Introduction

As is well known, representation theory began with the work of G. Frobenius, who invented the notion of a character of a noncommutative finite group and solved a highly nontrivial problem of describing irreducible characters of symmetric groups $S(n)$ (see his paper [41] of 1900, and also Curtis [30]). Twenty-five years later, H. Weyl computed the irreducible characters of the compact classical Lie groups $U(N)$, $SO(N)$, $Sp(N)$ (see his famous book [133] and references therein). These results of Frobenius and Weyl form the basis of the whole representation theory of groups, and in one way or another they usually appear in any introductory course on finite-dimensional group representations (e.g., Fulton and Harris [42], Goodman and Wallach [52], Simon [111], Zhelobenko [136]).

It is a remarkable fact that character theory can be built for *infinite-dimensional analogs* of symmetric and classical groups if one suitably modifies the notion of the character. This was independently discovered by E. Thoma in the 1960s [117] for the infinite symmetric group $S(\infty)$ and by D. Voiculescu in the 1960s [130], [131] for infinite-dimensional classical groups $U(\infty)$, $SO(\infty)$, $Sp(\infty)$. It turned out that, for all these groups, the so-called extreme characters (analogs of the irreducible characters) depended on countably many continuous parameters, and in the two cases, i.e., for $S(\infty)$ and infinite-dimensional classical groups, the formulas looked very similar.

In spite of all the beauty of Thoma's and Voiculescu's results, they looked too unusual and even exotic, and were largely away from the principal routes of representation theory that formed the mainstream in the 1960s and 70s. It took time to appreciate their depth and realize what kind of mathematics lies behind them. Thoma's and Voiculescu's original motivation came from the theory of von Neumann factors and operator algebras. Nowadays we can point out some connections with other areas of mathematics. First of all, those are:

(a) combinatorics of symmetric functions and multivariate special functions of hypergeometric type;

and

(b) probabilistic models of mathematical physics: random matrices, determinantal point processes, random tiling models, Markov processes of infinitely many interacting particles. Let us emphasize that such connections to probability theory and mathematical physics are new; previously known ones were of a different kind (see, e.g., P. Diaconis' book [31]).

Pioneering work of A.M. Vershik and S.V. Kerov (see [63], [121], [122], [123], [124], [125], [126]) played a key role in bringing forward the combinatorial and probabilistic aspects of the representation theory of $S(\infty)$ and infinite-dimensional classical groups. We will say more about their work below.

—

The goal of this book is to provide a detailed introduction to the representation theory of the infinite symmetric group that would be accessible to graduate and advanced undergraduate students. The amount of material that would be required for the reader to know in advance is rather modest: some familiarity with representation theory of finite groups and basics of functional analysis (measure theory, Stone–Weierstrass' theorem, Choquet's theorem on extreme points of a compact set, Hilbert spaces) would suffice. Theory of symmetric functions plays an important role in our approach, and while some previous exposure to it would be useful, we also provide all the necessary background along the way.

We aimed at writing a relatively short, simple, and self-contained book and did not try to include everything people know about representations of $S(\infty)$. We also do not touch upon representations of infinite-dimensional classical groups – while being analogous, that theory is somewhat more involved. The infinite symmetric group can be viewed as a "toy model" for infinite-dimensional classical groups. We feel that it makes sense to begin the exposition with $S(\infty)$, in parallel to how the subject of representation theory historically and logically started from the finite symmetric groups $S(n)$ that form the simplest and most natural family of finite noncommutative groups.

Knowledge of the material in this book should be sufficient for understanding research papers on representations of $S(\infty)$ and infinite-dimensional groups as well as their applications. Speaking of applications, we first of all mean the class of probabilistic models of mathematical physics where representation theoretic ideas turned out to be remarkably successful. We hope that probabilists and mathematical physicists interested in representation theoretic mechanisms behind such applications would find the book useful.

—

We have not defined our main object of study yet; let us do that now. One could give different (meaningful) definitions of the infinite symmetric group. In this book we define $S(\infty)$ as the group of all finite permutations of the set $\mathbb{Z}_{>0} := \{1, 2, 3, \dots\}$, where the condition of a permutation $s : \mathbb{Z}_{>0} \to \mathbb{Z}_{>0}$ being finite means that $s(j) \neq j$ for finitely many $j \in \mathbb{Z}_{>0}$. Equivalently, $S(\infty)$ can be defined as the inductive limit (or simply the union) of the infinite chain $S(1) \subset S(2) \subset S(3) \subset \cdots$ of growing finite symmetric groups. The group $S(\infty)$ is countable, and it is reasonable to view it as a natural (yet not canonical) infinite analog of the finite symmetric groups $S(n)$. Infinite-dimensional classical groups are defined in a similar fashion as inductive limits of finite-dimensional classical Lie groups.

The first part of the book deals with characters of $S(\infty)$. Similarly to the case of finite groups, a substantial part of representation theory can be built in the language of characters without even mentioning actual representations. Many applications would just require operating with characters. We believe, however, that ignoring representations behind the characters takes away an essential part of the subject and may eventually negatively influence future developments. For that reason, in the second part we turn to unitary representations.

Let us now describe the content of the book in a little more detail.

—

The first part of the book is devoted to Thoma's theorem and related topics. As was mentioned above, Thoma's theorem is an analog of the classical Frobenius theorem on irreducible characters of the finite symmetric groups. The word "analog" should be taken with a pinch of salt here. The point is that $S(\infty)$ does not have conventional irreducible characters (except for two trivial examples), and the notion needs to be revised. Here is a definition given by Thoma that we use:

(a) A *character* of a given group K is a function $\chi : K \to \mathbb{C}$ that is positive definite, constant on conjugacy classes, and normalized to take value 1 at the unity of the group. (For topological groups one additionally assumes that the function χ is continuous.)

(b) An *extreme character* is an extreme point of the set of all characters. (This makes sense as the characters, as defined in (a), form a convex set.)

In the case where the group K is finite or compact, the set of all characters (in the sense of the above definition) is a simplex whose vertices are the extreme characters. Those are exactly the *normalized* irreducible characters, i.e., functions of the form $\chi^\pi(g)/\chi^\pi(e)$, where π denotes an arbitrary irreducible representation (it is always finite-dimensional), $\chi^\pi(g) = \operatorname{Tr}\pi(g)$ is the trace of the operator $\pi(g)$ corresponding to a group element $g \in K$, and the denominator $\chi^\pi(e)$ (the value of χ^π at the unity of the group) coincides with the dimension of the representation. For $S(\infty)$ and its relatives the numerator and denominator in

$$\frac{\chi^\pi(g)}{\chi^\pi(e)} = \frac{\operatorname{Tr}\pi(g)}{\operatorname{Tr}\pi(e)}$$

do not make sense when considered separately, but the notion of extreme character assigns a meaning to their ratio.

According to Frobenius' theorem, irreducible characters of $S(n)$ are parameterized by Young diagrams with n boxes. According to Thoma's theorem, extreme characters of $S(\infty)$ are parameterized by points of an infinite-dimensional space Ω situated inside the infinite-dimensional unit cube; a point $\omega \in \Omega$ is a pair (α, β) of infinite sequences with entries from [0, 1] such that

$$\alpha = (\alpha_1 \geq \alpha_2 \geq \cdots), \quad \beta = (\beta_1 \geq \beta_2 \geq \cdots), \quad \sum_{i=1}^\infty \alpha_i + \sum_{i=1}^\infty \beta_i \leq 1.$$

Direct comparison of the two theorems leads to a somewhat perplexing conclusion that the extreme characters of $S(\infty)$ are both simpler and more complicated than the irreducible characters of $S(n)$. They are more complicated because, instead of the finite set of Young diagrams with n boxes, one gets a countable set of continuous parameters α_i, β_i. But at the same time they are simpler because, for the extreme characters, there is an explicit elementary formula (found by Thoma), while the irreducible characters of finite symmetric groups should be viewed as special functions – there are algorithms for computing them but no explicit formulas.

Thoma's theorem can be (nontrivially) reformulated in several ways. In particular, it is equivalent to:

(1) classifying infinite upper-triangular Toeplitz matrices, all of whose minors are nonnegative (matrices with nonnegative minors are called *totally positive*);
(2) describing all multiplicative functionals on the algebra of symmetric functions that take nonnegative values on the basis of the Schur symmetric functions;
(3) describing the entrance boundary for a certain Markov chain related to the Young graph or, equivalently, describing the extreme points in a suitably defined space of Gibbs measures on paths of the Young graph.

The original proof by Thoma consisted in reduction to (1) and solving the classification problem. Apparently, Thoma did not know that the latter problem had been studied earlier by Schoenberg and his followers ([109], [1], [2]) and its solution had been finalized by Edrei [36].[1]

Interpretations (2) and (3) are due to Vershik and Kerov [122], [124]. The proof of Thoma's theorem that we give is based on (3) and it is an adaptation of an argument from the paper [64] by Kerov, Okounkov, and Olshanski, where a more general result had been proved. We do not focus on extreme characters but rather establish an isomorphism between the convex set of all characters and the convex set of all probability measures on Ω, which immediately implies Thoma's theorem. Although our approach deviates from the proof outlined by Vershik and Kerov in [122], we substantially rely on their *asymptotic method*, which, in particular, explains the nature of Thoma's parameters $\{\alpha_i, \beta_i\}$.

The solution of problem (1) given by Edrei [36] and Thoma [117] is largely analytic, while our approach to the equivalent problem (3) is in essence algebraic; we work with symmetric functions and rely on results of Okounkov, Olshanski [80], and Olshanski, Regev, and Vershik [97], [98].

—

In the second part of the book we move from characters to representations. Our exposition is based on the works of Olshanski [85] and Kerov, Olshanski, and Vershik [67].

There exist two approaches that relate characters to unitary representations. To be concrete, let us discuss extreme characters $\chi = \chi^\omega$ of $S(\infty)$, where $\omega = (\alpha, \beta) \in \Omega$ and (α, β) are Thoma's parameters of χ. The first approach gives the corresponding factor-representations Π^ω of the group $S(\infty)$ (Thoma [117]), while in the second approach one deals with irreducible representations T^ω of the "bisymmetric group" $S(\infty) \times S(\infty)$ (Olshanski [84]). Factor-representations Π^ω are analogs of the irreducible representations π^λ of the finite symmetric groups, but they are (excluding two trivial cases) not irreducible at all: the term "factor" means that these are unitary representations that generate a von Neumann factor (in our case, this is the hyperfinite factor of type II_1). The representations T^ω are exactly those irreducible representations of the bisymmetric group $S(\infty) \times S(\infty)$ that contain a nonzero vector that is invariant with respect to the subgroup diag $S(\infty)$ (the image of $S(\infty)$ under its

[1] Likewise, the extreme characters of $U(\infty)$ correspond to arbitrary (not necessarily upper triangular) totally positive Toeplitz matrices; such matrices were classified, prior to Voiculescu's work, in another paper by Edrei [37]. At the present time, the theory of total positivity became popular due to works of G. Lusztig, S. Fomin, and A. Zelevinsky, but in the sixties and seventies it was much less known.

diagonal embedding in $S(\infty) \times S(\infty)$). Such representations are called *spherical*. Similar representations for finite bisymmetric groups $S(n) \times S(n)$ have the form $\pi^\lambda \otimes \pi^\lambda$, but representations T^ω cannot be written as a (exterior) tensor product of two irreducibles. Such a phenomenon is typical for so-called *wild* (or non-type I) groups, and the infinite symmetric group is one of them.

For wild groups the space of equivalence classes of irreducible representations has pathological structure, and for that reason standard representation theoretic problem settings require modifications. There are many interesting examples of wild groups but there is no general recipe for addressing their representation theories. This highlights the remarkable fact that, for the infinite symmetric group and the infinite-dimensional classical groups, it is possible to develop a meaningful representation theory, and the purpose of the second part of the book is to give an introduction to this theory.

Returning to factor-representations Π^ω and irreducible spherical representations T^ω, let us note that they are closely related: namely, Π^ω is the restriction of T^ω to the subgroup $S(\infty) \times \{e\}$ in the bisymmetric group (that should not be confused with the diagonal subgroup diag $S(\infty)$!). Thus, to a certain extent, the choice between factor-representations of $S(\infty)$ and irreducible representations of $S(\infty) \times S(\infty)$ is a matter of taste (the theories do diverge in further developments, though). We follow the approach of Olshanski [84] and prefer to work with irreducible representations T^ω.

The existence of such representations is a simple corollary of Thoma's theorem. However, that theorem gives no information as to how such representations could be explicitly constructed. This is a typical representation theoretic situation, when it is known how to parameterize the representations but their explicit construction may be completely unobvious and very complicated.

The first realization of representations T^ω was found by Vershik and Kerov [121] (actually, they dealt with factor-representations Π^ω but from here the passage to T^ω is easy). We describe a modification of their construction that employs infinite tensor products of Hilbert spaces in the sense of von Neumann. For generic values of Thoma's parameters, when α- and β-parameters of ω are nonzero, the Hilbert spaces involved in the tensor product construction have a $\mathbb{Z}_2$-grading, i.e., they are super-spaces, and their tensor product has to be understood according to the sign rule from linear super-algebra. The super-algebra actually already appears in the first half of the book – it is present in the formula for Thoma's characters, where super-analogs of Newton power sums arise. Representation theory of the infinite symmetric group is an example of a subject where the need for the use of supersymmetric notions is dictated by the nature of the objects involved rather than by a pure wish to generalize known results to a super-setting.

Let us note that the infinite tensor product construction also gives more general, so-called admissible representations of the bisymmetric group (see Olshanski [85]), and that there is also one more realization due to Okounkov [78], [79].

Next, following the general philosophy of unitary representation theory, we proceed from the theory of irreducible representations to harmonic analysis. The problem of (noncommutative) harmonic analysis is, to a certain extent, similar to Fourier analysis or to expansion in eigenfunctions of a self-adjoint operator. One starts with a ("natural" in some sense) reducible unitary representation, and the problem consists in finding its decomposition on irreducible components. As for self-adjoint operators, the spectrum of the decomposition may be complicated, e.g., not necessarily discrete. In the case of a continuous spectrum one talks about decomposing a representation into a direct integral (rather than a direct sum).

It is a matter of discussion which representations should be considered as "natural" objects for harmonic analysis. However, each finite or compact group has one distinguished representation – the regular representation in the Hilbert space L^2 with respect to the Haar measure on the group. The action of the group is given by left (equivalently, right) shifts. It is even better to consider both left and right shifts together, i.e., the so-called *bi-regular representation* of the direct product of two copies of the group.

For compact (in particular, finite) groups, the decomposition of the bi-regular representation is well known and quite simple (Peter–Weyl's theorem). In particular, the bi-regular representation of the bisymmetric group $S(n) \times S(n)$ is a multiplicity free direct sum of irreducible spherical representations $\pi^\lambda \otimes \pi^\lambda$ that were already mentioned above. Note now that for infinite-dimensional classical groups there is no Haar measure (they are not locally compact) and, therefore, there is no (bi)regular representation.

At first glance, for $S(\infty)$ the situation is different – it is a discrete countable group that carries a Haar measure (which is simply the counting measure), and its bi-regular representation makes perfect sense. However, it ends up being irreducible and thus useless for harmonic analysis.

This dead end turns out to be illusory, and we explain how it can be overcome. The essence of the problem is in the fact that the discrete group $S(\infty)$ is too small to carry a suitable measure with respect to which we would like to build the L^2 space. The way out is in constructing a compactification $\mathfrak{S} \supset S(\infty)$, called the *space of virtual permutations*, that serves as the support of the measure (Kerov, Olshanski, and Vershik [66]).[2] The space $\mathfrak{S}$

[2] The construction of the space $\mathfrak{S}$ was inspired by Pickrell's paper [102]. A close (but not identical) construction is that of *Chinese Restaurant Process*; see Pitman [104].

does not have a group structure, but $S(\infty) \times S(\infty)$ acts on it, and there is a (unique) invariant measure μ on $\mathfrak{S}$ which is finite, as opposed to the counting measure on $S(\infty)$. It is that measure that should be viewed as the correct analog of the Haar measure. Furthermore, the measure μ is just a representative of a whole family of probability measures with good transformation properties that are equally suitable for constructing representations.[3] One thus obtains a whole family $\{T_z\}$ of *generalized bi-regular representations* that depend on a parameter $z \in \mathbb{C}$. The problem of harmonic analysis in our understanding is the problem of decomposing these representations on irreducibles.

We prove that each T_z decomposes on irreducible spherical representations T^ω, and that the decomposition spectrum is simple. We also prove that the spectral measures that govern the decomposition of T_z are mutually singular. This result is important as it implies that the representations T_z are pairwise disjoint, and thus the parameter z is not fictitious.

Investigating the spectral measures goes beyond the goals of this book. It turns out that their structure substantially depends on whether parameter z is an integer or not. These two cases are studied separately and using different means in Kerov, Olshanski, and Vershik [67], and in Borodin and Olshanski [9], respectively. The case of non-integral z is especially interesting as, in the course of its study, one discovers novel models of determinantal random point processes and close connections to random matrix theory.

We conclude with a (short and incomplete) guide to the literature encompassing other aspects of the theory and its further development:

- *A few expository papers* (unfortunately, already rather old): Borodin and Olshanski [8], [12], [14], Olshanski [92].
- *Asymptotic approach to characters of infinite-dimensional classical groups*: Vershik and Kerov [123], Okounkov and Olshanski [81], [82], Borodin and Olshanski [22].
- *Irreducible unitary representations of infinite-dimensional classical groups*: Olshanski [84], [87], [86], Pickrell [103].
- *Quasiinvariant measures for infinite-dimensional classical groups*: Pickrell [102], Neretin [75].
- *Harmonic analysis on $U(\infty)$*: Olshanski [91], Borodin and Olshanski [13], Gorin [54], Osinenko [99].

[3] These measures, called *Ewens measures*, are very interesting in their own right. They are closely related to the *Ewens sampling formula* that is widely used in the literature on mathematical models of population genetics; see, e.g., the survey paper by Ewens and Tavaré [38].

- *The Plancherel measure on partitions*: Kerov [60], [61], [63], Baik, Deift, and Johansson [4], Johansson [58], Borodin, Okounkov, and Olshanski [7], Strahov [115].
- *Other measures on partitions of representation-theoretic origin and their generalizations*: Borodin and Olshanski [9], [10], [15], [16], [18], [24], Borodin, Olshanski and Strahov, [25], Olshanski [90], [93], [96], Strahov [116].
- *Models of Markov dynamics of representation-theoretic origin*: Borodin and Olshanski [17], [19], [20], [21], Olshanski [94], [95].

Acknowledgments. We are grateful to Cesar Cuenca, Vadim Gorin, Leonid Petrov, and the anonymous referee whose remarks helped us to fix typos and improve the exposition. The work of the second-named author (G.O.) was partially supported by the Simons Foundation (the Simons–IUM fellowship).