Mathematical
Surveys
and
Monographs

Volume 77

The Semicircle Law, Free Random Variables and Entropy

Fumio Hiai
Dénes Petz

American Mathematical Society

Contents

Preface

This book is based on the recent brilliant discoveries of Dan Voiculescu, which started from free products of operator algebras, but grew rapidly to include all sorts of other interesting topics. Although we both were fascinated by Voiculescu's beautiful new world from the very beginning, our attitude changed and our interest became more intensive when we got an insight into its interrelations with random matrices, entropy (or large deviations) and the logarithmic energy of classical potential theory.

There are many ways to present these ideas. In this book the emphesis is not put on operator algebras (Voiculescu's original motivation), but on entropy and random matrix models. It is not our aim to make a complete survey of all aspects of free probability theory. Several important recent developments are completely missing from this book. Our emphasis is on the role of random matrices. However, we do our best to make the presentation accessible for readers of different backgrounds.

The basis of this monograph was provided by lectures delivered by the authors at Eötvös Loránd University in Budapest, at Hokkaido University in Sapporo, and at Ibaraki University in Mito.

The structure of the monograph is as follows. Chapter 1 makes the connection between the concepts of probability theory and linear operators in Hilbert spaces. A sort of ideological foundation of noncommutative probability theory is presented here in the form of many examples. Chapter 2 treats the fundamental free relation. Again several examples are included, and the algebraic and combinatorial aspects of free single and multivariate random variables are discussed. This chapter is a relatively concise, elementary and selfcontained introduction to free probability. The analytic aspects come in the next chapter. The infinitely divisable laws show an analogy with classical probability theory. This chapter is not much required to follow the rest of the monograph. Chapter 4 introduces the basic random matrix models and the limit of their eigenvalue distribution. Voiculescu's concept of asymptotic freeness originated from independent Gaussian random matrices. Since its birth, asymptotic freeness has been a very important bridge between free probability and random matrix theory. The strong analogy between the free relation and statistical independence is manifested in the asymptotic free relation of some independent matrix models. Entropy appears on the stage in Chapter 5—first the Boltzmann-Gibbs entropy, which is considered here as the rate function in some large deviation theorems. The frequent random matrix ensembles are charecterized by maximization of the Boltzmann-Gibbs entropy under certain constraints. Several large deviation results are given following the pioneering work of Ben Arous and Guionnet on symmetric Gaussian random matrices. The main ingredient of the rate functional is the logarithmic energy, familiar from potential theory. For an n-tuple of noncommuative random variables, the probabilistic–measure theoretic model is lacking; hence Chapter 6 is technically in the field of functional analysis. Properties of Voiculescu's multivariate free entropy are discussed in the setting of operator algebras, and we introduce an analogous concept for n-tuples of unitaries. Chapters 3–6 comprise the main part of the monograph. The last chapter is mostly on free group factors, and gives ideas on applications to operator algebras.

Since rather different areas in mathematics are often combined, it was our intention to make the material nearly self-contained for the sake of convenience. This was a heavy task, and we had to cope with the combination of probabilistic, analytic, algebraic and combinatorial arguments. Each chapter concludes with some notes giving information on our sources and hints on further develpoments. Furthermore, we supply standard references for the reader who is not familiar with the general background of the chapter. The "Overview" is an attempt to show the place of the subject and to give orientation. It replaces an introduction, and the reader is invited to consult this part either before or after studying the much more technical chapters.

We thank many colleagues for helping us to finish this enterprise. Imre Csiszár, Roland Speicher, and Masaki Izumi can be named specifically. We are also grateful for several institutions for supporting our collaboration: funds of the Hungarian Academy of Sciences (OTKA F023447 and AKP 96/2-6782), of the Canon Foundation, of the Grant-in-Aid for Scientific Research (C)09640152, and of the Erdős Center are acknowledged.

Fumio Hiai and Dénes Petz

Overview

0.1 The isomorphism problem of free group factors

John von Neumann established the theory of so-called *von Neumann algebras* in the 1930's. The comprehensive study of "rings of operators" (as von Neumann algebras were called at that time) was motivated by the spectral theorem of selfadjoint Hilbert space operators and by the needs of the mathematical foundation of quantum mechanics. A von Neumann algebra is an algebra of bounded linear operators acting on a Hilbert space which is closed with respect to the topology of pointwise convergence. *Factors* are in a sense the building blocks of general von Neumann algebras; they are von Neumann algebras with trivial center. In a joint paper with F.J. Murray, a classification of the factors was given. Von Neumann was fond of the type II_1 factors, which are continuous analogues of the finite-dimensional matrix algebras. The normalized trace functional on the algebra of $n \times n$ matrices is invariant under unitary conjugation, and it takes the values k/n $(k = 0, 1, \dots, n)$ on projections. A type II_1 factor admits an abstract trace functional τ which is invariant under unitary conjugation, and it can take any value in $[0, 1]$ on projections. The N-fold tensor product of 2×2 matrix algebras is nothing else than the matrix algebra of $2^N \times 2^N$ matrices on which the normalized trace of projections is in the set $\{k/2^N : k = 0, 1, 2, \dots, 2^N\}$. In the limit as $N \to \infty$, the dyadic rationals fill the interval $[0, 1]$ and we arrive at a type II_1 factor. What we are constructing in this way is the infinite tensor product of 2×2 matrices, and the construction yields the *hyperfinite* type II_1 factors. ("Hyperfinite" means a good approximation by finite-dimensional subalgebras; in the above case approximation is by the finite tensor products with growing size.) This was the first example of a type II_1 factor. Murray and von Neumann showed that any two hyperfinite type II_1 factors are isomorphic, and they were looking for a non-hyperfinite factor.

Countable discrete groups give rise to von Neumann algebras; in fact one can associate to a discrete group G a von Neumann algebra $\mathcal{L}(G)$ in a canonical way. On the Hilbert space $\ell_2(G)$ the group G has a natural unitary representation $g \mapsto L_g$, the so-called left regular one, which is given by

$$(L_g\xi)(h) := \xi(g^{-1}h) \qquad (\xi \in \ell^2(G),\ g, h \in G).$$

The group ring $R(G)$ is the linear hull of the set $\{L_g : g \in G\}$ of unitaries. The *group von Neumann algebra* $\mathcal{L}(G)$ associated to G is by definition the closure of $R(G)$ in the topology of pointwise convergence. If the group under consideration is ICC (i.e. all its non-trivial conjugacy classes contain infinitely many elements), then the von Neumann algebra $\mathcal{L}(G)$ is a factor. When the closure of $R(G)$ is taken with respect to the norm topology of $B(\ell^2(G))$, we arrive at another important object, that is, the *reduced group C^*-algebra $C_r^*(G)$*. There exists a *canonical trace* τ on $\mathcal{L}(G)$, which is given by the unit element e of G. Let $\delta_e \in \ell_2(G)$ stand for the characteristic function of $\{e\}$ and define

$$\tau(\,\cdot\,) := \langle \,\cdot\, \delta_e, \delta_e \rangle \,.$$

Then it is easy to check that τ is a trace, i.e. it satisfies

$$\tau(ab) = \tau(ba) \quad \text{for all } \ a, b \in \mathcal{L}(G).$$

Von Neumann started from the free group with two generators and proved that the corresponding factor is not hyperfinite. Historically this led to the first example of two non-isomorphic type II_1 factors. Much later it was discovered that the group factor of an ICC group is hyperfinite if and only if the group itself is amenable, and free groups are the simplest non-amenable groups. (The concept of *amenable groups* also goes back to von Neumann.) Actually, von Neumann showed that the free product of groups leads to a non-hyperfinite factor. It seems that Richard Kadison was the person who explicitly posed the question of whether free groups with different numbers of generators could produce the same factor. This question is still open, and it was the main motivation for Dan Voiculescu to study the free relation and to develop free probability theory.

Let $\mathbf{F}_n$ denote the *free group with n generators.* If $n \neq m$ then $\mathbf{F}_n$ and $\mathbf{F}_m$ are not isomorphic. This can be seen by considering the group homomorphisms from $\mathbf{F}_n$ to $\mathbb{Z}_2$, the two element group, which is actually a field. Consider the space X_n of all homomorphisms from $\mathbf{F}_n$ to $\mathbb{Z}_2$. This is a bimodule over $\mathbb{Z}_2$ of dimension exactly 2^n. Since a group isomorphism between $\mathbf{F}_n$ and $\mathbf{F}_m$ induces a module isomorphism between X_n and X_m, the groups $\mathbf{F}_n$ and $\mathbf{F}_m$ for $n \neq m$ cannot be isomorphic. Although the free group $\mathbf{F}_n$ is contained in $\mathcal{L}(\mathbf{F}_n)$ in the form of the group ring, $\mathcal{L}(\mathbf{F}_n)$ is much larger than $R(\mathbf{F}_n)$ due to the closure procedure in the definition of the group factor. Hence simple proofs, for example the proof that $\mathbf{F}_2$ is not isomorphic to $\mathbf{F}_3$, do not extend to the topological closures, to the group C^*-algebras, or to the group von Neumann algebras.

If one expects $\mathcal{L}(\mathbf{F}_n)$ and $\mathcal{L}(\mathbf{F}_m)$ to be non-isomorphic, a possible strategy to prove this is to read out n intrinsically from $\mathcal{L}(\mathbf{F}_n)$. Each von Neumann factor of type II_1 has a unique canonical (faithful normal) tracial state τ, so this is intrinsically at our disposal. A.N. Kolmogorov in 1958 and Ya.G. Sinai in 1959 introduced the so-called Kolmogorov-Sinai (or dynamical) entropy of a measure-preserving transformation on a probability space, which has been a very successful tool in the isomorphism problem of Bernoulli shifts. Bernoulli shifts are simple, but they are the most important probabilistic dynamical systems. It turned out that the Kolmogorov-Sinai entropy is a complete invariant for Bernoulli shifts; in 1970

Donald Ornstein proved that two Bernoulli shifts are conjugate if and only if they have the same dynamical entropy. It seems to the authors that Dan Voiculescu was deeply influenced by these ideas. Independently of the success of his approach toward the isomorphism problem, his analysis created a new world in which the free relation and the free entropy of noncommutative random variables play the leading roles. Random matrices give the simplest examples of noncommutative random variables; they can model the free relation and are present in the definition of free entropy.

0.2 From the relation of free generators to free probability

The free group factor $\mathcal{L}(\mathbf{F}_n)$ is an important example in von Neumann algebra theory, since it is the first-found and simplest non-hyperfinite type II_1 factor. The free group itself is an important object in harmonic analysis; in fact, there is a strong and intimate relation between harmonic analysis on the free group and the structure of the free group factor.

The fundamental relation of the generators of a free group is the algebraic "free relation"; namely, there are no identities expressed in operations of the group which are satisfied by the generators. Eventually, this algebraic free relation is present in non-free groups as well. For instance, if G is the free product of G_1 and G_2, then the elements of G_1 are in free relation to the elements of G_2. The study of harmonic analysis on free product groups is the obvious continuation of its study on free groups.

Next we outline an idea which starts from a random walk on free groups and leads to the key concept of free relation; on top of that, the role of the semicircle law becomes clear as well. Let $g_1, g_2, \dots, g_n$ be generators of $\mathbf{F}_n$ and consider a random walk on this group which starts from the group unit and goes from the group element g to hg with probability $1/2n$ if $h \in \{g_1, g_2, \dots, g_n, g_1^{-1}, g_2^{-1}, \dots, g_n^{-1}\}$. Then the probability of return to the unit e in m steps is of the form

$$P(n,m) := \frac{1}{(2n)^m} \big\langle \big(L_{g_1} + L_{g_1^{-1}} + L_{g_2} + L_{g_2^{-1}} + \dots + L_{g_n} + L_{g_n^{-1}}\big)^m \delta_e, \delta_e \big\rangle \,.$$

(This vanishes if m is not even.) In the probabilistic interpretation of quantum mechanics, it is standard to interpret the number $\langle A\xi, \xi\rangle$ as the expectation of the operator A in the state vector ξ. Adopting this view, we have the expectation

$$\Big\langle \Big(\sum_{i=1}^{n} A_i^{(n)}\Big)^m \delta_e, \delta_e \Big\rangle$$

of selfadjoint operators, $A_i^{(n)}$ standing for $(L_{g_i} + L_{g_i^{-1}})/\sqrt{2}$. The operators $A_i^{(n)}$

might be called random variables in the sense that they have expectations, and to distinguish this concept from classical probability theory, we speak of *noncommutative random variables*. The asymptotic behavior of the return probability $P(n, 2m)$ is given as follows:

$$P(n, 2m) \approx \frac{1}{(2n)^m} \frac{1}{m+1} \binom{2m}{m} \quad \text{as} \quad n \to \infty.$$

What we have is a sort of *central limit theorem* for the array

$$A_1^{(n)}, A_2^{(n)}, \dots, A_n^{(n)}$$

of noncommutative random variables, because

$$\frac{A_1^{(n)} + A_2^{(n)} + \ldots + A_n^{(n)}}{\sqrt{n}}$$

converges in moments (or in distribution) to the even classical probability density whose $2m$th moment is the so-called *Catalan number*

$$\frac{1}{m+1} \binom{2m}{m} \qquad (m \in \mathbb{N}).$$

This limit density is the *semicircle* or *Wigner law* $\frac{1}{2\pi}\sqrt{4 - x^2}$ with support on $[-2, 2]$.

The different generators g_i are free in the algebraic sense, and this led Voiculescu to call the relation of the selfadjoint operators $A_i^{(n)}$ (for a fixed n) free as well. When one aims to formulate the concept in the spirit of probability, the plausible free relation of these noncommutative random variables must be formulated in terms of expectation. The noncommutative random variables $A_1, A_2, \dots, A_n$ are called *free* with respect to the expectation φ if

$$\varphi\big(P_1(A_{i(1)})P_2(A_{i(2)}) \ldots P_k(A_{i(k)})\big) = 0$$

whenever $P_1, P_2, \dots, P_n$ are polynomials, $P_j(A_{i(j)}) = 0$ and $i(1) \neq i(2) \neq \ldots \neq i(k)$. In the motivating example we had $\varphi(\,\cdot\,) = \langle \cdot\, \delta_e, \delta_e \rangle$, and of course the above defined operators $A_1^{(n)}, A_2^{(n)}, \dots, A_n^{(n)}$ are free in the sense of Voiculescu's definition. It is not immediately obvious from the definition that the free relation of the noncommutative random variables $A_1, A_2, \dots, A_n$ is a particular rule to calculate the joint moments

$$\varphi\big(A_{i(1)}^{m(1)} A_{i(2)}^{m(2)} \ldots A_{i(k)}^{m(k)}\big)$$

(for positive integers $m(1), m(2), \dots, m(k)$) from the moments $\varphi(A_i^m)$ of the given variables.

It is always tempting to compare Voiculescu's free relation with the independence of classical random variables. The comparison cannot be formal, since the free relation and independence do not take place at the same time. What we have in mind is only the analogy of the free relation in noncommutative probability theory with independence in classical probability theory. The part of noncommutative probability in which the free relation plays a decisive role is called *free probability theory*.

Free probability theory has its celebrated distributions. The semicircle law is one of those celebrities; it comes from the free central limit theorem and its moment sequence is the Catalan numbers. The free analogue of the Poisson limit theorem can be established, and it gives the free analogue of the Poisson distribution. Given measures μ and ν, we can find noncommutative random variables a and b with these distributions. When a and b are in free relation, the distribution measure of $a+b$ (or ab) can be called the *additive* (or *multiplicative*) *free convolution* of μ and ν. The notations $\mu \boxplus \nu$ and $\mu \boxtimes \nu$ are used to denote the two kinds of free convolution. The class of semicircle laws is closed under additive free convolution. The distributions $\mu_0 \boxplus w_{\sqrt{4t}}$ form a convolution semigroup when $w_{\sqrt{4t}}$ stands for the semicircle law with variance t (which corresponds to radius $\sqrt{4t}$). The analytic machinery to handle additive free convolution is based on the Cauchy transform of measures. If μ_t is a freely additive convolution semigroup and $G(z,t)$ is the Cauchy transform of μ_t, then the complex Burger equation

$$\frac{\partial G(z,t)}{\partial t} + R(G(z,t)) = \frac{\partial G(z,t)}{\partial z}$$

is satisfied with initial condition $G(z,0) = G_{\mu_0}(z)$, and R is the so-called R-transform of μ_1. If μ_t is the semigroup of semicircle laws, then $R(z) = z$ and we have the analogue of the heat equation.

There is an obvious way to associate a measure to a noncommuting random variable via the spectral theorem if it is a normal operator. If the operator is not normal, then a measure can still be constructed by using type II_1 von Neumann algebra techniques. (What we have in mind is the Brown measure of an element of a type II_1 von Neumann algebra.) However, if we deal with several noncommuting random variables which are really noncommuting operators, then there is no way to reduce the discussion to measures. At this level of generality free probability theory has to work with joint moments, power series on noncommuting indeterminates, and a new kind of combinatorial arguments. The picture becomes very different from classical probability theory. Computation of the joint moments of free noncommutative random variables is a rather combinatorial procedure.

0.3 Random matrices

The joint eigenvalue density for certain symmetric random matrices has been known for a long time. J. Wishart found it for the so-called (real) Wishart matrix back in 1928. In quantum physics the energy is represented by the Hamiltonian operator, and one is interested in the point spectrum of the Hamiltonian, which is the set of eigenvalues. The problem is difficult; we do not know the exact Hamiltonian, and even if we did, it would be too complicated to find the eigenvalues. An approach to this problem is based on the following statistical hypothesis: The statistical behavior of the energy levels is identical with the behavior of the eigenvalues of a large random matrix. Reasoning along this line, in 1955 E.P. Wigner obtained the semicircle law for the limiting eigenvalue density of random matrices with independent Bernoulli entries. Wigner's work initiated a huge and deep interest in random matrices from theoretical and nuclear physicists. From the point of view of physical applications the most interesting question is to treat the correlation functions of the eigenvalues and the so-called "level spacing".

Random matrices are noncommutative random variables with respect to the expectation

$$\tau_n(H) := \frac{1}{n}\sum_{i=1}^{n} E(H_{ii})$$

for an $n \times n$ random matrix H, where E stands for the expectation of a classical random variable. It is a form of the *Wigner theorem* that

$$\tau_n(H(n)^{2m}) \to \frac{1}{m+1}\binom{2m}{m} \quad \text{as} \quad n \to \infty$$

if the $n \times n$ real symmetric random matrix $H(n)$ has independent identical Gaussian entries $N(0, 1/n)$ so that $\tau_n(H(n)^2) = 1$. The semicirle law is the limiting eigenvalue density of $H(n)$'s as well as the limiting law of the free central limit theorem in the previous section. The reason why this is so was made clear by Voiculescu. Let $X_1(n), X_2(n), \dots$ be independent random matrices with the same distribution as $H(n)$. It follows from the properties of Gaussians that the distribution of the random matrix

$$\frac{X_1(n) + X_2(n) + \cdots + X_n(n)}{\sqrt{n}}$$

is the same as that of $H(n)$. Hence the convergence in moments to the semicircle law would be understandable if $X_1(n), X_2(n), \dots, X_n(n)$ were in free relation. Their free relation with respect to the expectaion τ_n would include the condition

$$\tau_n\Big(\big[X_1(n)^k - \tau_n(X_1(n)^k)\big]\big[X_2(n)^l - \tau_n(X_2(n)^l)\big]\Big) = 0\,,$$

which is equivalently written as

$$\tau_n\big(X_1(n)^k X_2(n)^l\big) = \tau_n\big(X_1(n)^k\big)\tau_n\big(X_2(n)^l\big)\,. \tag{1}$$

For notational simplicity we write A and B for $X_1(n)$ and $X_2(n)$, respectively.

Then what we have on the left hand side is

$$\frac{1}{n^2}\sum E\big(A_{i(1)i(2)}A_{i(2)i(3)}\cdots A_{i(k)i(k+1)}B_{i(k+1)i(k+2)}\cdots B_{i(k+l)i(1)}\big)$$
$$= \frac{1}{n^2}\sum E\big(A_{i(1)i(2)}A_{i(2)i(3)}\cdots A_{i(k)i(k+1)}\big)E\big(B_{i(k+1)i(k+2)}\cdots B_{i(k+l)i(1)}\big)$$

with summation for all indices. The matrix elements are independent and have zero expectation. Hence a term in which a matrix element appears only once among the factors must vanish. On the right hand side of (1) we have

$$\frac{1}{n}\sum E\big(A_{i(1)i(2)}A_{i(2)i(3)}\cdots A_{i(k)i(1)}\big)$$
$$\times\frac{1}{n}\sum E\big(B_{i(k+1)i(k+2)}\cdots B_{i(k+l)i(k+1)}\big)\,.$$

The two expressions are equal in many cases, in particular when k or l is odd or when both are 0. However the summation is over slightly different sequences of indices. The difference goes to 0 as $n \to \infty$. In this way, instead of the equality in (1) for a finite n, we have identical limits as $n \to \infty$. The free relation appears only in the limit; this is called the *asymptotic freeness* of $X_1(n)$ and $X_2(n)$. More generally, the asymptotic freeness of the sequence $X_1(n), X_2(n), \dots$ is formulated. The free relation is present in the random matrix context asymptotically, and this fact explains why the semicircle law is the limiting eigenvalue distribution of the random matrix $H(n)$.

Independent symmetric Gaussian matrices are asymptotically free, but there are several other interesting examples too. For instance, independent Haar distributed unitary matrices are asymptotically free (as the matrix size tends to infinity). The asymptotic freeness may serve as a bridge connecting random matrix theory with free probability theory.

In a very abstract sense, the distribution of a not neccessarily selfadjoint noncommutative random variable A is the collection of all joint moments

$$\varphi(A^{m(1)}A^{*m(2)}A^{m(3)}A^{*m(4)}\cdots)$$

of A and its adjoint A^*. A random matrix model of A is a sequence of random matrices $X(n)$ such that $X(n)$ is $n\times n$ and the joint moments of these matrices reproduce those of A as $n\to\infty$, that is,

$$\tau_n(X(n)^{m(1)}X(n)^{*m(2)}X(n)^{m(3)}\cdots)\to\varphi(A^{m(1)}A^{*m(2)}A^{m(3)}\cdots)$$

for all finite sequences $m(1),\dots,m(k)$ of nonnegative integers. It is really amazing that many important distributions appearing in free probability theory admit a suitable random matrix model. The selfadjoint Gaussian matrices with independent entries were mentioned above in connection with the Wigner theorem. The free analogue of the Poisson distribution is related to the Wishart matrix, and the circular and elliptic laws come from non-selfadjoint Gaussian matrices. These facts provide room for the interaction between random matrices and free probability. On the one hand, approximation by random matrices gives a powerful method to study free probability theory, and on the other hand techniques of free probability can be used to determine the limiting eigenvalue distribution of some random matrices, for example. It seems that bi-unitarily invariant matrix ensembles form an important class; they give the random matrix model of R-diagonal distributions. One way to define an R-diagonal noncommutative random variable is to consider its polar decomposition uh. If u is a Haar unitary and it is free from h, then uh is called R-diagonal. A bi-unitarily invariant random matrix has a polar decomposition UH in which U is a Haar distributed unitary matrix independent of H. As the matrix size grows, independence is converted into freeness according to some asymptotic freeness result.

The *empirical eigenvalue distribution* of an $n \times n$ random matrix H is the random atomic measure

$$R_H := \frac{1}{n}\big(\delta(\lambda_1) + \delta(\lambda_2) + \cdots + \delta(\lambda_n)\big)\,,$$

where $\lambda_1, \lambda_2, \dots, \lambda_n$ are the eigenvalues of H. It is a stronger form of the Wigner theorem that $R_{H(n)}$ goes to the semicircle law almost everywhere when $H(n)$ is symmetric with independent identically distributed entries. The almost sure limit of the empirical eigenvalue distribution is known to be a non-random measure in many examples, and it is often called the *density of states*. Sometimes it cannot be given explicitly, but is determined by a functional equation for the Cauchy transform or by a variational formula. Results are available for non-selfadjoint random matrices as well.

The best worked out example of symmetric random matrices is the case of independent identically distributed Gaussian entries. If the entries are Gaussian $N(0, 1/n)$ and if G is an open subset of the space of probability measures on $\mathbb{R}$ such that the semicircle law (the density of states) w is not in the closure, then we have

$$\mathbf{Prob}\big(R_{H(n)} \in G\big) \approx \exp(-n^2 C(w, G))$$

with a positive constant $C(w, G)$ depending on the distance of w from G. The bigger the distance of w from G, the larger the constant $C(w, G)$ is. Large deviation theory expresses this constant as the infimum of a so-called *rate function* I defined on the space of measures:

$$C(w, G) = \inf\{I(\mu) : \mu \in G\}\,.$$

In our example, I is a stricly convex function which is stricly positive for $\mu \neq w$. Ingredients of $I(\mu)$ are the logarithmic energy and the second moment of μ. What we are sketching now is the pioneering large deviation theorem of Ben Arous and Guionnet for symmetric Gaussian matrices. More details can be found in the next section.

0.4 Entropy and large deviations

Originally entropy was a quantity from physics. Entropy as a mathematical concept is deeply related to large deviations, although the two had independent lives for a long time. A typical large deviation result was discovered by I.N. Sanov in 1957; however, the general abstract framework of large deviations was given by S.R.S. Varadhan in 1966.

Let $\xi_1, \xi_2, \dots$ be independent standard Gaussian random variables and let G be an open set in the space $\mathcal{M}(\mathbb{R})$ of probability measures on $\mathbb{R}$ (with the weak topology). The *Sanov theorem* says that if the standard Gaussian measure ν is not in the closure of G, then

$$\mathbf{Prob}\Big(\frac{\delta(\xi_1)+\delta(\xi_2)+\cdots+\delta(\xi_n)}{n} \in G\Big) \approx \exp\big(-nC(\nu, G)\big)$$

and

$$C(\nu, G) = \inf\{I(\mu) : \mu \in G\}\,.$$

In the above case, the rate function $I(\mu)$ is the *relative entropy* (or the *Kullback-Leibler divergence*) $S(\mu, \nu)$ of μ with respect to ν. So it is also written as

$$I(\mu) = -S(\mu) + \frac{1}{2}\int x^2\, d\mu(x) + \frac{1}{2}\log(2\pi) \tag{2}$$

with the *Boltzmann-Gibbs entropy*

$$S(\mu) := -\int p(x)\log p(x)\, dx\,, \tag{3}$$

whenever μ has the density $p(x)$ and the logarithmic integral is meaningful. This rate function I is a strictly convex function such that $I(\mu) > 0$ if $\mu \neq \nu$.

The rate functions in some large deviation results are called entropy functionals. Eventually, this could be the definition of entropy. The logarithmic integral (3) of a probability distribution $p(x)$ has been used for a long time, but it was identified much later as a component of the rate function in the Sanov theorem.

The Boltzmann-Gibbs entropy $S(\mu)$ can be recovered from the asymptotics of probabilities. Let ν^n be the n-fold product of the standard Gaussian measure on $\mathbb{R}$. For each $x \in \mathbb{R}^n$ we have the discrete measure

$$\kappa_x := \frac{\delta(x_1)+\delta(x_2)+\ldots+\delta(x_n)}{n},$$

which can be used to approximate the given measure μ. The asymptotic volume of the approximating $\mathbb{R}^n$-vectors up to the first r moments is given by

$$\frac{1}{n}\log\nu^n\big(\{x \in \mathbb{R}^n : |m_k(\kappa_x) - m_k(\mu)| \le \varepsilon,\, k \le r\}\big). \tag{4}$$

A suitable limit of this as $n \to \infty, r \to \infty$ and $\varepsilon \to 0$ is exactly the above described rate function (2). Furthermore, from (4) the entropy $S(\mu)$ can be recovered. This crucial argument deduces the entropy from the asymptotics of probabilities. The extension of this argument works for multivariables, but first we consider the analogous situation in which the atomic measures κ_x are replaced by symmetric matrices.

In the large deviation result of Ben Arous and Guionnet the explicit form of the rate function I on $\mathcal{M}(\mathbb{R})$ is

$$I(\mu) = -\frac{1}{2}\iint \log|x-y|\,d\mu(x)\,d\mu(y) + \frac{1}{4}\int x^2\,d\mu(x) + \text{const.} \tag{5}$$

The above double integral is the (negative) logarithmic energy of μ, which is very familiar from potential theory.

Here we give an outline of how the rate function in (5) arises in the large deviation theorem of Ben Arous and Guionnet. In an abstract setting, a large deviation is considered for a sequence (P_n) of probability distributions, usually on a Polish space $\mathcal{X}$ in the scale (L_n); in our example $\mathcal{X} = \mathcal{M}(\mathbb{R})$ and $P_n(G) = \mathbf{Prob}(R_{H(n)} \in G)$. A standard way of proving the large deviation in this setting is to show the following equality:

$$I(x) = \sup\left[-\limsup_{n\to\infty} L_n \log P_n(G)\right] = \sup\left[-\liminf_{n\to\infty} L_n \log P_n(G)\right],$$

where the supremum is over neighborhoods G of $x \in \mathcal{X}$. This equality gives the large deviation of (P_n) with the rate function I if (P_n) satisfies an additional property of a stronger form of tightness. The scale in the Sanov large deviation is $L_n = n^{-1}$, but the scale in large deviations related to random matrices is $L_n = n^{-2}$, corresponding to the number of entries of an $n \times n$ matrix. The joint eigenvalue density of the relevant random matrix $H(n)$ is known to be

$$\frac{1}{Z_n}\exp\biggl(-\frac{n+1}{4}\sum_{i=1}^{n} x_i^2\biggr)\prod_{i<j}|x_i - x_j| \tag{6}$$

on $\mathbb{R}^n$ with the normalizing constant Z_n. This means that for a neighborhood G of $\mu \in \mathcal{M}(\mathbb{R})$ one has

$$\begin{aligned}&\mathbf{Prob}\bigl(R_{H(n)} \in G\bigr)\\ &\quad = \frac{1}{Z_n}\int\cdots\int_{\tilde{G}} \exp\biggl(\sum_{i<j}\log|x_i - x_j| - \frac{n+1}{4}\sum_{i=1}^{n} x_i^2\biggr) dx_1\cdots dx_n\,,\end{aligned}$$

where $\tilde{G} \subset \mathbb{R}^n$ is defined by

$$\tilde{G} := \left\{x \in \mathbb{R}^n : \frac{1}{n}\sum_{i=1}^{n}\delta(x_i) \in G\right\}.$$

Very roughly speaking, when G goes to a point μ, the approximation

$$\sum_{i<j} \log|x_i - x_j| - \frac{n+1}{4}\sum_{i=1}^n x_i^2 \approx n^2\left(\frac{1}{2}\iint \log|x-y|\,d\mu(x)\,d\mu(y) - \frac{1}{4}\int x^2\,d\mu(x)\right)$$

holds for $x \in \tilde{G}$, and one gets

$$-\frac{1}{n^2}\log \mathbf{Prob}\big(R_{H(n)} \in G\big) \approx -\frac{1}{2}\iint \log|x-y|\,d\mu(x)\,d\mu(y) + \frac{1}{4}\int x^2\,d\mu(x) + \frac{1}{n^2}\log Z_n\,.$$

This gives rise to the rate function (5), and the constant term there comes from $\lim_{n\to\infty} n^{-2}\log Z_n$.

Besides the symmetric Gaussian matrix, we know the exact form of the joint eigenvalue density for several other random matrices, such as the selfadjoint or non-selfadjoint Gaussian matrix, the Wishart matrix, the Haar distributed unitary matrix, and so on. The joint densities are distributed on $\mathbb{R}^n, \mathbb{C}^n, (\mathbb{R}^+)^n, \mathbb{T}^n$ depending on the type of matrices, but they have a common form which is a product of two kernels as in (6); one is the kernel of Vandermonde determinant type

$$\prod_{i<j} |x_i - x_j|^{2\beta}$$

with some constant $\beta > 0$, and the other is of the form

$$\exp\Bigg(-\sum_{i=1}^n Q_n(x_i)\Bigg)$$

with some function Q_n depending on n. Applying the method outlined above to this form of joint density, we can show the large deviations for the empirical eigenvalue distribution of random matrices as above. Corresponding to the form of joint density, the rate function is a weighted logarithmic potential as in (5) and its main term is always the logarithmic energy.

What is the free probabilistic analogue of the Boltzmann-Gibbs entropy (3) of a probability distribution μ on $\mathbb{R}$? Voiculescu answered this question by looking at the asymptotic behavior of the Bolzmann-Gibbs entropy of random matrices: The free entropy of μ should be

$$\Sigma(\mu) := \iint \log|x-y|\,d\mu(x)\,d\mu(y)\,, \tag{7}$$

that is, minus the logarithmic energy. Below we shall explain that the above double integral is the real free analogue of the Boltzmann-Gibbs entropy.

Let $\mathcal{M}$ be a von Neumann algebra with a faithful normal tracial state τ; so τ gives the expectation of elements of $\mathcal{M}$. Moreover, $m_k(a) := \tau(a^k)$ is viewed as the kth moment of a noncommutative random variable $a \in \mathcal{M}$. In order to approximate a selfadjoint a with $\|a\| \le R$ in distribution, Voiculescu suggested using symmetric matrices $A \in M_n(\mathbb{R})$; approximation means that $|\mathrm{tr}_n(A^k) - \tau(a^k)|$ is small. Hence the analogue of (4) is

$$\frac{1}{n^2} \log \nu_n\big(\big\{A \in M_n(\mathbb{R})^{sa} : \|A\| \le R,\ |\mathrm{tr}_n(A^k) - \tau(a^k)| \le \varepsilon,\ k \le r\big\}\big)\,.$$

The scaling is changed into n^2 corresponding to the higher degree of freedom, and the measure ν_n must be a measure on the space of real symmetric matrices; again the standard Gaussian measure will do. The limit as $n \to \infty$ and then $r \to \infty$, $\varepsilon \to 0$ is

$$\frac{1}{2} \iint \log|x-y|\, d\mu(x)\, d\mu(y) - \frac{1}{4}\tau(a^2) + \mathrm{const.}\,, \tag{8}$$

where μ is the probability measure on $\mathbb{R}$ which has the same moments as a: $\int x^k\, d\mu(x) = \tau(a^k)$. This limit is minus the rate function (5), and the first term gives the free entropy $\Sigma(\mu)$. Instead of $M_n(\mathbb{R})^{sa}$ one may use the space $M_n(\mathbb{C})^{sa}$ of selfadjont matrices together with the standard Gaussian measure on it.

Another analogy between the two entropies $S(\mu)$ and $\Sigma(\mu)$ is clarified by their maximization results. The entropy $S(\mu)$ can take any value in $[-\infty, +\infty]$. Instead of the value itself, rather important is the difference of $S(\mu)$ from the maximum under some constraint. For instance, under the constraint of the second moment $m_2(\mu) \le 1$, the Boltzmann-Gibbs entropy has the upper bound $S(\mu) \le \frac{1}{2}\log(2\pi e)$, and equality is attained here if and only if μ has the normal distribution $N(0,1)$. This fact is readily verified from the positivity of the relative entorpy $S(\mu, \nu)$ with $\nu = N(0,1)$. On the other hand, under the constraint of the second moment $m_2(\mu) \le 1$, the free entropy has the upper bound $\Sigma(\mu) \le -1/4$, and equality is attained if and only if μ is the semicircle law of radius 2. This maximization of $\Sigma(\mu)$ resembles that of $S(\mu)$; their maximizers are the normal law and the semicircle law, and the latter is the free analogue of the former. The Gaussian and semicircle maximizations are linked by random matrices. The symmetric random matrix with maximal Boltzmann-Gibb entropy under the constraint $\tau_n(H^2) \le 1$ is the standard Gaussian matrix, which is a random matrix model of the semicircle law.

0.5 Voiculescu's free entropy for multivariables

The free entropy as well as the Boltzmann-Gibbs entropy can be extended to multivariables. The multivariate case is slightly more complicated, but conceptually it is exactly the same. First we consider the Boltzmann-Gibbs entropy of multi-random variables (i.e. random vectors). For a random vector $(X_1, X_2, \dots, X_N)$ one has the joint distribution μ on $\mathbb{R}^N$, and the logarithmic integral (3) is meaningful and functions well whenever μ has the density $p(x)$ on $\mathbb{R}^N$. For $x = (x_1, x_2, \dots, x_n) \in (\mathbb{R}^N)^n$ let κ_x be the atomic measure on $\mathbb{R}^N$ defined analogously to the above single variable case. Now $k = (k(1), k(2), \dots, k(p))$ must be a multi-index of length $|k| := p$. For a measure μ on $\mathbb{R}^N$ whose support is in $[-R, R]^N$, we define

$$m_k(\mu) := \int x_{k(1)} x_{k(2)} \cdots x_{k(p)} \, d\mu(x)$$

and consider

$$\frac{1}{n} \log \nu^n \big(\big\{ x \in ([-R, R]^N)^n : |m_k(\kappa_x) - m_k(\mu)| \le \varepsilon, \ |k| \le r \big\} \big) , \tag{9}$$

where ν^n is the n-fold product of Gaussian measures on $\mathbb{R}^N$. The usual limit as $n \to \infty$ and then $r \to \infty$, $\varepsilon \to 0$ is

$$-S(\mu) + \frac{1}{2} \int \big(x_1^2 + x_2^2 + \cdots + x_N^2 \big) \, d\mu(x) + \text{const.}$$

Now let $(a_1, a_2, \dots, a_N)$ be an N-tuple of selfadjoint noncommutative random variables. Due to the noncommutativity we cannot have the joint distribution (as a measure); however the mixed joint moments of $(a_1, a_2, \dots, a_N)$ with respect to the tracial state τ are available and we can consider the analogue of (9). For a multi-index k we set $m_k(a_1, a_2, \dots, a_N) := \tau(a_{k(1)} a_{k(2)} \cdots a_{k(p)})$, and similarly $m_k(A_1, A_2, \dots, A_N) := \mathrm{tr}_n(A_{k(1)} A_{k(2)} \cdots A_{k(p)})$ for an N-tuple $(A_1, A_2, \dots, A_N)$ of $n \times n$ matrices. To deal with the quantity

$$\begin{aligned} \frac{1}{n^2} \log \nu_n \big(\big\{ &(A_1, A_2, \dots, A_N) \in (M_n(\mathbb{C})^{sa})^N : \|A_i\| \le R, \\ &|m_k(A_1, A_2, \dots, A_N) - m_k(a_1, a_2, \dots, a_N)| \le \varepsilon, \ |k| \le r \big\} \big) , \end{aligned} \tag{10}$$

we again put a product measure ν_n on the set of selfadjoint matrices. Since it is not known whether the limit as $n \to \infty$ of (10) exists, the $\limsup$ as $n \to \infty$ may be considered. The limit as $r \to \infty$, $\varepsilon \to 0$ of $\limsup_{n\to\infty}$ of the quantity (10) is of the form

$$\chi(a_1, a_2, \dots, a_N) + \frac{1}{2} \tau(a_1^2 + a_2^2 + \cdots + a_N^2) + \text{const.},$$

independently of the choice of $R > \|a_i\|$. This defines Voiculescu's free entropy $\chi(a_1, a_2, \dots, a_N)$. The multivariate free entropy generalizes the above free entropy $\Sigma(\mu)$ (up to an additive constant); to be more precise, the equality

$$\chi(a) = \Sigma(\mu) + \frac{1}{2} \log(2\pi) + \frac{3}{4}$$

is valid with the distribution μ of a. The term "free" has nothing to do with thermodynamics; it comes from the additivity property:

$$\chi(a_1, a_2, \dots, a_N) = \chi(a_1) + \chi(a_2) + \cdots + \chi(a_N) \tag{11}$$

when $a_1, a_2, \dots, a_N$ are in free relation with respect to the expectation τ. In this spirit the Boltzmann-Gibbs entropy must be called "independent" since it is additive if and only if $X_1, X_2, \dots, X_N$ are independent; that is, the joint distribution μ is a product measure. When the noncommutative random variables are far away from freeness (in particular, when an algebraic relation holds for $a_1, a_2, \dots, a_N$), their free entropy becomes $-\infty$. This is another reason for the terminology "free entropy". The additivity (11) is equivalent to the free relation of the a_i's when $\chi(a_i) > -\infty$, but the subadditivity

$$\chi(a_1, a_2, \dots, a_N) \le \chi(a_1) + \chi(a_2) + \cdots + \chi(a_N)$$

always holds. The free entropy $\chi(a_1, a_2, \dots, a_N)$ is upper semicontinuous in the convergence in joint moments. Furthermore, certain kinds of change of variable formulas are available. Under the constraint for $a_i = a_i^*$ that $\sum_i \tau(a_i^2)$ is fixed, $\chi(a_1, a_2, \dots, a_N)$ is maximal when (and only when) all a_i's are free and semicircular. There are possibilities to extend $\chi(a_1, a_2, \dots, a_N)$ to the case of non-selfadjoint noncommutative random variables. One possibility is to allow non-selfadjoint matrices in the definition (10), and another is to split the non-selfadjoint operators into their real and imaginary parts. The two approaches give the same result, say $\hat{\chi}(a_1, a_2, \dots, a_N)$, where the N-tuple is arbitrary and not necessarily selfadjoint. The subadditivity is still true, and $\hat{\chi}(a_1, a_2, \dots, a_N) = -\infty$ when one of the a_i's is normal, in paricular when all of them are unitaries.

For an N-tuple of unitaries $(u_1, u_2, \dots, u_N)$ the appropriate way leading to a good concept of entropy is to use unitary matrices in a definition similar to (10), and to measure the volume of the approximating unitary matrices by the Haar measure. In this way we arrive at $\chi_u(u_1, u_2, \dots, u_N)$. The free entropy of unitary variables has properties similar to the free entropy of selfadjoint ones; namely, the subadditivity and the upper semicontinuity hold, and additivity is equivalent to freeness. The three kinds of free entropies are connected under the polar decompositions $a_i = u_i h_i$ in the following way:

$$\hat{\chi}(a_1, a_2, \dots, a_N) \le \chi_u(u_1, u_2, \dots, u_N) + \chi(h_1^2, h_2^2, \dots, h_N^2) + \frac{N}{2}\Big(\log\frac{\pi}{2} + \frac{3}{2}\Big),$$

and, furthermore. equality is valid under a freeness assumption.

0.6 Operator algebras

The study of (selfadjoint) operator algebras is divided into two major categories, C^*-algebras and von Neumann algebras (i.e. W^*-algebras). *C^*-algebras* are usually introduced in an axiomatic way: A C^*-algebra is an involutive Banach algebra satisfying the C^*-norm condition $\|a^*a\| = \|a\|^2$. But any C^*-algebra is represented as a norm-closed *-algebra of bounded operators on a Hilbert space (Gelfand-Naimark representation theorem). Von Neumann algebras are included in the class of C^*-algebras; however, the ideas and methods in the two categories are very different. A commutative C^*-algebra with unit is isomorphic to $C(\Omega)$, the C^*-algebra of continuous complex functions on a compact Hausdorff space Ω with sup-norm (another Gelfand-Naimark theorem). So a general C^*-algebra is sometimes viewed as a "noncommutative topological space". On the other hand, a von Neumann algebra is a noncommutative analogue of (probability) measure spaces. In fact, a commutative von Neumann algebra with a faithful normal state is isomorphic to the space $L^\infty(\Omega,\mu)$ over a standard Borel space (Ω,μ).

According to von Neumann's reduction theory, a von Neumann algebra $\mathcal{M}$ on a separable Hilbert space is a sort of direct integral of factors:

$$\mathcal{M} = \int_\Gamma^\oplus \mathcal{M}(\gamma)\,d\nu(\gamma)\,.$$

Therefore factors are building blocks of general von Neumann algebras. Factors are classified into the types I_n $(n = 1,2,3,\dots,\infty)$, II_1, II_∞, and III. The I_n $(n<\infty)$ factor is the matrix algebra $M_n(\mathbb{C})$ and the I_∞ factor is $B(\mathcal{H})$ with $\dim\mathcal{H} = \infty$. A II_1 factor is sometimes said to have continuous dimensions because, as already mentioned in the first section, it has a normal tracial state whose values of projections in $\mathcal{M}$ are all reals in $[0,1]$. A type II_∞ factor is written as the tensor product of a II_1 factor and the I_∞ factor, and it has a normal semifinite trace. The type I factors are trivial from the operator algebra point of view. Infinite tensor products of matrix algebras with normalized traces and the group von Neumann algebras of ICC discrete groups are typical examples of type II_1 factors; the simplest construction is $\mathcal{R} := \bigotimes_{n=1}^\infty (M_2(\mathbb{C}),\mathrm{tr}_2)$, as described in the first section. All factors except type I or II are said to be of type III, and they are further classified into the types III_λ $(0\le\lambda\le 1)$; the latter subclasses were introduced by A. Connes. For $0<\lambda<1$ a typical example of a III_λ factor is the *Powers factor*

$$\mathcal{R}_\lambda := \bigotimes_{n=1}^\infty (M_2(\mathbb{C}),\omega_\lambda), \quad \text{where} \quad \omega_\lambda(\,\cdot\,) := \mathrm{tr}_2\left(\begin{bmatrix} \frac{1}{1+\lambda} & 0 \\ 0 & \frac{\lambda}{1+\lambda}\end{bmatrix}\,\cdot\,\right).$$

Furthermore, the tensor product $\mathcal{R}_\lambda\otimes\mathcal{R}_\mu$ of two Powers factors with $\log\lambda/\log\mu\notin\mathbb{Q}$ becomes a III_1 factor. The Tomita-Takesaki theory is fundamental in the structure analysis of type III factors.

A von Neumann algebra $\mathcal{M}$ (on a separable Hilbert space) is said to be *approximately finite dimensional* (AFD), or sometimes *hyperfinite*, if it is generated by an increasing sequence of finite-dimensional subalgebras. On the other hand, $\mathcal{M}\subset B(\mathcal{H})$ is said to be *injective* if there exists a conditional expectation (i.e. norm one projection) from $B(\mathcal{H})$ onto $\mathcal{M}$. The epoch-making result of Connes in 1976 shows that the injectivity of $\mathcal{M}$ is equivalent to the AFD of $\mathcal{M}$, and there is a unique injective factor for each type II_1, II_∞, III_λ $(0<\lambda<1)$. The fact that the above $\mathcal{R}$ is a unique hyperfinite type II_1 factor had been proved long ago by Murray and von Neumann, and the uniqueness of the injective III_1 factor was later proved by U. Haagerup in 1987. Furthermore, all AFD factors of type III are completely classified in terms of the flow of weights introduced by Connes and Takesaki. In this way, $\mathcal{R}\otimes B(\mathcal{H})$, the Powers factor $\mathcal{R}_\lambda$ and the above $\mathcal{R}_\lambda\otimes\mathcal{R}_\mu$ are unique AFD factors of type II_∞, III_λ $(0<\lambda<1)$ and III_1, respectively. There are many AFD III_0 factors; all of them are constructed as *Krieger factors* $L^\infty(\Omega,\mu)\rtimes_T\mathbb{Z}$, where T is an ergodic transformation on a standard Borel space (Ω,μ). The so-called measure space-group construction is to make the crossed product von Neumann algebra $L^\infty(\Omega,\mu)\rtimes_\alpha G$, where α is an action of a group G on (Ω,μ). According to J. Feldman and C.C. Moore, there is a more general construction of von Neumann algebras from a measurable equivalence relation on (Ω,μ) with countable orbits, and $L^\infty(\Omega,\mu)$ is a *Cartan subalgebra* of the constructed von Neumann algebra. Moreover, the existence of a Cartan subalgebra is sufficient for a von Neumann algebra to enjoy such a measure space construction. Any injective factor has a Cartan subalgebra as a consequence of the above classification result. The notion of amenability can be defined for a measurable relation, and a profound result of Connes, Feldman and Weiss in 1982 says that a measurable relation is amenable if and only if it is generated by a single measurable transformation. A consequence of this is that any two Cartan subalgebras of an injective factor $\mathcal{M}$ are conjugate by an automorphism of $\mathcal{M}$.

After the AFD factors were classified as above, a huge class of non-AFD factors remained unclassified. Since any type III factor can be canonically decomposed into the crossed product $\mathcal{N}\rtimes_\theta\mathbb{R}$ with a II_∞ von Neumann algebra $\mathcal{N}$ according to the Takesaki duality, and since type II_∞ can be somehow reduced to type II_1, we may say that the problem returns to the type II_1 theory in some sense. A type II_1 von Neumann algebra with a faithful normal tracial state is a noncommutative probability space most appropriate to free probability theory. A typical example of non-AFD factors is the free group factor $\mathcal{L}(\mathbf{F}_n)$, and it is generated by a free family of noncommutative random variables. This is the reason why some techniques from free probability and free entropy are so useful in analyzing free group factors. There has been much progress in the theory of free group factors; for instance, the non-existence of a Cartan subalgebra in $\mathcal{L}(\mathbf{F}_n)$, proved by Voiculescu, is remarkable because it means the impossibility of the measure space construction as above for $\mathcal{L}(\mathbf{F}_n)$.

The K-theory of C^*-algebras yields algebraic invariants to study isomorphism problems of C^*-algebras. Two abelian groups $K_0(\mathcal{A})$ and $K_1(\mathcal{A})$ are associated with each C^*-algebra $\mathcal{A}$. The construction of $K_0(\mathcal{A})$ follows the old idea of Murray and von Neumann for the classification of von Neumann factors. Since a C^*-algebra may not have enough projections, we pass to the algebra $\bigcup_{n=1}^\infty M_n(\mathcal{A})$. An equivalence relation is defined on the collection of all projections in $\bigcup_{n=1}^\infty M_n(\mathcal{A})$, whose equivalence classes form an abelian semigroup $K_0(\mathcal{A})^+$ with zero element. Then the Grothendieck group of $K_0(\mathcal{A})^+$ is the K_0-group $K_0(\mathcal{A})$. On the other hand, the K_1-group $K_1(\mathcal{A})$ is defined as the inductive limit of the quotient groups $\mathcal{U}(M_n(\mathcal{A}))/\mathcal{U}_0(M_n(\mathcal{A}))$, where $\mathcal{U}(M_n(\mathcal{A}))$ is the group of unitaries in $M_n(\mathcal{A})$ and $\mathcal{U}_0(M_n(\mathcal{A}))$ is the connected component of the identity.

The first major contribution toward classification of C^*-algebras was made by

G.A. Elliott in 1976. He showed that the AF C^*-algebras are completely classified by the ordered group $(K_0(\mathcal{A}), K_0(\mathcal{A})^+)$, the *dimension group*. A C^*-algebra $\mathcal{A}$ is said to be *nuclear* if the minimal C^*-norm is a unique C^*-norm on the algebraic tensor product of $\mathcal{A}$ with any C^*-algebra $\mathcal{B}$. The class of nuclear C^*-algebras is in some sense a C^*-counterpart of the class of injective von Neumann algebras. Many characterizations of nuclear C^*-algebras are known; for example, $\mathcal{A}$ is nuclear if and only if $\pi(\mathcal{A})''$ is injective for any representation of $\mathcal{A}$. The nuclear C^*-algebras are closed under basic operations such as inductive limits, quotients by closed ideals, tensor products and crossed products by actions of amenable groups. In particular, AF algebras are nuclear, and the amenability of a discrete group G is equivalent to the nuclearity of the reduced C^*-algebra $C^*_r(G)$ (similarly to the hyperfiniteness of the group von Neumann algebra $\mathcal{L}(G)$).

Exact C^*-algebras form an important class. A C^*-algebra is said to be *exact* if the sequence of the minimal C^*-tensor products

$$0 \to \mathcal{A} \otimes \mathcal{J} \to \mathcal{A} \otimes \mathcal{B} \to \mathcal{A} \otimes (\mathcal{B}/\mathcal{J}) \to 0$$

is exact whenever $\mathcal{J}$ is a closed ideal of an arbitrary C^*-algebra $\mathcal{B}$. This class includes the nuclear C^*-algebras. For example, the reduced C^*-algebra $C^*_r(\mathbf{F}_n)$ of the free group $\mathbf{F}_n$ is not nuclear but exact. Indeed, it is open whether $C^*_r(G)$ is exact for every countable discrete group G. A C^*-subalgebra of an exact C^*-algebra is exact, and the exact C^*-algebras are closed under the operations of inductive limits, minimal tensor poducts and quotients. It seems that the role of exact C^*-algebras has increased in recent development of C^*-algebra theory since the appearance of the work of Kirchberg.

The group C^*-algebra $C^*_r(\mathbf{F}_n)$ is out of the scope of well-established algebraic invariants; nevertheless, K-theory can detect n from the reduced C^*-algebra of $\mathbf{F}_n$. In 1982 Pimsner and Voiculescu computed the K-groups

$$K_0(C^*_r(\mathbf{F}_n)) = \mathbb{Z} \quad \text{and} \quad K_1(C^*_r(\mathbf{F}_n)) = \mathbb{Z}^n\,,$$

and this computation proves that $C^*_r(\mathbf{F}_n)$ is not isomorphic to $C^*_r(\mathbf{F}_m)$ for $n \neq m$. The isomorphism question of whether $\mathcal{L}(\mathbf{F}_n) \not\cong \mathcal{L}(\mathbf{F}_m)$ if $n \neq m$ is still open, and a possible approach uses the free entropy dimension, which is a candidate for a reasonable entropic invariant.

Let $a_1, \dots, a_N \in \mathcal{M}^{sa}$ and assume that $S_1, \dots, S_N \in \mathcal{M}^{sa}$ is a free family of semicircular elements which are in free relation to $\{a_1, \dots, a_N\}$. Then the *free entropy dimension* $\delta(a_1, \dots, a_N)$ is defined in terms of the multivariate free entropy by

$$\delta(a_1, \dots, a_N) := N + \limsup_{\varepsilon \to +0} \frac{\chi(a_1 + \varepsilon S_1, \dots, a_N + \varepsilon S_N)}{|\log \varepsilon|}\,.$$

Note that the above $S_1, \dots, S_N$ always exist when we enlarge $\mathcal{M}$ by taking a free product with another von Neumann algebra, and that the joint distribution of

$a_1 + \varepsilon S_1, \ldots, a_N + \varepsilon S_N$ is independent of the choice of $S_1, \ldots, S_N$, so $\delta(a_1, \ldots, a_N)$ is well-defined.

Let $g_1, g_2, \ldots, g_N$ be generators of $\mathbf{F}_N$. Then $a_i = L_{g_i} + L_{g_i^{-1}}$ form a free family of semicircular noncommutative random variables, and $\delta(a_1, \ldots, a_k) = k$ for $k \leq N$. Proving the lower semicontinuity of the free entropy dimension would be very exciting, for the following reason. If $a_1, a_2, \ldots, a_N$ are in free relation and $b_i \in \{a_1, a_2, \ldots, a_N\}''$, then

$$\delta(a_1, a_2, \ldots, a_N) \leq \delta(a_1, a_2, \ldots, a_N, b_1, \ldots, b_M) \tag{12}$$

with equality when the b_i's are polynomials of the a_j's. This is a proven fact. If the lower semicontinuity held, we would have equality in (12) without any further hypothesis on the b_i's. The isomorphism $\mathcal{L}(\mathbf{F}_N) \cong \mathcal{L}(\mathbf{F}_M)$ would imply that in this von Neumann algebra $\mathcal{M}$ there exist two systems $a_1, \ldots, a_N$ and $b_1, \ldots, b_M$ of generators, each of which consists of free semicircular variables. Hence, the above equality in (12) gives

$$N = \delta(a_1, \ldots, a_N) = \delta(a_1, \ldots, a_N, b_1, \ldots, b_M) = \delta(b_1, \ldots, b_M) = M\,.$$

In this way, the lower semicontinuity of $\delta(a_1, \ldots, a_N)$ would imply the solution of the isomorphism problem.

In the parametrization of the von Neumann algebras $\mathcal{L}(\mathbf{F}_n)$ the integer n is a discrete parameter when free group factors are considered. However, in the work of K. Dykema and F. Rădulescu a continuous interpolation $\mathcal{L}(\mathbf{F}_r)$ appears, where r is real and $r > 1$. Those are the so-called *interpolated free group factors*. It turned out that they are either isomorphic for all parameter values or non-isomorphic for any two different values of r. So one of the two extreme cases holds true. However, the stable isomorphism $\mathcal{L}(\mathbf{F}_r) \otimes B(\mathcal{H}) \cong \mathcal{L}(\mathbf{F}_s) \otimes B(\mathcal{H})$ is known.

The existence of the interpolation of the free group factors may suggest that the $\mathcal{L}(\mathbf{F}_n)$ are all isomorphic, contrary to the indication from the free entropy dimension.