a Gaussian random measure. Remember that in this case the control space usually is taken to be $(\mathbb{R}, \mathcal{B}, |\cdot|/2)$, so that the variance of M_2 is $\mathbb{E}M^2(B) = |B|$. The process defined by

$$B_t = \int_{\mathbb{R}} 1_{[0,t]}(x) M_2(\mathrm{d}x) = \int_0^t M_2(\mathrm{d}x), \qquad t \ge 0.$$

is Brownian motion. (One can also define for $t < 0$, $B_t = \int_{-t}^0 M_2(\mathrm{d}x)$.) It is the *only* Gaussian H-sssi process with

$$H = 1/2.$$

Its scale parameter is $\mathbb{E}B_t^2 = \|B_t\|_2^2 = t$ by (2.6) (and $\mathbb{E}B_t = 0$). This is actually *standard Brownian motion* since $\mathbb{E}B_1^2 = 1$. Its covariance $\mathrm{Cov}\,(B_{t_1}, B_{t_2})$ satisfies

$$\mathrm{Cov}\,(B_{t_1}, B_{t_2}) = \mathbb{E}B_{t_1}B_{t_2} = \int_{\mathbb{R}} 1_{[0,t_1]}(x) 1_{[0,t_2]}(x)\mathrm{d}x = \min(t_1, t_2). \tag{2.10}$$

Moreover, the increments over disjoint intervals are (mutually) independent.

What happens if “Gaussian” in Brownian motion is replaced by “$S\alpha S$, $0 < \alpha < 2$”?

Replacing the Gaussian random measure M_2 with the $S\alpha S$ random measure M_α, we obtain the *stable Lévy motion*:

$$L_t = \int_0^t M_\alpha(\mathrm{d}x), \qquad t \ge 0.$$

Also called α-stable motion, it is a $S\alpha S$ process with $\|L_t\|_\alpha^\alpha = t$. Its increments over disjoint intervals are independent, a feature that distinguishes it from other $S\alpha S$ processes. Moreover, it is H-sssi with

$$H = 1/\alpha.$$

One can verify heuristically the self-similarity: for $a > 0$,

$$L_{at} = \int_0^{at} M_\alpha(\mathrm{d}x) \stackrel{d}{=} \int_0^t M_\alpha(a\mathrm{d}x) \stackrel{d}{=} a^{1/\alpha} \int_0^t M_\alpha(\mathrm{d}x) = a^{1/\alpha} L_t.$$

(This can be checked precisely using characteristic functions.) A striking fact is that when $0 < \alpha < 1$, there is no other nondegenerate $S\alpha S$ $1/\alpha$-sssi process besides $\{L_t, t \ge 0\}$.

Proposition 2.1. *For $0 < \alpha < 1$ α-stable motion is the only nondegenerate $S\alpha S$-stable $1/\alpha$-sssi process.*

Proof. We will follow the proof of [12, Theorem 7.5.4, p. 351], citing several referenced results from that monograph. That theorem is stated more generally for *arbitrary* α-stable $1/\alpha$-sssi processes, not necessarily symmetric.

Let $\{X_t, t \geq 0\}$ be a nondegenerate $S\alpha S$ $1/\alpha$-sssi process for fixed $\alpha, 0 < \alpha < 1$. In particular, X_1 is non-constant almost surely (a.s.). Denote by σ_t the scale parameter of X_t, namely, $\sigma_t = \|X_t\|_\alpha$. If $s < t$, then

$$\|X_t - X_s\|_\alpha^\alpha = \sigma_{t-s}^\alpha = (t-s)\sigma_1^\alpha, \tag{2.11}$$

since $X_t - X_s \stackrel{d}{=} X_{t-s} \stackrel{d}{=} (t-s)^{1/\alpha} X_1$ by stationarity and $1/\alpha$-self-similarity. Observe first that $\sigma_1 = \|X_1\|_\alpha \neq 0$. Indeed, if $\sigma_1 = 0$ then $\{X_t\}, t \geq 0$ would be degenerate since by (2.11),

$$\sigma_t = \|X_t - X_0\|_\alpha = t^{1/\alpha}\|X_1 - X_0\|_\alpha = t^{1/\alpha}\sigma_1 = 0.$$

We must prove X_t has independent increments, that is, for any $d \geq 3$ and $0 < t_1 \leq \cdots \leq t_d$ the random variables $\{X_{t_j} - X_{t_{j-1}}\}, 2 \leq j \leq d$ are (mutually) independent.

Consider arbitrary epochs $t_1 < t_2 \leq t_3 < t_4$. Since the vector $(X_{t_1}, X_{t_2}, X_{t_3}, X_{t_4})$ is jointly $S\alpha S$, then there exist a $S\alpha S$ random measure M_α with Lebesgue control space $([0,1], \mathcal{B}, |\cdot|)$ and functions $\{f_{t_j}(x)\}, x \in [0,1]$, satisfying $\int_0^1 |f_{t_j}(x)|^\alpha \mathrm{d}(x) < \infty$, $j = 1,2,3,4$, such that

$$X_t = \int_0^1 f_t(x) M_\alpha(\mathrm{d}x)$$

for each $t = t_1, t_2, t_3, t_4$ (Theorem 3.5.6, pp. 131–132). We are now going to verify that the pair of increments $X_{t_2} - X_{t_1}$ and $X_{t_4} - X_{t_3}$ are independent, by showing $f_{t_2} - f_{t_1}$ and $f_{t_4} - f_{t_3}$ have almost-$[\mathrm{d}x]$ disjoint supports, i.e.

$$\big(f_{t_2}(x) - f_{t_1}(x)\big)\big(f_{t_4}(x) - f_{t_3}(x)\big) = 0 \quad \text{a.e.} [\mathrm{d}x]. \tag{2.12}$$

Using the inequality $|a+b|^\alpha \leq |a|^\alpha + |b|^\alpha$, valid for $0 < \alpha \leq 1$,

$$\begin{aligned}
(t_4 - t_1)\sigma_1^\alpha = \sigma_{t_4-t_1}^\alpha &= \int_0^1 |f_{t_4}(x) - f_{t_1}(x)|^\alpha \mathrm{d}x \qquad (\text{by}(2.11)) \\
&\leq \int_0^1 |f_{t_4}(x) - f_{t_3}(x)|^\alpha \mathrm{d}x + \int_0^1 |f_{t_3}(x) - f_{t_2}(x)|^\alpha \mathrm{d}x + \\
&\quad + \int_0^1 |f_{t_2}(x) - f_{t_1}(x)|^\alpha \mathrm{d}x \\
&= \sigma_{t_4-t_3}^\alpha + \sigma_{t_3-t_2}^\alpha + \sigma_{t_2-t_1}^\alpha = \\
&= \big(t_4 - t_3\big)\sigma_1^\alpha + \big(t_3 - t_2\big)\sigma_1^\alpha + \big(t_3 - t_2\big)\sigma_1^\alpha = (t_4 - t_1)\sigma_1^\alpha.
\end{aligned}$$

The preceding inequality is therefore an equality. Applying Lemma 2.7.14 (1), p. 92, we can conclude (2.12) holds. This proves that $X_{t_2} - X_{t_1}$ and $X_{t_4} - X_{t_3}$ are independent (Theorem 3.5.3, p. 128).

Since jointly α-stable random variables are independent if and only if they are pairwise independent (Corollary 3.5.4, p. 129), then $\{X_{t_j} - X_{t_{j-1}}\}, 2 \le j \le d, d \ge 3$ are independent. Thus, the increments of $\{X(t)\}$ are independent, which establishes in turn that $\{X(t)\}$ must be α-stable motion. ∎

When $1 \le \alpha < 2$, there are other $S\alpha S$ processes besides α-stable motion that are H-sssi with $H = 1/\alpha$. In the sequel we will concentrate on the *log-fractional stable motion*.

3 – Log-fractional stable motion

DEFINITION 3.1. The process defined by

$$X_t = \int_{\mathbb{R}} (\ln|t-x| - \ln|x|)\, M_\alpha(\mathrm{d}x), \quad t \in \mathbb{R}, \tag{3.1}$$

where for $1 < \alpha < 2$, M_α is a $S\alpha S$ random measure having Lebesgue control measure, is called log-fractional stable motion (log-FSM).

Log-FSM was introduced by Kasahara et al. [7]. It is well-defined only for $1 < \alpha \le 2$. Indeed, $\int_{-\infty}^{\infty} |\ln|t-x| - \ln|x||^\alpha \,\mathrm{d}x$ is finite since, when $x \sim 0, \int_0^\delta (\ln|x|)^\alpha \,\mathrm{d}x < \infty$ but if $|x| \sim \infty$, then $\ln|t-x| - \ln|x| \sim -t/x$, and for $A > 0$, $\int_A^\infty x^{-\alpha} \mathrm{d}x < \infty$ if and only if $\alpha > 1$. (See also [12, ch. 7.6] for additional information.)

PROPOSITION 3.1. *Log-FSM is H-sssi with $H = 1/\alpha$.*

PROOF. We will show that (2.9) holds. Let $c > 0$ and $s, \theta_1, \dots, \theta_d, -\infty < t_1 \le \cdots \le t_d < \infty \in \mathbb{R}$. The change of variables $x \mapsto (s-x)/c$ gets

$$\begin{aligned}
&\int_{-\infty}^{\infty} \left| \sum_{j=1}^{d} \theta_j \left(\ln|ct_j + s - x| - \ln|s-x|\right) \right|^\alpha \mathrm{d}x = \\
&= \int_{-\infty}^{\infty} \left| \sum_{j=1}^{d} \theta_j \left(\ln|c(t_j - x)| - \ln|cx|\right) \right|^\alpha c\,\mathrm{d}x \\
&= c \int_{-\infty}^{\infty} \left| \sum_{j=1}^{d} \theta_j \left(\ln|t_j - x)| - \ln|x|\right) \right|^\alpha \mathrm{d}x.
\end{aligned}$$

This verifies (2.9) with $\alpha H = 1$. ∎

What happens when we consider log-FSM with $\alpha = 2$? It becomes Gaussian. Since it is also H-sssi with $H = 1/2$, is it different from Brownian motion? The answer is "no." To see that log-FSM and Brownian motion are the *same* Gaussian process, it suffices to observe that they have identical variance-covariance structures. Indeed, by self-similarity, $H = 1/2$ implies $\mathbb{E}X_t^2 = t\mathbb{E}X_1^2$, and this leads to

$$\begin{aligned}\mathbb{E}X_{t_1}X_{t_2} &= \frac{1}{2}\left(\mathbb{E}X_{t_1}^2 + \mathbb{E}X_{t_1}^2 - \mathbb{E}\,|X_{t_1} - X_{t_2}|^2\right) = \\ &= \frac{1}{2}\left(t_1 + t_2 - |t_1 - t_2|\right)\mathbb{E}X_1^2 = \min(t_1, t_2)\mathbb{E}X_1^2,\end{aligned}$$

which is the covariance of Brownian motion (compare it to (2.10)). Thus, when $\alpha = 2$, (3.1) is merely a different representation of Brownian motion.

What about the case $1 < \alpha < 2$? Is log-FSM the same process as α-stable motion? Observe that they are both H-sssi with $H = 1/\alpha$. However,

Proposition 3.2. *When $1 < \alpha < 2$, log-FSM and α-stable motion are different processes.*

We have verified in Proposition 2.1 that α-stable motion has independent increments. We will show momentarily that log-FSM has *dependent* increments. To do so, we consider the increment process of log-FSM called *log-fractional stable noise.*

Definition 3.2. Let $1 < \alpha \le 2$. Log-fractional stable noise (Log-FSN) is the $S\alpha S$ process,

$$Y_t : = X_{t+1} - X_t = \int_{\mathbb{R}} \left(\ln|t+1-x| - \ln|t-x|\right) M(\mathrm{d}x) \qquad t \in \mathbb{R}. \tag{3.2}$$

It is the increment process of log-FSM, $\{X_t, t \in \mathbb{R}\}$.

Do not confuse "log-FSM" with "log-FSN." The first, with "M" standing for motion, refers to the process with stationary increments. The second with "N" standing for noise refers to the corresponding stationary process obtained by taking the increments of log-FSM.

We proceed to prove Proposition 3.2.

Proof. Two α-stable variables, $0 < \alpha < 2$, $\int_{\mathbb{R}} f(x)M_\alpha(\mathrm{d}x)$ and $\int_{\mathbb{R}} g(x)$ $M_\alpha(\mathrm{d}x)$ are independent if and only if their kernels f and g have disjoint support, a.e. $[\mathrm{d}x]$ [12, Theorem 3.5.3, p. 128]. For any $t \in \mathbb{R}$, the support of Y_t in (3.2) is evidently $\mathbb{R}$. Therefore, Y_{t_1} and Y_{t_2} can never be independent for any $t_1 \neq t_2$. $\square$

Having established Proposition 3.2, our goal is to analyze the dependence of the increments using the codifference and the covariation.

4 – Two measures of dependence

Suppose X_1 and X_2 are jointly $S\alpha S$. In particular, $X_1 - X_2$ is $S\alpha S$. The *codifference* between two jointly $S\alpha S$ random variables is defined by

$$\tau_{X_1,X_2} = \|X_1\|_\alpha^\alpha + \|X_2\|_\alpha^\alpha - \|X_1 - X_2\|_\alpha^\alpha. \tag{4.1}$$

The codifference arises from comparing the joint characteristic function of (X_1, X_2) to the product of their marginal characteristic functions:

$$U_{X_1,X_2}(\theta_1,\theta_2) = \mathbb{E}e^{i(\theta_1 X_1+\theta_2 X_2)} - \mathbb{E}e^{i\theta_1 X_1}\mathbb{E}e^{i\theta_2 X_2},$$

whereupon setting $\theta_1 = 1, \theta_2 = -1$, one gets

$$\begin{aligned} U_{X_1,X_2}(1,-1) &= \mathbb{E}e^{i(X_1-X_2)} - \mathbb{E}e^{iX_1}\mathbb{E}e^{-iX_2} \\ &= e^{-\|X_1-X_2\|_\alpha^\alpha} - e^{-\|X_1\|_\alpha^\alpha - \|X_2\|_\alpha^\alpha} \\ &= e^{-\|X_1\|_\alpha^\alpha - \|X_2\|_\alpha^\alpha}\left(e^{\tau_{X_1,X_2}} - 1\right). \end{aligned}$$

The last term behaves asymptotically like a constant times τ_{X_1,X_2} as $\tau_{X_1,X_2} \to 0$.

Note that independence of X_1 and X_2 certainly implies $\tau_{X_1,X_2} = 0$. If, on the other hand, $\tau_{X_1,X_2} = 0$ then $U_{X_1,X_2}(1,-1) = 0$, but this does not imply independence unless $0 < \alpha < 1$. We mention some of the properties of the codifference (see also [12, ch. 2.10]).

<u>*Properties*</u> :

(i) τ_{X_1,X_2} is well-defined for $0 < \alpha \le 2$.
(ii) For $\alpha = 2$, $\tau_{X_1,X_2} = \mathrm{Cov}(X_1, X_2)$.
(iii) The codifference is symmetric: $\tau_{X_1,X_2} = \tau_{X_2,X_1}$.
(iv) τ_{X_1,X_2} is non-negative definite.

In order to define the *covariation*, take $\alpha > 1$ and suppose that $X_1 = \int_{\mathbb{R}} f_1(x)\, M_\alpha(\mathrm{d}x)$ and $X_2 = \int_{\mathbb{R}} f_2(x) M_\alpha(\mathrm{d}x)$. The covariation of X_1 and X_2 is given by

$$[X_1, X_2]_\alpha = \int_{\mathbb{R}} f_{t_1}(x) f_{t_2}(x)^{\langle \alpha-1\rangle} \mathrm{d}x \tag{4.2}$$

where $a^{\langle\alpha-1\rangle} = |a|^{\alpha-1}\mathrm{sign}(a)$. It is defined for $1 < \alpha \le 2$.

<u>*Properties*</u> : We refer to [12, ch. 2.7].

(i) If $\alpha = 2, [X_1, X_2]_\alpha = (1/2)\mathrm{Cov}(X_1, X_2)$.

(ii) It shows up naturally in linear regression ([12, ch. 4.1]). If $1 < \alpha \leq 2$, then the *regression of* X_1 *on* X_2 is not only linear (as a function of X_2) but also satisfies

$$\mathbb{E}\,(X_1|X_2) = \frac{[X_1, X_2]_\alpha}{\|X_2\|_\alpha^\alpha} X_2 \quad \text{a.s.}$$

This relation generalizes the well-known relation for jointly Gaussian mean-zero variables X_1, X_2:

$$\mathbb{E}\,(X_1|X_2) = \frac{\mathrm{Cov}\,(X_1, X_2)}{\mathbb{E}X_2^2} X_2 \quad \text{a.s.}$$

(iii) If X_1 and X_2 are independent, then $[X_1, X_2]_\alpha = 0$. The converse is false, unless X_2 is *James orthogonal* to X_1. X_2 is James orthogonal to X_1, symbolized by $X_2 \perp_J X_1$, means

$$\|\lambda X_1 + X_2\|_\alpha \geq \|X_2\|_\alpha$$

for all $\lambda \in \mathbb{R}$. Thus, by [12, Proposition 2.9.2, p. 98]

$$[X_1, X_2]_\alpha = 0 \quad \iff \quad X_2 \perp_J X_1.$$

There are, however, a few “drawbacks” with the covariation.

(i) (4.2) is defined just for $\alpha > 1$. This can be appreciated by applying Hölder's inequality with the exponents $p = \alpha$ and $q = \alpha/(\alpha - 1)$:

$$|[X_1, X_2]_\alpha| \leq \left(\int_{\mathbb{R}} |f_{t_1}(x)|^\alpha \,\mathrm{d}x\right)^{\frac{1}{\alpha}} \left(\int_{\mathbb{R}} |f_{t_2}(x)|^\alpha \,\mathrm{d}x\right)^{\frac{\alpha-1}{\alpha}} = \|X_1\|_\alpha \|X_2\|_\alpha^{\alpha-1}.$$

(ii) It is not symmetric for $\alpha < 2$: $[X_1, X_2]_\alpha \neq [X_2, X_1]_\alpha$.

(iii) It is linear in the first argument, but not in the second, if $\alpha \leq 2$

$$[X_1, X_2 + X_3]_\alpha \neq [X_1, X_2]_\alpha + [X_1, X_3]_\alpha\,,$$

unless X_2 and X_3 are independent.

5 – Application to log-fractional stable noise

We want to study the asymptotic behavior as $t \to \infty$ of log-fractional stable noise (log-FSN), namely the increment process Y_t of log-FSM. Since it is stationary, we need only consider (Y_t, Y_0). From (4.1) its codifference is

$$\tau_{Y_t,Y_0} = \|Y_t\|_\alpha^\alpha + \|Y_0\|_\alpha^\alpha - \|Y_t - Y_0\|_\alpha^\alpha. \tag{5.1}$$

Its covariation, using (3.2) and (4.2), is

$$[Y_t, Y_0]_\alpha = \int_{\mathbb{R}} (\ln|t+1-x| - \ln|t-x|)\,(\ln|1-x| - \ln|-x|)^{\langle\alpha-1\rangle}\,\mathrm{d}x. \tag{5.2}$$

We noted that the codifference is always symmetric, but this is not true for the covariation. However, the covariation of log-FSN is symmetric. Indeed, substituting $y = t+1-x$ in (5.2), we get

$$\begin{aligned}[Y_t, Y_0]_\alpha &= \int_{\mathbb{R}} (\ln|y| - \ln|y-1|)\,(\ln|y-t| - \ln|y-t-1|)^{\langle\alpha-1\rangle}\,\mathrm{d}y\\ &= \int_{\mathbb{R}} (\ln|1-y| - \ln|y|)\,(\ln|t+1-y| - \ln|t-y|)^{\langle\alpha-1\rangle}\,\mathrm{d}y \;=\; [Y_0, Y_t]_\alpha\end{aligned}$$

since $(-1)(-1)^{\langle\alpha-1\rangle} = 1$.

Y_t is a moving average,

$$Y_t = \int_{\mathbb{R}} g(t-x) M_\alpha(\mathrm{d}x)\,.$$

As a consequence, as $t \to \infty$, Y_t and Y_0 are asymptotically independent. Y_t is actually mixing because, denoting it by the map Y, then

$$\lim_{t\to\infty} \mathbb{P}Y^{-1}(S_t(A) \cap B) = \mathbb{P}Y^{-1}(A)\mathbb{P}Y^{-1}(B),$$

where $S_t : \Omega \longrightarrow \Omega$ is the shift transformation on $\Omega = \mathbb{R}^{\mathbb{R}}$ that is defined by $(S_t\omega)(s) = \omega(s+t)$. ($\{S_t\}$ is a family of measure-preserving transformations on Ω [12, ch. 14.4].) One therefore expects as $t \to \infty$

$$\tau_{Y_t,Y_0} \to 0 \qquad \text{and} \qquad [Y_t, Y_0]_\alpha \to 0.$$

The precise rate of convergence of these measures is important, since this rate will characterize the form of asymptotic dependence.

THEOREM 5.1. *Suppose $S\alpha S$ log-FSN, Y_t, is given by* (3.2).

(i) *Its codifference* (5.1) *satisfies*

$$\tau_{Y_t,Y_0} \sim Pt^{1-\alpha} \quad \textit{as } t \to \infty$$

where

$$\begin{aligned} P = &\int_{-\infty}^{1} \left[\left| \frac{1}{1-x} \right|^{\alpha} + \left| \frac{1}{x} \right|^{\alpha} - \left| \frac{1}{1-x} + \frac{1}{x} \right|^{\alpha} \right] \mathrm{d}x \\ &+ \int_{0}^{\infty} \left[\left(\frac{1}{1+x} \right)^{\alpha} + \left(\frac{1}{x} \right)^{\alpha} - \left(\frac{1}{x} - \frac{1}{1+x} \right)^{\alpha} \right] \mathrm{d}x \end{aligned}$$

and $P > 0$.

(ii) *Its covariation* (5.2) *satisfies*

$$[Y_t, Y_0]_\alpha \sim Qt^{1-\alpha} \quad \textit{as } t \to \infty$$

where

$$Q = \int_0^1 \left[(1+x)^{1-\alpha} \left(x^{-1} + x^{\alpha-2} \right) - (1-x)^{1-\alpha} \left(x^{-1} - x^{\alpha-2} \right) \right] \mathrm{d}x$$

and $Q > 0$.

Theorem 5.1 was proved in [8] and the codifference of log-FSN was initially examined in [1](1).

The results show that the codifference and covariation converge to zero hypergeometrically, ct^p, where c is a positive constant and the *rate* $p = 1 - \alpha$ is the same for both. In particular, the non-vanishing of c renders this rate *exact* for either measure. Since $1 < \alpha < 2$, the rate is slow enough so that the series $\sum_{t=1}^{\infty} \tau_{Y_t,Y_0}$ and $\sum_{t=1}^{\infty} [Y_t, Y_0]_\alpha$ diverge. One often asserts in this case that log-FSN and, in turn, log-FSM exhibit long-range dependence.

6 – Comparison with fractional Gaussian noise

Consider the Gaussian H-sssi process

$$B_H(t) = \int_{\mathbb{R}} \left(|t-x|^{H-1/2} - |x|^{H-1/2} \right) M_2(\mathrm{d}x), \quad t \in \mathbb{R}$$

(1) There are some typographical errors in [12, Theorem 7.10.1, p. 368 and Theorem 7.10.2, p. 369]. The constant $F(\theta_1, \theta_2)$ is correct but the constants $B(\theta_1, \theta_2)$ and $G(\theta_1, \theta_2)$ are not. To correct $B(\theta_1, \theta_2)$ in Theorem 7.10.1, the constant $-b\theta_2$ should replace $b\theta_2$ in the first term of the integrand of $\int_0^1$. In Theorem 7.10.2, replace $1 + x$ by $1 - x$ in the integrand of $\int_{-\infty}^1$. The correct versions are stated in [1, Theorem 2.1 and Theorem 2.4].

where $0 < H \leq 1$, known as fractional Brownian motion (FBM). Its increment process, fractional Gaussian noise (FGN) is

$$\begin{aligned}\Delta B_H(t) &= B_H(t+1) - B_H(t) = \\ &= \int_{\mathbb{R}} \left(|t+1-x|^{H-1/2} - |t-x|^{H-1/2}\right) M_2(\mathrm{d}x).\end{aligned} \tag{6.1}$$

The covariance of FGN satisfies

$$r_t = \mathrm{Cov}(\Delta B_H(t), \Delta B_H(0)) \sim C_H t^{2H-2} \quad \text{as } t \to \infty$$

where $C_H = \mathbb{E}B_H^2(1)H(2H-1)$ ([12, Proposition 7.2.10, p. 335]).

Now restrict H to the range

$$1/2 < H < 1.$$

Then $C_H > 0$ and

$$\sum_{t=1}^{\infty} r_t \sim \sum_{t=1}^{\infty} C_H t^{2H-2} = \infty,$$

so that FGN exhibits long-range dependence.

The dependence structures of FGN and log-FSN ((3.2)) share some common attributes.

(i) The constants of asymptoticity are positive for both processes: $C_H > 0$ for FGN and $P > 0$ and $Q > 0$ in Theorem 5.1.
(ii) The exponents $2H-2$ (FGN) and $1-\alpha$ (log-FSN) have the same extreme values: the exponent is -1 for FGN with $H \to 1/2$ and for log-FSN with $\alpha \to 2$, while it is 0 for FGN with $H \to 1$ and for log-FSN with $\alpha \to 1$. Thus, the ranges of the exponents are the same interval $(-1, 0)$ of values.
(iii) In that range $(-1, 0)$ we have long-range dependence displayed by both processes. For FGN, the sum of the covariances diverges ($\sum_{t=1}^{\infty} r_t = \infty$), and for log-FSN, the sum of the codifferences diverges ($\sum_{t=1}^{\infty} \tau_{Y_t,Y_0} = \infty$) and the sum of the covariations diverges ($\sum_{t=1}^{\infty} [Y_t, Y_0]_\alpha = \infty$).

The dependence nevertheless is due to different sources. Both processes are parametrized by a single parameter, H for FGN and α for log-FSN. The dependence for FGN arises from the presence of H in the integrand in (6.1). By contrast, the integrand is fixed in log-FSN, $Y_t = \int_{\mathbb{R}} (\ln|t+1-x| - \ln|t-x|)\, M_\alpha(\mathrm{d}x)$, but the dependence is due to the presence of α in the random measure.

7 – Concluding remarks and extensions

We have observed that log-FSM becomes Brownian motion when $\alpha = 2$. What if one alters the kernel of log-FSM in (3.1), replacing the logarithm by a power function? One gets

$$X_t = \int_{\mathbb{R}} \left(|t-x|^{H-1/\alpha} - |x|^{H-1/\alpha}\right) M_\alpha(\mathrm{d}x), \quad t \in \mathbb{R}. \tag{7.1}$$

This process is called *linear fractional stable motion* (LFSM) ([12, ch. 7.4]). It is defined for $0 < \alpha < 2$ and $0 < H < 1$, provided $H \neq 1/\alpha$. If $H = 1/\alpha$, it is ordinarily identified as a generalization of α-stable motion, which has independent increments.

When $H = 1/\alpha$, one could also identify LFSM with log-FSM, since

$$\begin{aligned}\frac{1}{H-1/\alpha}\left(|t-x|^{H-1/\alpha} - |x|^{H-1/\alpha}\right) &= \frac{|t-x|^{H-1/\alpha}-1}{H-1/\alpha} - \frac{|x|^{H-1/\alpha}-1}{H-1/\alpha}\\ &= \frac{e^{(H-1/\alpha)\ln|t-x|}-1}{H-1/\alpha} - \frac{e^{(H-1/\alpha)\ln|x|}-1}{H-1/\alpha}\\ &\longrightarrow \ln|t-x| - \ln|x|\end{aligned}$$

as $H \to 1/\alpha$, for any $x \neq 0, t$.

LFSM also becomes FBM when $\alpha = 2$.

There are also extensions of LFSM obtained by substituting for the absolute value in (7.1) a linear combination of the positive and negative parts:

$$\begin{aligned}X_{a,b;t} = \int_{\mathbb{R}} \Big(a\left[(t-x)_+^{H-1/\alpha} - (-x)_+^{H-1/\alpha}\right] &+ \\ + b\left[(t-x)_-^{H-1/\alpha} - (-x)_-^{H-1/\alpha}\right]\Big) &M_\alpha(\mathrm{dx}),\end{aligned} \tag{7.2}$$

where a and b are real-valued constants, not both equal to 0, and

$$x_+ = \begin{cases} x & \text{if } x \geq 0 \\ 0 & \text{if } x < 0, \end{cases} \qquad x_- = \begin{cases} 0 & \text{if } x \geq 0 \\ -x & \text{if } x < 0. \end{cases}$$

The process $X_{a,b;t}, t \in \mathbb{R}$ in (7.2) is also called LFSM, although it has essentially different finite-dimensional distributions as a and b take different values ([12, Theorem 7.4.5, p. 347]). The instance (7.1) is recovered by setting $a = b$. The case $a \neq 0, b = 0$ is non-anticipative (or causal) and the case $a = 0, b \neq 0$ is anticipative. These processes have stationary increments; in fact, the difference process

$$Y_{a,b;t} = X_{a,b;t+1} - X_{a,b;t}$$

is known as *linear fractional stable noise* (LFSN). Since $Y_{a,b;t}$ is stationary and $S\alpha S$, one can inquire about the asymptotic behavior of its codifference if $0 < \alpha < 2$ and its covariation when $1 < \alpha < 2$. There is burgeoning research on this topic. Refer to [9] for related results when a and b are restricted to $a \neq 0, b = 0$ and $a = 0, b \neq 0$. In both cases this behavior is also hypergeometric, ct^p with $p < 0$ and, more importantly, $c \neq 0$, so again the rates are exact. On the other hand, their precise asymptotic behavior for *arbitrary* a and b is more complicated and currently is being examined by the authors.

In view of the preceding discussion comparing (7.2) and (7.1), one may wonder what happens if also the absolute values in the representation (3.1) of log-FSM are replaced by a linear combination of positive and negative parts; that is, if one considers the process

$$Z(t) = \int_{\mathbb{R}} \left(a\,[\ln_0(t-x)_+ - \ln_0(-x)_+] + b\,[\ln_0(t-x)_- - \ln_0(-x)_-]\right) M_\alpha(\mathrm{d}x),$$

where $\ln_0 x = \ln x$ if $x > 0$ and $= 0$ otherwise. We also intend to study its asymptotic dependence structure. Observe, however, such a process falls outside our present framework because it is no longer H-ss ([12, p. 355]), unless $a = b$, in which case it is log-FSM.

REFERENCES

[1] A. Astrauskas – J. B. Levy – M. S. Taqqu: *The asymptotic dependence structure of the linear fractional Lévy motion*, Lithuanian Journal of Mathematics **31** (1) (1991), 1-28.

[2] D. M. Cifarelli – E. Regazzini: *De Finetti's contribution to probability and statistics*, Statistical Science **11** (4) (1996), 253-282.

[3] B. de Finetti: *Sulle funzioni ad incremento aleatorio*, Rendiconti della R. Accademia Nazionale dei Lincei **10** (1929), 163-168.

[4] P. Doukhan – G. Oppenheim – M. S. Taqqu: *Theory and Applications of Long-Range Dependence*, Birkhäuser, Boston, 2003.

[5] P. Embrechts – M. Maejima: *Selfsimilar Processes*, Princeton University Press, Princeton, NJ, 2002.

[6] P. Embrechts – C. Klüppelberg – T. Mikosch: *Modelling Extremal Events*, Springer-Verlag, Berlin, 1997.

[7] Y. Kasahara – M. Maejima – W. Vervaat: *Log-fractional stable processes*, Stochastic Processes and Their Applications **30** (1988), 329-339.

[8] J. B. Levy – M. S. Taqqu: *The asymptotic codifference and covariation of log-fractional stable noise*, Preprint (2006).

[9] J. B. Levy – M. S. Taqqu: *The asymptotic dependence of one-sided linear-fractional stable noise*, Preprint (2005).

[10] F. Mainardi – S. Rogosin: *The origin of infinitely divisible distributions: from de Finetti's problem to Lévy-Khintchine formula*, in: Mathematical Methods in Economics and Finance, Vol. 1 (2006), 37-55.

[11] S. Rachev – S. Mittnik: *Stable Paretian Models in Finance*, Wiley, Chichester, 2000.

[12] G. Samorodnitsky – M. S. Taqqu: *Stable Non-Gaussian Processes: Stochastic Models with Infinite Variance*, Chapman and Hall, New York, 1994.

[13] M. S. Taqqu: *Fractional Brownian motion and long-range dependence*, in: Theory and Applications of Long-Range Dependence (P. Doukhan, G. Oppenheim, M. S. Taqqu, eds.), Birkhäuser, Boston, 2003, 5–38.

Lavoro pervenuto alla redazione il 12 ottobre 2007
ed accettato per la pubblicazione il 18 ottobre 2007.
Bozze licenziate il 9 aprile 2008

Indirizzo degli Autori:

Murad S. Taqqu – Department of Mathematics and Statistics – Boston University – 111 Cummington St. – Boston, MA 02215, USA
E-mail: murad@math.bu.edu

Joshua B. Levy – School of Business – The University of Texas of the Permian – 4901 E. University Blvd. – Odessa, TX 79762, USA
E-mail: levy_j@utpb.edu

This research was partially supported by the NSF Grant DMS-0505747 at Boston University.