Entropy and Transport for Non-commutative Laws

David A. Jekel

University of California, Los Angeles

Oberwolfach March 5, 2020

Outline:

- Non-commutative laws.
- Microstates free entropy χ.
- Change of variables for χ.
- A non-commutative Bézout bound?
- Non-commutative transport of measure.
- Trace polynomials.

Non-commutative laws

Classical theory:

The law (probability distribution) of a random tuple $(X_1, \ldots, X_n)$ is equivalent to a probability measure on $[-R, R]^n$.

Probability measures on $[-R, R]^n$ can be identified with linear functionals $\lambda : \mathbb{C}[x_1, \ldots, x_n] \to \mathbb{C}$ satisfying

1. $\lambda(1) = 1$.
2. $\lambda(\overline{p(x)}p(x)) \geq 0$ for all p.
3. $|\lambda(x_{i_1} \ldots x_{i_\ell})| \leq R^\ell$ for all $\ell \in \mathbb{N}$ and $i_1, \ldots, i_\ell \in \{1, \ldots, n\}$.

Non-commutative laws

Definition

Let $\mathbb{C}\langle x_1, \ldots, x_n \rangle$ be the $*$-algebra of non-commutative polynomials in indeterminates x_1, ..., x_n with $x_i = x_i^*$.

A (tracial) *non-commutative law* (of an n-tuple, with radius bound R) is a linear functional $\lambda : \mathbb{C}\langle x_1, \ldots, x_n \rangle \to \mathbb{C}$ satisfying

1. Unital: $\lambda(1) = 1$.
2. Positive: $\lambda(p(x)^* p(x)) \geq 0$ for all p.
3. Radius bounds: $|\lambda(x_{i_1} \ldots x_{i_\ell})| \leq R^\ell$ for all $\ell \in \mathbb{N}$ and i_1, ..., $i_\ell \in \{1, \ldots, n\}$.
4. Tracial: $\lambda(p(x)q(x)) = \lambda(q(x)p(x))$.

The space of these laws is denoted $\Sigma_{n,R}$, and it is equipped with the weak-$*$ topology (pointwise convergence as functions on $\mathbb{C}\langle x_1, \ldots, x_d \rangle$).

Approximation by matrices

Observation/Definition

Let $\mathbf{A} = (A_1, \dots, A_n)$ be a d-tuple of $k \times k$ self-adjoint matrices. Then the linear functional $\lambda_{\mathbf{A}} : \mathbb{C}\langle x_1, \dots, x_n \rangle \to \mathbb{C}$ given by

$$\lambda_{\mathbf{A}}(p) = \frac{1}{k} \operatorname{Tr}(p(A_1, \dots, A_n))$$

is a non-commutative law, which we call the *law of* $\mathbf{A}$. Same is true if $\mathbf{A}$ is a tuple of self-adjoint operators in a tracial von Neumann algebra.

We think of the laws of tuples in $M_k(\mathbb{C})_{sa}^n$ (where "sa" stands for "self-adjoint") as the non-commutative analogue of probability measures on $[-R, R]^n$ given as $(1/k) \sum_{i=1}^k \delta_{\xi_i}$ for vectors $\xi_1, \dots, \xi_k$.

Approximation by matrices

Classically, we know that any measure on $[-R, R]^n$ can be approximated in the weak-$*$ topology by linear combinations of delta masses. This can be equivalently stated in terms of approximating certain representations of $\mathbb{C}[x_1, \dots, x_n]$ by finite-dimensional representations.

The *Connes embedding problem* is the analogous question for $\Sigma_{n,R}$: Are the non-commutative laws that arise from matrix tuples dense in $\Sigma_{n,R}$? This has an equivalent formulation in terms of approximating $*$-representations from $\mathbb{C}\langle x_1, \dots, x_n \rangle$ in a *tracial von Neumann algebra* by $*$-representations in $M_k(\mathbb{C})$.

Microstates free entropy

Voiculescu's *microstates free entropy* asks a related question: Instead of asking *whether* a non-commutative law λ can be simulated by matrices, we ask *how many* matrices there are which simulate it (how many in the sense of asymptotic volume).

Definition (Voiculescu)

For a neighborhood $\mathcal{U}$ of λ in $\Sigma_{n,R}$, we define the *microstate space*

$$\Gamma^{(k)}(\mathcal{U}) = \{\mathbf{A} \in M_k(\mathbb{C})_{sa}^n : \lambda_{\mathbf{A}} \in \mathcal{U}\}.$$

Then we define

$$\chi(\lambda) = \inf_{\mathcal{U} \ni \lambda} \limsup_{k \to \infty} \frac{1}{k^2} \log(\operatorname{vol} \Gamma^{(k)}(\mathcal{U}))$$

(with appropriate normalization of the volume, depending on k).

It's a fact that $\chi(\lambda)$ is independent of R so long as $\lambda \in \Sigma_{n,R}$.

Microstates free entropy — remarks and questions

This is really an analogue of the classical entropy of a probability measure $h(\mu) = -\int \rho \log \rho$ where ρ is the density (and $-\infty$ if there is no density) because the classical entropy can be characterized using microstates in a similar way. We replace $\Gamma^{(k)}(\mathcal{U})$ by the set of vectors $(\xi_1, \dots, \xi_k) \in \mathbb{R}^{n \times k}$ such that $(1/k) \sum_{i=1}^k \delta_{\xi_i}$ is in $\mathcal{U}$, and we take the limit of $(1/k)$ times the log of the volume.

For free entropy, unlike the classical case, we do not know whether using $\liminf$ instead of $\limsup$ in the definition would make a difference. We would like to understand the asymptotic behavior of the volume better, and even more, we want to understand the geometry of these non-commutative microstates spaces.

Microstates free entropy — remarks and questions

To put this in terms of real algebraic geometry, fix a non-commutative law λ, let $p_1, \ldots, p_m \in \mathbb{C}\langle x_1, \ldots, x_n \rangle$ and $\epsilon_1, \ldots, \epsilon_m > 0$. Consider the set

$$\mathcal{U} = \{\mu : |\mu(p_j) - \lambda(p_j)| < \epsilon_j \text{ for } j = 1, \ldots, m\}.$$

Then the microstate space would be

$$\Gamma^{(k)}(\mathcal{U}) = \{(A_1, \ldots, A_n) \in M_k(\mathbb{C})_{sa}^n : |(1/k) \operatorname{Tr}(p_j(A_1, \ldots, A_n)) - \lambda(p_j)| < \epsilon_j\}$$

which is a semi-algebraic set (where the coordinates would be given by the real/imaginary parts of the entries of the matrices).

One could ask, e.g., how many connected components does this set have? Also, can we characterize the non-commutative polynomials p such that $p(\mathbf{A})$ is self-adjoint and positive semi-definite for all $\mathbf{A} \in \Gamma^{(k)}(\mathcal{U})$ (either for a fixed k or for all k)?

Change of variables for free entropy

Definition (Push-forward)

Let $\lambda \in \Sigma_{n,R}$. Let $\mathbf{f} = (f_1, \dots, f_n)$ be an n-tuple of self-adjoint polynomials in $\mathbb{C}\langle x_1, \dots, x_n \rangle$. We can define $\mathbf{f}_* \lambda$ by the formula

$$\mathbf{f}_* \lambda(p) = \lambda(p \circ \mathbf{f}).$$

- Note $\mathbf{f}_* \lambda$ will be in $\Sigma_{n,R'}$ for a large enough R'.
- The definition can be generalized to replace $\mathbf{f}$ by a non-commutative convergent power series, or more generally "non-commutative C^k functions."
- Note also that $\mathbf{f}_* \lambda_{\mathbf{A}} = \lambda_{\mathbf{f}(\mathbf{A})}$ for a matrix tuple $\mathbf{A}$ (and this is also true when $\mathbf{A}$ is a tuple of self-adjoint operators in a tracial von Neumann algebra).

Change of variables for free entropy

The following theorem is analogous to the classical change of variables for entropy. We will not explain the precise definitions and hypotheses involved here, but the non-commutative derivatives will be discussed in Ian Charlesworth's talk.

Theorem (Voiculescu)

Suppose that $\mathbf{f}$ *is an n-tuple of "non-commutative functions" which is sufficiently smooth and has a sufficiently smooth inverse function. Then*

$$\chi(\mathbf{f}_* \lambda) = \chi(\lambda) + (\mathrm{Tr} \otimes \lambda \otimes \lambda)[\log(D\mathbf{f})],$$

where "$(\mathrm{Tr} \otimes \lambda \otimes \lambda)[\log(D\mathbf{f})]$" *is a non-commutative version of the log-determinant of the Jacobian matrix of a function.*

Change of variables — remarks and questions

The idea of Voiculescu's proof is that $\mathbf{f}$ should map microstate spaces for λ into microstate spaces for $\mathbf{f}_*\lambda$. The volume of the microstate spaces can be computed using the classical change of variables formula. The classical log-determinant of the transformation $\mathbf{f}$ acting on $M_k(\mathbb{C})_{sa}^n$ can be well-approximated by the constant $(\operatorname{Tr} \otimes \lambda \otimes \lambda)[\log(D\mathbf{f})]$ times k^2.

It turns out that for $n = 1$, the free entropy χ has an explicit formula, and one can show that the change of variables holds *without assuming that f is invertible*, which yields a stronger result than for classical entropy. It would be quite interesting if we could remove the assumption that $\mathbf{f}$ is invertible in the free setting with $n \geq 2$.

References — free entropy

1. D.-V. Voiculescu, "The analogues of entropy and of Fisher's information measure in free probability" series:
   - I, Communications in Mathematical Physics, 1993;
   - II, Inventiones Mathematicae, 1994;
   - V, Inventiones Mathematics, 1998.
2. D.-V. Voiculescu, "Free entropy" (survey), Bulletin of the London Mathematical Society, 1992.

Change of variables — remarks and questions

The proof sketched above should work if $\mathbf{f}$ is only locally invertible and surjective rather than globally invertible, provided that the fibers of the map $\mathbf{f}$ are not too large. For this purpose, it would be sufficient that for any $\delta > 0$, the fibers of $\mathbf{f}$ have cardinality less than $e^{-k^2\delta}$ for large enough k. Then the over/under-counting that we do when computing the volume by the classical change of variables formula would become irrelevant when we evaluate $(1/k^2)$ times the log of the volume and then send $k \to \infty$.

Question (Shlyakhtenko)

Let $\mathbf{f}$ be an n-tuple of self-adjoint non-commutative polynomial in n self-adjoint variables. Let $\mathbf{B}$ be an n-tuple of $k \times k$ self-adjoint matrices. Then how many solutions can there be to the equation $\mathbf{f}(\mathbf{A}) = \mathbf{B}$ for $\mathbf{A} \in M_k(\mathbb{C})_{sa}^n$?

A non-commutative Bézout bound?

Theorem (Classical Bézout bound)

Let $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^n$ be a polynomial transformation of degree d. Then each fiber of $\mathbf{f}$ is either infinite or has cardinality bounded by d^n.

In particular, if $\mathbf{f} : M_k(\mathbb{C})_{sa}^n \to M_k(\mathbb{C})_{sa}^n$ is a non-commutative polynomial transformation of degree d, then the finite fibers have cardinality at most d^{nk^2} since $\dim_{\mathbb{R}} M_k(\mathbb{C})_{sa} = k^2$.

However, if we use the non-commutative structure, we might hope to do better than that. In particular, one might optimistically conjecture that it is bounded by d^{nk} rather than d^{nk^2}.

Such optimism is informed by the following result for $n = 1$:

A non-commutative Bézout bound?

Observation (Shlyakhtenko)

Let $f \in \mathbb{C}[x]$. If B is a self-adjoint matrix with distinct eigenvalues, then the number of self-adjoint matrices satisfying $f(A) = B$ is at most d^k.

Proof.

Note that if v is an eigenvector of A with eigenvalue λ, then v is also an eigenvector of B with eigenvalue $f(\lambda)$. Since B has distinct eigenvalues, every eigenvector of B must in fact be an eigenvector of A, and A must have distinct eigenvalues.

Let $\beta_1, \ldots, \beta_k$ and $v_1, \ldots, v_k$ be the eigenvalues and eigenvectors of B. Then A must have the same eigenvectors with eigenvalues $\alpha_1, \ldots, \alpha_k$ satisfying $f(\alpha_j) = \beta_j$. There are at most d choices of α_j for each j, hence the number of solutions is at most d^k. $\square$

Non-commutative transport of measure

Question

Let μ and ν be non-commutative laws. Under what conditions does there exist an invertible non-commutative polynomial map (or more general type of non-commutative function) $\mathbf{f}$ such that $\mathbf{f}_* \mu = \nu$?

Recall that if μ and ν are classical Borel probability measures with no atoms (on any Polish space), then there exists an invertible measurable transformation $\mathbf{f}$ such that $\mathbf{f}_* \mu = \nu$.

The situation in the non-commutative case is wildly different: Having such a transport function $\mathbf{f}$ from μ to ν would imply that the associated tracial von Neumann algebras are isomorphic. But we know that there are uncountably many non-isomorphic separable tracial von Neumann algebras with "no atoms" (McDuff), and there is no separable tracial von Neumann algebra that contains an isomorphic copy of all others (Ozawa).

Non-commutative transport of measure

Free Gibbs laws from a convex potential behave much better. These laws can be described through random matrix theory.

Let $V^{(k)}(\mathbf{A}) = (1/k)\,\mathrm{Tr}(f(\mathbf{A}))$, where $f(\mathbf{A})$ is a self-adjoint non-commutative polynomial or rational function that is well-defined for all self-adjoint tuples $\mathbf{A}$, for instance.

$$f(\mathbf{A}) = \sum_{j=1}^{n} A_j^2 + (A_1 - i)^{-1}(A_2 + i)^{-1} + (A_2 - i)^{-1}(A_1 + i)^{-1} + A_3^4.$$

Consider the probability measure on $M_k(\mathbb{C})_{sa}^n$ given by

$$d\mu^{(k)}(\mathbf{A}) = C_k e^{-k^2 V^{(k)}(\mathbf{A})}\, d\mathbf{A}.$$

(Assume that $e^{-k^2 V^{(k)}}$ is integrable.)

Non-commutative transport of measure

Denote $\|\mathbf{A}\|_2 = ((1/k)\sum_{j=1}^n \operatorname{Tr}(A_j^2))^{1/2}$.

Theorem (by Guionnet, Maurel-Segala, Shlyakhtenko, Dabrowksi, J.)

Let $0 < c < C$ and suppose that $V^{(k)}$ is as above and also satisfies $V^{(k)}(\mathbf{A}) - c\|\mathbf{A}\|_2^2$ is convex and $V^{(k)}(\mathbf{A}) - C\|\mathbf{A}\|_2^2$ is concave. Let $\mathbf{X}^{(k)}$ be a random tuple of matrices chosen according to the measure $\mu^{(k)}$. Then

1. *There exists a non-commutative law λ such that the (random) non-commutative law $\lambda_{\mathbf{X}^{(k)}}$ almost surely converges in $\Sigma_{d,R}$ to λ. [1], [2], [5].*
2. *The normalized classical entropy of $\mu^{(k)}$ converges to $\chi(\lambda)$. [5], [6].*
3. *There exist classical functions $\mathbf{f}^{(k)} : M_k(\mathbb{C})_{sa}^n \to M_k(\mathbb{C})_{sa}^n$ which transport $\mu^{(k)}$ to the Gaussian measure $\sigma^{(k)}$ on $M_k(\mathbb{C})_{sa}^n$, and such that $\mathbf{f}^{(k)}$ "behaves asymptotically like" some non-commutative function $\mathbf{f}$ which transports λ to the "free Gaussian" non-commutative law σ. [3], [4], [6].*

Non-commutative transport of measure

Theorem (J. 2019)

Under the hypotheses above, the transport function $\mathbf{f}$ *can be chosen to be triangular in the sense that*

$$\mathbf{f}(x_1, \dots, x_n) = (f_1(x_1), f_2(x_1, x_2), \dots, f(x_1, \dots, x_d)).$$

- The existence of triangular transport in the classical case is easy if there are nice conditional distributions.
- If the $\mathbf{f}$ is smooth enough, then the triangular transport provides a nice way to compute the entropy.
- The above theorem has interesting consequences for von Neumann algebras. [6]

References — Random matrices from convex potentials

1. A. Guionnet and E. Maurel-Segala, "Combinatorial apsects of random matrix models," ALEA, 2006.
2. A. Guionnet and D. Shlyakhtenko, "Free diffusions and random matrix models with strictly convex interaction," Geometric and Functional Analysis, 2009.
3. A. Guionnet and D. Shlyakhtenko, "Free monotone transport," Inventiones Mathematicae, 2014.
4. Y. Dabrowski, A. Guionnet, and D. Shlyakhtenko, "Free transport for convex potentials," arXiv, 2016.
5. D. Jekel. "An elementary approach to free entropy theory for convex potentials." To appear in Analysis & PDE Journal.
6. D. Jekel. "Conditional expectation, entropy, and transport for convex Gibbs laws in free probability." Submitted to IMRN.

Trace Polynomials

The theorem is proved by using some well-known PDE techniques to construct the transport of $\mathbf{f}^{(k)}$, then showing that $\mathbf{f}^{(k)}$ asymptotically behaves like some non-commutative function (which requires work!).

A crucial role is played by *trace polynomials*, which are a certain generalization of non-commutative polynomials.

In the remainder of the talk, I will persuade you to study trace polynomials.

Trace polynomials

Informal definition

Trace polynomials are linear combinations of terms of the form $f_0\tau(f_1)\dots\tau(f_\ell)$ where

$$f_0, f_1, \dots, f_\ell \in \mathbb{C}\langle x_1, \dots, x_n \rangle,$$

τ is a formal symbol representing "the trace," and where we impose the relation that $\tau(x_{i_1} \dots x_{i_\ell}) = \tau(x_{i_\ell} x_1 \dots x_{i_{\ell-1}})$, and that $\tau(f_j)$ commutes with everything. Trace polynomials can be evaluated on tuples of $k \times k$ self-adjoint matrices by substituting $(1/k)\operatorname{Tr}$ for the formal symbol τ, that is,

$$[f_0\tau(f_1)\dots\tau(f_\ell)](\mathbf{A}) = f_0(\mathbf{A})\frac{1}{k}\operatorname{Tr}(f_1(\mathbf{A}))\dots\frac{1}{k}\operatorname{Tr}(f_\ell(\mathbf{A}))$$

Trace polynomials can be added, multiplied, and composed in a natural way.

Trace polynomials

Formal definition

Let V_d be the vector space of non-commutative polynomials of degree d in variables $x_1, \ldots, x_n$, quotiented by the relation of cyclic symmetry, and let $V = \bigoplus_{d=0}^{\infty} V_d$. Let A be the (symmetric) tensor algebra of V. Define TrP_n to be $\mathbb{C}\langle x_1, \ldots, x_n \rangle \otimes A$.

History

Trace polynomials were studied in an algebraic setting by Razmyslov (1974 and 1987). Their use in free probability is due largely to Cébron (2013), Driver–Hall–Kemp (2013), Dabrowski–Guionnet–Shylaktenko (2016), and is a current area of progress.

Trace polynomials — the one variable case

Let $\mathrm{TrP}_n|_{M_k(\mathbb{C})_{sa}^n}$ denote the functions on $M_k(\mathbb{C})_{sa}^n$ obtained as the evaluation of trace polynomials.

For $n = 1$, the algebra $\mathrm{TrP}_1|_{M_k(\mathbb{C})_{sa}}$ is the same as the algebra of symmetric polynomials in the eigenvalues of the matrix A.

Trace polynomials and the Laplacian

Let f be a trace polynomial. Let $f^{(k)}$ be the evaluation of f on $M_k(\mathbb{C})_{sa}^n$.

1. The Laplacian $\Delta f^{(k)}$ is another trace polynomial function.
2. In fact, there exist trace polynomials g and h (that can be computed combinatorially from f) such that $(1/k)\Delta f^{(k)} = g^{(k)} + (1/k^2)h^{(k)}$.
3. However, if f is a non-commutative polynomial, then $\Delta f^{(k)}$ is usually *not* a non-commutative polynomial.
4. Since the Laplacian respects trace polynomials, there is a precise asymptotic formula for $e^{(t/k)\Delta} f^{(k)}$ in powers of $1/k^2$.
5. Looking at this heat semigroup on trace polynomials, we obtain the asymptotic formula for the moments of Gaussian random matrix tuples which is well known in random matrix theory.

Trace polynomials — Questions

- Is there an interesting algebraic geometry for sets defined by trace polynomials?
- Study optimization of trace polynomials over x in an operator norm ball. This relates to model theory of operator algebras. Also, optimize over a closed microstate space.
- Can we estimate the size of fibers of trace polynomial maps $M_k(\mathbb{C})_{sa}^n \to M_k(\mathbb{C})_{sa}^n$?
- In what ways do trace polynomials behave differently than non-commutative polynomials?
- Find a nice formula for the dimension of the space of trace polynomials of degree $\leq d$. Easy?
- For each k, find a nice formula for the dimension of the space of trace polynomial functions on $M_k(\mathbb{C})_{sa}^n$ of degree $\leq d$.

References — trace polynomials

1. Y. P. Razmyslov, "Trace identities of full matrix algebras over a field of characteristic zero," Math. of USSR-Izvestiya, 1974.
2. C. Procesi, "The invariant theory of $n \times n$ matrices," Advances in Mathematics, 1976.
3. Y. P. Razmyslov, "Trace identities and centrl polynomials in the matrix superalgebras $M_{n,k}$, Math. of USSR-Sbornik, 1987.
4. G. Cébron, "Free convolution powers and the free Hall transform," Journal of Functional Analysis, 2013.
5. B. K. Driver, B. C. Hall, and T. Kemp, "The large-N limit of the Segal-Bargmann transform on U_N," Journal of Functional Analysis, 2013.
6. Y. Dabrowski, A. Guionnet, and D. Shlyakhtenko, "Free transport for convex potentials," arXiv, 2016.
7. D. Jekel, "Conditional expectation, entropy, and transport for convex Gibbs laws in free probability," 2019.