Free Entropy for Free Gibbs Laws Given by Convex Potentials

David A. Jekel

University of California, Los Angeles

Berkeley Free Probability Seminar

Motivation

Motivation

We will discuss Voiculescu's *free entropy* of a *non-commutative law* μ *of an m-tuple*. This is an analogue in *free probability theory* of the *continuous entropy* of a probability measure.

Motivation

We will discuss Voiculescu's *free entropy* of a *non-commutative law* μ *of an m-tuple*. This is an analogue in *free probability theory* of the *continuous entropy* of a probability measure.

Voiculescu defined two types of free entropy, $\chi(\mu)$ and $\chi^*(\mu)$. They both measure the "regularity" of the law μ.

Motivation

We will discuss Voiculescu's *free entropy* of a *non-commutative law* μ *of an m-tuple*. This is an analogue in *free probability theory* of the *continuous entropy* of a probability measure.

Voiculescu defined two types of free entropy, $\chi(\mu)$ and $\chi^*(\mu)$. They both measure the "regularity" of the law μ.

They are based on two different viewpoints for classical entropy: χ is based on the microstates interpretation of entropy and is defined by "counting" matrix approximations to μ, while χ^* is defined in terms of free Fisher information, which describes how μ interacts with derivatives.

Motivation

Problem

When does $\chi = \chi^*$?

Motivation

Problem

When does $\chi = \chi^*$?

Theorem (Biane, Capitaine, Guionnet 2003)

$\chi(\mu) \leq \chi^*(\mu)$.

Motivation

Problem

When does $\chi = \chi^*$?

Theorem (Biane, Capitaine, Guionnet 2003)

$\chi(\mu) \leq \chi^*(\mu)$.

Until recently, the problem was unresolved even for convex Gibbs laws μ. These Gibbs laws are the free analogue of $(1/Z)e^{-V(x)}\,dx$, and are the best understood non-commutative laws.

Motivation

Problem

When does $\chi = \chi^*$?

Theorem (Biane, Capitaine, Guionnet 2003)

$\chi(\mu) \leq \chi^*(\mu)$.

Until recently, the problem was unresolved even for convex Gibbs laws μ. These Gibbs laws are the free analogue of $(1/Z)e^{-V(x)}\,dx$, and are the best understood non-commutative laws.

Theorem (Dabrowski 2017, J. 2018)

If μ is a free Gibbs state given by a nice enough convex potential V, then $\chi(\mu) = \chi^*(\mu)$.

Background: Free Probability

What is non-commutative probability?

classical	non-commutative
$L^\infty(\Omega, P)$	W^*-algebra M
expectation E	trace τ
bdd. real rand. var. X	self-adjoint $X \in M$
law of X	spectral distribution of X w.r.t. τ

What is free probability?

We replace classical independence by *free independence*.

What is free probability?

We replace classical independence by *free independence*.

Definition by Example

For groups G_1 and G_2, the algebras $L(G_1)$ and $L(G_2)$ are freely independent in $(L(G_1 * G_2), \tau)$.

What is free probability?

We replace classical independence by *free independence*.

Definition by Example

For groups G_1 and G_2, the algebras $L(G_1)$ and $L(G_2)$ are freely independent in $(L(G_1 * G_2), \tau)$.

Free Central Limit Theorem: There's a free central limit theorem with normal distribution replaced by semicircular distribution.

What is free probability?

We replace classical independence by *free independence*.

Definition by Example

For groups G_1 and G_2, the algebras $L(G_1)$ and $L(G_2)$ are freely independent in $(L(G_1 * G_2), \tau)$.

Free Central Limit Theorem: There's a free central limit theorem with normal distribution replaced by semicircular distribution.

Free Convolution: If X and Y are classically independent, then $\mu_{X+Y} = \mu_X * \mu_Y$. If X and Y are freely independent, then $\mu_{X+Y} = \mu_X \boxplus \mu_Y$.

What is the law of a tuple?

Classically, the law of $X = (X_1, \dots, X_m)$ is a measure on $\mathbb{R}^m$ given by

$$\mu_X(A) = P(X \in A).$$

What is the law of a tuple?

Classically, the law of $X = (X_1, \dots, X_m)$ is a measure on $\mathbb{R}^m$ given by

$$\mu_X(A) = P(X \in A).$$

Assuming finite moments, this can be viewed as a map

$$\mu_X : \mathbb{C}[x_1, \dots, x_m] \to \mathbb{C}, \quad p(x_1, \dots, x_m) \mapsto E[p(X_1, \dots, X_m)].$$

What is the law of a tuple?

Classically, the law of $X = (X_1, \ldots, X_m)$ is a measure on $\mathbb{R}^m$ given by

$$\mu_X(A) = P(X \in A).$$

Assuming finite moments, this can be viewed as a map

$$\mu_X : \mathbb{C}[x_1, \ldots, x_m] \to \mathbb{C}, \quad p(x_1, \ldots, x_m) \mapsto E[p(X_1, \ldots, X_m)].$$

In the non-commutative case, the *law of* $X = (X_1, \ldots, X_m) \in M_{sa}^m$ is defined as the map

$$\mu_X : \mathbb{C}\langle x_1, \ldots, x_m \rangle \to \mathbb{C}, \quad p(x_1, \ldots, x_m) \mapsto \tau[p(X_1, \ldots, X_m)],$$

What is the law of a tuple?

Classically, the law of $X = (X_1, \dots, X_m)$ is a measure on $\mathbb{R}^m$ given by

$$\mu_X(A) = P(X \in A).$$

Assuming finite moments, this can be viewed as a map

$$\mu_X : \mathbb{C}[x_1, \dots, x_m] \to \mathbb{C}, \quad p(x_1, \dots, x_m) \mapsto E[p(X_1, \dots, X_m)].$$

In the non-commutative case, the *law of* $X = (X_1, \dots, X_m) \in M_{sa}^m$ is defined as the map

$$\mu_X : \mathbb{C}\langle x_1, \dots, x_m \rangle \to \mathbb{C}, \quad p(x_1, \dots, x_m) \mapsto \tau[p(X_1, \dots, X_m)],$$

The *moment topology* on laws is given by pointwise convergence on $\mathbb{C}\langle x_1, \dots, x_m \rangle$.

Background: Microstates Free Entropy χ

What is classical entropy?

The *continuous entropy* of a probability measure $d\mu(x) = \rho(x)\,dx$ on $\mathbb{R}^m$ is given by

$$h(\mu) = -\int \rho \log \rho.$$

If μ does not have a density, we set $h(\mu) = -\infty$.

What is classical entropy?

The *continuous entropy* of a probability measure $d\mu(x) = \rho(x)\,dx$ on $\mathbb{R}^m$ is given by

$$h(\mu) = -\int \rho \log \rho.$$

If μ does not have a density, we set $h(\mu) = -\infty$.

"Entropy measures regularity."

1. If μ is highly concentrated, then there is large negative entropy.
2. For mean zero and variance 1, the highest entropy is achieved by Gaussian.
3. If you smooth μ out by convolution, the entropy increases.

Microstates Interpretation

Since there's no nice integral formula for entropy in the free case, the definition of χ is based on the microstates interpretation.

Microstates Interpretation

Since there's no nice integral formula for entropy in the free case, the definition of χ is based on the microstates interpretation.

Classical case: Given a vector in $x = (x_1, \dots, x_m) \in (\mathbb{R}^N)^m$, let's define its *empirical distribution* as

$$\mu_x = \frac{1}{N} \sum_{j=1}^{N} \delta_{((x_1)_j, \dots, (x_m)_j)}.$$

Then $\{x : \mu_x \text{ is close to } \mu\}$ has measure approximately $\exp(-Nh(\mu))$.

Microstates Interpretation

Since there's no nice integral formula for entropy in the free case, the definition of χ is based on the microstates interpretation.

Classical case: Given a vector in $x = (x_1, \dots, x_m) \in (\mathbb{R}^N)^m$, let's define its *empirical distribution* as

$$\mu_x = \frac{1}{N} \sum_{j=1}^{N} \delta_{((x_1)_j, \dots, (x_m)_j)}.$$

Then $\{x : \mu_x \text{ is close to } \mu\}$ has measure approximately $\exp(-Nh(\mu))$. Thus, $h(\mu)$ can be expressed as

$$\inf_{\text{(nbhd's of } \mu)} \lim_{N \to \infty} \frac{1}{N} \log \operatorname{vol}\{x : \mu_x \text{ close to } \mu\}.$$

Microstates Interpretation

Since there's no nice integral formula for entropy in the free case, the definition of χ is based on the microstates interpretation.

Classical case: Given a vector in $x = (x_1, \dots, x_m) \in (\mathbb{R}^N)^m$, let's define its *empirical distribution* as

$$\mu_x = \frac{1}{N} \sum_{j=1}^N \delta_{((x_1)_j, \dots, (x_m)_j)}.$$

Then $\{x : \mu_x \text{ is close to } \mu\}$ has measure approximately $\exp(-Nh(\mu))$. Thus, $h(\mu)$ can be expressed as

$$\inf_{(\text{nbhd's of } \mu)} \lim_{N \to \infty} \frac{1}{N} \log \operatorname{vol}\{x : \mu_x \text{ close to } \mu\}.$$

Intuition: If μ is more regular and spread out, then there are more microstates because most choices of N vectors are "evenly distributed."

Microstates Free Entropy

Idea for free case: Replace $\mathbb{R}^N$ (self-adjoints in $L^\infty(\{1, \ldots, N\})$) by $M_N(\mathbb{C})_{sa}$.

Microstates Free Entropy

Idea for free case: Replace $\mathbb{R}^N$ (self-adjoints in $L^\infty(\{1, \dots, N\})$) by $M_N(\mathbb{C})_{sa}$.

Given $(x_1, \dots, x_m) \in M_N(\mathbb{C})^m$, the *empirical distribution* μ_x is the non-commutative law of x w.r.t. normalized trace on $M_N(\mathbb{C})$. For a neighborhood $\mathcal{U}$ of μ in the moment topology and $R > 0$, define

$$\Gamma_{N,R}(\mathcal{U}) = \{x : \|x_j\| \leq R \text{ and } \mu \in \mathcal{U}\}.$$

Define

$$\chi(\mu) = \sup_{R>0} \inf_{\mathcal{U} \ni \mu} \limsup_{N \to \infty} \left(\frac{1}{N^2} \log \operatorname{vol} \Gamma_{N,R}(\mathcal{U}) + \frac{m}{2} \log N \right).$$

Microstates Free Entropy

Idea for free case: Replace $\mathbb{R}^N$ (self-adjoints in $L^\infty(\{1, \dots, N\})$) by $M_N(\mathbb{C})_{sa}$.

Given $(x_1, \dots, x_m) \in M_N(\mathbb{C})^m$, the *empirical distribution* μ_x is the non-commutative law of x w.r.t. normalized trace on $M_N(\mathbb{C})$. For a neighborhood $\mathcal{U}$ of μ in the moment topology and $R > 0$, define

$$\Gamma_{N,R}(\mathcal{U}) = \{x : \|x_j\| \leq R \text{ and } \mu \in \mathcal{U}\}.$$

Define

$$\chi(\mu) = \sup_{R>0} \inf_{\mathcal{U} \ni \mu} \limsup_{N \to \infty} \left(\frac{1}{N^2} \log \operatorname{vol} \Gamma_{N,R}(\mathcal{U}) + \frac{m}{2} \log N \right).$$

(Voiculescu) χ has properties similar to h, and also relates to properties of the W^*-algebra generated by a tuple with the law μ.

Background: Non-microstates Free Entropy χ^*

Classical Fisher Information

Classical case: Let μ be a probability measure on $\mathbb{R}^m$ with density ρ. Let γ_t be the law of a Gaussian random vector with variance tI. Then

$$\frac{d}{dt} h(\mu * \gamma_t) = \int |\nabla \rho_t|^2 / \rho_t = \|\nabla \rho_t / \rho_t\|^2_{L^2(\mu * \gamma_t)}.$$

Classical Fisher Information

Classical case: Let μ be a probability measure on $\mathbb{R}^m$ with density ρ. Let γ_t be the law of a Gaussian random vector with variance tI. Then

$$\frac{d}{dt} h(\mu * \gamma_t) = \int |\nabla \rho_t|^2 / \rho_t = \|\nabla \rho_t / \rho_t\|^2_{L^2(\mu * \gamma_t)}.$$

The quantity $\|\nabla \rho_t / \rho_t\|^2_{L^2(\mu * \gamma_t)}$ is called the *Fisher information* of $\mu * \gamma_t$. The entropy can be recovered by integrating the Fisher information.

Classical Fisher Information

Classical case: Let μ be a probability measure on $\mathbb{R}^m$ with density ρ. Let γ_t be the law of a Gaussian random vector with variance tI. Then

$$\frac{d}{dt} h(\mu * \gamma_t) = \int |\nabla \rho_t|^2 / \rho_t = \|\nabla \rho_t / \rho_t\|^2_{L^2(\mu * \gamma_t)}.$$

The quantity $\|\nabla \rho_t / \rho_t\|^2_{L^2(\mu * \gamma_t)}$ is called the *Fisher information* of $\mu * \gamma_t$. The entropy can be recovered by integrating the Fisher information.

Intuition: The Fisher information measures the regularity of μ by looking at its derivatives.

Non-Microstates Free Entropy

In the free case, we don't know an analogue of density, so we want to rephrase the definition using integration by parts.

Non-Microstates Free Entropy

In the free case, we don't know an analogue of density, so we want to rephrase the definition using integration by parts.

Classical Fisher information is L^2 norm of the conjugate variable $\xi = (\nabla \rho / \rho)(X)$, which is characterized by an integration-by-parts formula $E[\xi f(X)] = E[\nabla f(X)]$.

Non-Microstates Free Entropy

In the free case, we don't know an analogue of density, so we want to rephrase the definition using integration by parts.

Classical Fisher information is L^2 norm of the conjugate variable $\xi = (\nabla \rho / \rho)(X)$, which is characterized by an integration-by-parts formula $E[\xi f(X)] = E[\nabla f(X)]$.

Voiculescu used the free version $\tau[\xi_j f(X)] = \tau \otimes \tau[\mathcal{D}_{X_j} f(X)]$ to define the *free conjugate variables* and hence the *free Fisher information*.

Non-Microstates Free Entropy

In the free case, we don't know an analogue of density, so we want to rephrase the definition using integration by parts.

Classical Fisher information is L^2 norm of the conjugate variable $\xi = (\nabla \rho / \rho)(X)$, which is characterized by an integration-by-parts formula $E[\xi f(X)] = E[\nabla f(X)]$.

Voiculescu used the free version $\tau[\xi_j f(X)] = \tau \otimes \tau[\mathcal{D}_{X_j} f(X)]$ to define the *free conjugate variables* and hence the *free Fisher information*.

$\chi^*(\mu)$ is defined by integrating the free Fisher information of $\mu \boxplus \sigma_t$, where σ_t is the law of a free semicircular family where each variable has mean zero and variance t.

Background: Free Gibbs Laws

Free Gibbs Laws

Classically, a Gibbs measure on $\mathbb{R}^m$ is a measure of the form $(1/Z)e^{-V(x)}\,dx$. This can be characterized by the equation

$$\int \partial_j V \cdot f \, d\mu = \int \partial_j f \, d\mu.$$

Free Gibbs Laws

Classically, a Gibbs measure on $\mathbb{R}^m$ is a measure of the form $(1/Z)e^{-V(x)}\,dx$. This can be characterized by the equation

$$\int \partial_j V \cdot f \, d\mu = \int \partial_j f \, d\mu.$$

If $g(X)$ is a non-commutative polynomial in $X_1, \dots, X_m$, then μ is said to be a free Gibbs law for g if

$$\mu[\mathcal{D}_j^\circ g(X) f(X)] = \mu \otimes \mu[\mathcal{D}_j f(X)],$$

where $\mathcal{D}_j^\circ v(X)$ is the cyclic derivative with respect to X_j. In other words, $\mathcal{D}^\circ g(X)$ is the conjugate variable of X with respect to μ.

Free Gibbs Laws

(Guionnet, Maurel-Segala, Shylakhtenko, Dabrowski) If $g(X)$ is a small (or a convex) perturbation of $(1/2)\sum_j X_j^2$, then there exists a unique free Gibbs law for g.

Free Gibbs Laws

(Guionnet, Maurel-Segala, Shylakhtenko, Dabrowski) If $g(X)$ is a small (or a convex) perturbation of $(1/2)\sum_j X_j^2$, then there exists a unique free Gibbs law for g.

These Gibbs laws also have good random matrix models. Let's just look at the case where $g(X)$ is uniformly convex (globally). Define a random matrix model μ_N (a probability measure on $M_N(\mathbb{C})_{sa}^m$) by

$$d\mu_N(x) = \frac{1}{Z_N} e^{-N \operatorname{Tr}(g(x))}\, dx,$$

where dx is Lebesgue measure. If X_N is a random variable given by μ_N, then the non-commutative laws of X_N converge almost surely to μ as $N \to \infty$.

Free Gibbs Laws

(Herbst, Bakry-Emery, Ledoux, Guionnet, Maurel-Segala, ...) Moreover, since $g(x)$ is uniformly convex, the measures μ_N satisfy log-Sobolev inequalities and concentration of measure.

Free Gibbs Laws

(Herbst, Bakry-Emery, Ledoux, Guionnet, Maurel-Segala, ...) Moreover, since $g(x)$ is uniformly convex, the measures μ_N satisfy log-Sobolev inequalities and concentration of measure.

Specifically, if f is a real-valued function that is Lipschitz in $\|\cdot\|_2$, then $f(X_N)$ is exponentially unlikely to be more than δ away from its expectation.

Free Gibbs Laws

(Herbst, Bakry-Emery, Ledoux, Guionnet, Maurel-Segala, ...) Moreover, since $g(x)$ is uniformly convex, the measures μ_N satisfy log-Sobolev inequalities and concentration of measure.

Specifically, if f is a real-valued function that is Lipschitz in $\|\cdot\|_2$, then $f(X_N)$ is exponentially unlikely to be more than δ away from its expectation.

Also, in this case, the operator norm $X_N - E[X_N]$ is bounded by some constant R with very high probability, so as $N \to +\infty$, we can restrict our measures to operator norm balls without losing much.

Results and Approach

Notation

τ_N is the normalized trace on $M_N(\mathbb{C})$.

$\|\cdot\|_2$ is the corresponding 2-norm, that is, for $x \in M_N(\mathbb{C})_{sa}^m$, we set $\|x\|_2^2 = \sum_{j=1}^m \tau_N(x_j^2)$.

$\|\cdot\|$ is the operator norm.

$\sigma_{N,t}$ denotes the law of m independent $N \times N$ GUE matrices which each have mean zero and variance t.

σ_t denotes the non-commutative law of m freely independent semicirculars which each have mean zero and variance t.

Setup

Consider a sequence of potentials $V_N : M_N(\mathbb{C})_{sa}^m \to \mathbb{R}$ that are uniformly convex and semi-concave, that is, for some $0 < c < C$, we have

$$V_N(x) - \frac{c}{2}\|x\|_2^2 \text{ convex and } V_N(x) - \frac{C}{2}\|x\|_2^2 \text{ concave.}$$

Let μ_N be the probability measure on $M_N(\mathbb{C})_{sa}^m$ given by

$$d\mu_N(x) = \frac{1}{Z_N} e^{-N^2 V_N(x)}\, dx$$

where dx is Lebesgue measure and Z_N is a normalizing constant.

Setup

For example, if $g(x)$ is a non-commutative polynomial in $x_1, \dots, x_m$, then we could take $V_N(x) = \tau_N(g(x))$. In this case, the gradient of V_N (with respect to the $\|\cdot\|_2$ given by the normalized trace) would be

$$DV_N(x) = \mathcal{D}^\circ g(x)$$

Setup

For example, if $g(x)$ is a non-commutative polynomial in $x_1, \dots, x_m$, then we could take $V_N(x) = \tau_N(g(x))$. In this case, the gradient of V_N (with respect to the $\|\cdot\|_2$ given by the normalized trace) would be

$$DV_N(x) = \mathcal{D}^\circ g(x)$$

We won't assume that V_N has this form, or that it is unitarily invariant on the nose. Rather, we will use the more flexible assumption that the gradients DV_N can be approximated on operator norm balls by trace polynomials. (Precise definition later.)

Results

Theorem

Let V_N be uniformly convex and semi-concave as above (note that the standard concentration results apply). Suppose that the sequence of normalized gradients DV_N is asymptotically approximable by trace polynomials. Suppose that the expected values $\int x \, d\mu_N$ are bounded in operator norm as $N \to \infty$. Then

1. $\mu(p) := \lim_{N\to\infty} \int \tau_N(p(x)) \, d\mu_N(x)$ *exists for every non-commutative polynomial p.*
2. *The non-commutative law μ has finite free Fisher information and finite free entropy.*
3. $\chi(\mu) = \chi^*(\mu) = \lim_{N\to\infty}[N^{-2} h(\mu_N) + (m/2)\log N]$.
4. *The normalized Fisher information of $\mu_N * \sigma_{N,t}$ converges to the free Fisher information of $\mu \boxplus \sigma_t$ for every $t \geq 0$.*
5. *The free Fisher information is locally Lipschitz in t.*

Approach

- The quantities we want to study ($\int \phi(x)\, \mu_N(x)$ or Fisher information of $\mu_N * \sigma_{N,t}$) can be expressed by solving some PDE (parallel to SDE approach of Dabrowski, Guionnet, Shlyakhtenko).

Approach

- The quantities we want to study ($\int \phi(x)\,\mu_N(x)$ or Fisher information of $\mu_N * \sigma_{N,t}$) can be expressed by solving some PDE (parallel to SDE approach of Dabrowski, Guionnet, Shlyakhtenko).
- We want to show that the solution at a given time t is asymptotically approximable by trace polynomials (AATP).

Approach

- The quantities we want to study ($\int \phi(x)\,\mu_N(x)$ or Fisher information of $\mu_N * \sigma_{N,t}$) can be expressed by solving some PDE (parallel to SDE approach of Dabrowski, Guionnet, Shlyakhtenko).
- We want to show that the solution at a given time t is asymptotically approximable by trace polynomials (AATP).
- AATP is preserved under certain operations, and by combining / iterating these simple operations we can build an approximation to the solution.

Approach

- The quantities we want to study ($\int \phi(x)\, \mu_N(x)$ or Fisher information of $\mu_N * \sigma_{N,t}$) can be expressed by solving some PDE (parallel to SDE approach of Dabrowski, Guionnet, Shlyakhtenko).
- We want to show that the solution at a given time t is asymptotically approximable by trace polynomials (AATP).
- AATP is preserved under certain operations, and by combining / iterating these simple operations we can build an approximation to the solution.
- We give dimension-independent estimates for the errors in the approximation.
- Hence, the solution has AATP.

Approach

We can apply this strategy to several different situations:

Approach

We can apply this strategy to several different situations:

1. By studying the diffusion associated to V_N, we can analyze $\int \phi(x)\, d\mu_N(x)$ for test functions ϕ, and thus prove that the limit as $N \to \infty$ of the laws exists.

Approach

We can apply this strategy to several different situations:

1. By studying the diffusion associated to V_N, we can analyze $\int \phi(x)\, d\mu_N(x)$ for test functions ϕ, and thus prove that the limit as $N \to \infty$ of the laws exists.
2. By studying the evolution of the potential V_N as we convolve μ_N with Gaussians, we get a handle on the conjugate variables for $\mu_N * \sigma_{N,t}$, and hence on the Fisher information.

Approach

We can apply this strategy to several different situations:

1. By studying the diffusion associated to V_N, we can analyze $\int \phi(x)\, d\mu_N(x)$ for test functions ϕ, and thus prove that the limit as $N \to \infty$ of the laws exists.
2. By studying the evolution of the potential V_N as we convolve μ_N with Gaussians, we get a handle on the conjugate variables for $\mu_N * \sigma_{N,t}$, and hence on the Fisher information.
3. (Future work.) The same idea as (1) can be applied to study the conditional expectation of $\phi(X)$ given $X_1, \dots, X_r$ for some $r < m$.

Approach

We can apply this strategy to several different situations:

1. By studying the diffusion associated to V_N, we can analyze $\int \phi(x)\, d\mu_N(x)$ for test functions ϕ, and thus prove that the limit as $N \to \infty$ of the laws exists.
2. By studying the evolution of the potential V_N as we convolve μ_N with Gaussians, we get a handle on the conjugate variables for $\mu_N * \sigma_{N,t}$, and hence on the Fisher information.
3. (Future work.) The same idea as (1) can be applied to study the conditional expectation of $\phi(X)$ given $X_1, \dots, X_r$ for some $r < m$.
4. (Future work.) Using what we know from (2), we can construct a transport map from $\mu_N * \sigma_{N,t}$ to Gaussian, and hence in the limit, we can establish free transport to semicircular for the law μ.

Asymptotic Approximation by Trace Polynomials

Trace Polynomials

Trace polynomials in $x_1, \ldots, x_m$ are linear combinations of functions of the form $p_0 \tau(p_1) \ldots \tau(p_n)$ where p_j is a non-commutative polynomial in $x_1, \ldots, x_m$. For example,

$$\tau(x_1 x_2) x_1 + 3\tau(x_2^2)\tau(x_1 x_3) x_3 x_2 + 5\tau(x_3^2)$$

Trace Polynomials

Trace polynomials in $x_1, \dots, x_m$ are linear combinations of functions of the form $p_0 \tau(p_1) \dots \tau(p_n)$ where p_j is a non-commutative polynomial in $x_1, \dots, x_m$. For example,

$$\tau(x_1 x_2) x_1 + 3\tau(x_2^2)\tau(x_1 x_3) x_3 x_2 + 5\tau(x_3^2)$$

If p is a trace polynomial, then p defines a function $M_N(\mathbb{C})_{sa}^m \to M_N(\mathbb{C})$. We interpret τ as the normalized trace on $M_N(\mathbb{C})$ and evaluate p at the point x.

Trace Polynomials

Trace polynomials in $x_1, \ldots, x_m$ are linear combinations of functions of the form $p_0 \tau(p_1) \ldots \tau(p_n)$ where p_j is a non-commutative polynomial in $x_1, \ldots, x_m$. For example,

$$\tau(x_1 x_2) x_1 + 3\tau(x_2^2)\tau(x_1 x_3) x_3 x_2 + 5\tau(x_3^2)$$

If p is a trace polynomial, then p defines a function $M_N(\mathbb{C})_{sa}^m \to M_N(\mathbb{C})$. We interpret τ as the normalized trace on $M_N(\mathbb{C})$ and evaluate p at the point x.

More generally, if (M, τ) is a tracial von Neumann algebra, then p defines a map $M_{sa}^m \to M$.

Asymptotic Approximation by Trace Polynomials

Definition

A sequence of functions $\phi_N : M_N(\mathbb{C})_{sa}^m \to M_N(\mathbb{C})_{sa}^m$ is *asymptotically approximable by trace polynomials* if for every $\epsilon > 0$ and $R > 0$, there exists an m-tuple of trace polynomials f such that

$$\limsup_{N \to \infty} \sup_{\substack{x \in M_N(\mathbb{C})_{sa}^m \\ \|x_j\| \leq R}} \|\phi_N(x) - f(x)\|_2 \leq \epsilon.$$

Linear Combinations and Limits

AATP is preserved by various natural operations. Obviously, it's preserved under linear combinations. Also, you can take limits.

Lemma

Suppose that $\phi_{N,k}$ is a bi-indexed sequence of functions and ϕ_N another sequence. If $(\phi_{N,k})_{N\in\mathbb{N}}$ has AATP for each k and if for every $R > 0$, we have

$$\lim_{k\to 0} \limsup_{N\to\infty} \sup_{\|x\|\leq R} \|\phi_{N,k} - \phi_N\|_2 = 0,$$

then $(\phi_N)_{N\in\mathbb{N}}$ also has AATP.

The proof is trivial.

Convolution with Gaussian

Lemma

Suppose that $\phi_N : M_N(\mathbb{C})_{sa}^m \to M_N(\mathbb{C})_{sa}^m$ *has AATP. Suppose that* $\|\phi_{N,t}(x)\|_2$ *has some reasonable growth as* $x \to \infty$*, for instance,* $\|\phi_{N,t}(x)\|_2 \leq A + B\|x\|^\alpha$ *is sufficient. Then* $\phi_N * \sigma_{N,t}$ *also has AATP for each* t*, where* $\sigma_{N,t}$ *is the GUE measure.*

Convolution with Gaussian

Lemma

Suppose that $\phi_N : M_N(\mathbb{C})_{sa}^m \to M_N(\mathbb{C})_{sa}^m$ *has AATP. Suppose that* $\|\phi_{N,t}(x)\|_2$ *has some reasonable growth as* $x \to \infty$*, for instance,* $\|\phi_{N,t}(x)\|_2 \leq A + B\|x\|^\alpha$ *is sufficient. Then* $\phi_N * \sigma_{N,t}$ *also has AATP for each* t*, where* $\sigma_{N,t}$ *is the GUE measure.*

Proof.

Fix $R > 0$. Choose a trace polynomial f that asymtotically approximates $\phi_N(x)$ within ϵ for $\|x\| \leq R + 3t^{1/2}$. Since GUE has operator norm bounded by $2t^{1/2}$ with high probability (with good tail bounds), we see that $\|f * \sigma_{N,t} - \phi_N * \sigma_{N,t}\|_2 \to 0$ uniformly on $\|x\| \leq R$. We can explicitly compute the convolution of GUE with a given trace polynomial f, and this is a trace polynomial which depends on N but has a limit as $N \to +\infty$. So $f * \sigma_{N,t}$ provides an ϵ-approximation for $\phi_N * \sigma_{N,t}$ on the ball of radius R. □

Composition

Lemma

Suppose that $\phi_N : M_N(\mathbb{C})_{sa}^m \to M_N(\mathbb{C})_{sa}^m$ *and* $\psi_N : M_N(\mathbb{C})_{sa}^m \to M_N(\mathbb{C})_{sa}^m$ *have AATP. Suppose that* ϕ_N *is uniformly continuous in* $\|\cdot\|_2$ *(uniformly in* N*). Then* $\phi_N \circ \psi_N$ *has AATP.*

Composition

Lemma

Suppose that $\phi_N : M_N(\mathbb{C})_{sa}^m \to M_N(\mathbb{C})_{sa}^m$ *and* $\psi_N : M_N(\mathbb{C})_{sa}^m \to M_N(\mathbb{C})_{sa}^m$ *have AATP. Suppose that* ϕ_N *is uniformly continuous in* $\|\cdot\|_2$ *(uniformly in* N*). Then* $\phi_N \circ \psi_N$ *has AATP.*

Proof.

Fix R and ϵ. Choose δ so that $\|x - y\|_2 < \delta$ implies $\|\phi_N(x) - \phi_N(y)\| < \epsilon/2$. Then choose a trace polynomial g which is an asymptotic δ-approximation of ψ_N on $\|x\| \leq R$. For $\|x\| \leq R$, the function g is bounded in operator norm by some R'. Let f be asymptotic $\epsilon/2$-approximation of ϕ_N on $\|x\| \leq R'$. Then

$$\|\phi_N \circ \psi_N - f \circ g\|_2 \leq \|\phi_N \circ \psi_N - \phi_N \circ g\|_2 + \|\phi_N \circ g - f \circ g\|_2.$$

As $N \to +\infty$, the sup of each term on $\|x\| \leq R$ is bounded by $\epsilon/2$. □

Ordinary Differential Equations

Lemma

Let $\phi_N(x,t)$ *be a function* $M_N(\mathbb{C})_{sa}^m \times [0,T] \to M_N(\mathbb{C})_{sa}^m$*. Suppose that* ϕ_N *is globally Lipschitz in* (x,t) *and that* $\phi_N(\cdot,t)$ *is AATP for each* t*. Let* ψ_N *solve the equation*

$$\partial_t \psi_N(x,t) = \phi_N(\psi_N(x,t),t)$$

and suppose that $\psi_N(\cdot,0)$ *has AATP. Then* $\psi_N(\cdot,t)$ *has AATP for each* t*.*

Ordinary Differential Equations

Sketch of proof.

The solution $\psi_N(x,t)$ can be evaluated through Picard iteration. Since $\phi_N(x,t)$ is Lipschitz in (x,t) and AATP for each fixed t, we know that for $R, \epsilon > 0$, there is a function $f(x,t)$ that is piecewise constant in each t and is a trace polynomial for each x, such that

$$\limsup_{N \to \infty} \sup_{\|x\| \leq R, t \in [0,T]} \|\phi_N(x,t) - f(x,t)\|_2 \leq \epsilon.$$

This property is preserved by the composition and integration operations of Picard iteration. □

Convergence of Laws μ_N

Goal

We need to prove that for a non-commutative polynomial p, the limit $\lim_{N\to\infty} \int \tau_N(p)\, d\mu_N$ exists. It's more convenient to replace p by something globally Lipschitz. For instance, let $f : \mathbb{R} \to \mathbb{R}$ be a C_c^∞ function such that $f(x) = x$ for $|x| \leq R$, and consider

$$\phi(x) = \tau_N(p(f(x_1), \ldots, f(x_m))).$$

If R is large enough, then $\int \tau_N(p) - \phi\, d\mu_N \to 0$. Also, ϕ is globally Lipschitz in $\|\cdot\|_2$ and it has AATP.

Goal

So it suffices to prove the following statement:

Theorem

Suppose that $\phi_N : M_N(\mathbb{C})_{sa}^m \to \mathbb{R}$ *is AATP. Then* $\lim_{N \to \infty} \int \phi_N \, d\mu_N$ *exists.*

Strategy

The fact that DV_N has AATP is the only assumption that relates the potentials V_N for different values of N. Thus, we'll need to evaluate $\int \phi_N \, d\mu_N$ in terms of ϕ_N and DV_N.

Strategy

The fact that DV_N has AATP is the only assumption that relates the potentials V_N for different values of N. Thus, we'll need to evaluate $\int \phi_N \, d\mu_N$ in terms of ϕ_N and DV_N.

Various papers have considered μ_N as the stationary distribution of the SDE

$$dX_t = dB_t - \frac{1}{2} DV_N(X_t) \, dt,$$

where B_t is a GUE Brownian motion.

Strategy

The PDE viewpoint is that if u_N solves

$$u_N(x,0) = \phi_N(x)$$
$$\partial_t u_N = \frac{1}{2N}\Delta u_N - \frac{1}{2}\langle Du_N, DV_N\rangle,$$

then

$$\lim_{N\to\infty} u_N(x,t) = \int \phi_N \, d\mu_N \text{ for all } x.$$

The first term

We consider the two terms on the right hand side separately. If we removed the DV_N term, we would have the heat equation

$$\partial_t u = \frac{1}{2N} \Delta u.$$

The first term

We consider the two terms on the right hand side separately. If we removed the DV_N term, we would have the heat equation

$$\partial_t u = \frac{1}{2N} \Delta u.$$

This equation can be solved by application of the heat semigroup $P_t \phi = \phi * \sigma_{N,t}$. We already showed that P_t preserves AATP.

The second term

If we only considered the second term of the equation, we would have

$$\partial_t u = -\frac{1}{2} \langle Du, DV_N \rangle.$$

The second term

If we only considered the second term of the equation, we would have

$$\partial_t u = -\frac{1}{2}\langle Du, DV_N \rangle.$$

This equation can be solved by the semigroup

$$S_t \phi(x) = \phi(W_N(x,t)),$$

where

$$W_N(x,0) = x \text{ and } \partial_t W_N(x,t) = \frac{1}{2} DV_N(W_N(x,t)).$$

The second term

If we only considered the second term of the equation, we would have

$$\partial_t u = -\frac{1}{2}\langle Du, DV_N \rangle.$$

This equation can be solved by the semigroup

$$S_t \phi(x) = \phi(W_N(x,t)),$$

where

$$W_N(x,0) = x \text{ and } \partial_t W_N(x,t) = \frac{1}{2} DV_N(W_N(x,t)).$$

Since solving the ODE preserves AATP and so does composition, we know that S_t preserves AATP.

Trotter's formula

As in Trotter's formula, we want to define a semigroup T_t by

$$T_t \phi = \lim_{n \to \infty} (P_{t/n} S_{t/n})^n \phi,$$

and we claim that $T_t \phi$ will solve the main PDE.

Trotter's formula

As in Trotter's formula, we want to define a semigroup T_t by

$$T_t \phi = \lim_{n \to \infty} (P_{t/n} S_{t/n})^n \phi,$$

and we claim that $T_t \phi$ will solve the main PDE.

Actually, it's more convenient to first define for dyadic rational values of t

$$T_{k,t} = (P_{1/2^k} S_{1/2^k})^{2^k t}.$$

Trotter's formula

Now we estimate $\|T_{k,t} - T_{k+1,t}\|_{L^\infty}$. Note that $T_{k+1,t}$ is obtained from $T_{k,t}$ by replacing each copy of $P_{1/2^k} S_{1/2^k}$ with $P_{1/2^{k+1}} S_{1/2^{k+1}} P_{1/2^{k+1}} S_{1/2^{k+1}}$.

Trotter's formula

Now we estimate $\|T_{k,t} - T_{k+1,t}\|_{L^\infty}$. Note that $T_{k+1,t}$ is obtained from $T_{k,t}$ by replacing each copy of $P_{1/2^k} S_{1/2^k}$ with $P_{1/2^{k+1}} S_{1/2^{k+1}} P_{1/2^{k+1}} S_{1/2^{k+1}}$.

One can check directly that $\|S_\delta P_\delta \phi - P_\delta S_\delta \phi\|_{L^\infty} \leq \text{const} \delta^{3/2} \|\phi\|_{\text{Lip}}$.

Trotter's formula

Now we estimate $\| T_{k,t} - T_{k+1,t} \|_{L^\infty}$. Note that $T_{k+1,t}$ is obtained from $T_{k,t}$ by replacing each copy of $P_{1/2^k} S_{1/2^k}$ with $P_{1/2^{k+1}} S_{1/2^{k+1}} P_{1/2^{k+1}} S_{1/2^{k+1}}$.

One can check directly that $\|S_\delta P_\delta \phi - P_\delta S_\delta \phi\|_{L^\infty} \leq \text{const}\delta^{3/2} \|\phi\|_{\text{Lip}}$.

So we must make $O(2^k)$ swaps and each one gives us an error $O(2^{-3k/2})$.

Trotter's formula

Now we estimate $\|T_{k,t} - T_{k+1,t}\|_{L^\infty}$. Note that $T_{k+1,t}$ is obtained from $T_{k,t}$ by replacing each copy of $P_{1/2^k} S_{1/2^k}$ with $P_{1/2^{k+1}} S_{1/2^{k+1}} P_{1/2^{k+1}} S_{1/2^{k+1}}$.

One can check directly that $\|S_\delta P_\delta \phi - P_\delta S_\delta \phi\|_{L^\infty} \leq \text{const}\delta^{3/2} \|\phi\|_{\text{Lip}}$.

So we must make $O(2^k)$ swaps and each one gives us an error $O(2^{-3k/2})$.

The errors do not grow as they propagate through additional applications of $S_{1/2^k}$ and $P_{1/2^k}$ because the operators S_δ and P_δ will not increase the Lipschitz norm or L^∞ norm of functions.

Trotter's formula

Now we estimate $\|T_{k,t} - T_{k+1,t}\|_{L^\infty}$. Note that $T_{k+1,t}$ is obtained from $T_{k,t}$ by replacing each copy of $P_{1/2^k} S_{1/2^k}$ with $P_{1/2^{k+1}} S_{1/2^{k+1}} P_{1/2^{k+1}} S_{1/2^{k+1}}$.

One can check directly that $\|S_\delta P_\delta \phi - P_\delta S_\delta \phi\|_{L^\infty} \leq \text{const}\delta^{3/2} \|\phi\|_{\text{Lip}}$.

So we must make $O(2^k)$ swaps and each one gives us an error $O(2^{-3k/2})$.

The errors do not grow as they propagate through additional applications of $S_{1/2^k}$ and $P_{1/2^k}$ because the operators S_δ and P_δ will not increase the Lipschitz norm or L^∞ norm of functions.

So the overall error between $T_{k,t}$ and $T_{k+1,t}$ is $O(2^{-k/2})$. This is summable in k, so we get convergence as $k \to \infty$.

Conclusion of the Proof

So the operators T_t are well-defined for dyadic t, but one can check Hölder continuity in t, and hence extend them to all real t.

Conclusion of the Proof

So the operators T_t are well-defined for dyadic t, but one can check Hölder continuity in t, and hence extend them to all real t.

Then one checks that they give the solution to the PDE.

Conclusion of the Proof

So the operators T_t are well-defined for dyadic t, but one can check Hölder continuity in t, and hence extend them to all real t.

Then one checks that they give the solution to the PDE.

The operators T_t preserve AATP for Lipschitz functions.

Conclusion of the Proof

We have $\int T_t \phi_N \, d\mu_N = \int \phi_N \, d\mu_N$ by a direct computation (and justification).

Conclusion of the Proof

We have $\int T_t \phi_N \, d\mu_N = \int \phi_N \, d\mu_N$ by a direct computation (and justification).

Due to the uniform convexity of V_N, we have

$$\|S_t \phi\|_{\text{Lip}} \leq e^{-ct/2} \|\phi\|_{\text{Lip}},$$

which follows from estimating the Lipschitz norm the function $W_N(x,t)$ used to define S_t.

Conclusion of the Proof

We have $\int T_t \phi_N \, d\mu_N = \int \phi_N \, d\mu_N$ by a direct computation (and justification).

Due to the uniform convexity of V_N, we have

$$\|S_t \phi\|_{\text{Lip}} \leq e^{-ct/2} \|\phi\|_{\text{Lip}},$$

which follows from estimating the Lipschitz norm the function $W_N(x,t)$ used to define S_t.

Thus, $T_t \phi_N \to \int \phi_N \, d\mu_N$ as $N \to \infty$ with a dimension-independent rate of convergence. Since $T_t \phi_N$ preserves AATP, we know that the sequence of constant functions $\int \phi_N \, d\mu_N$ has AATP.

Conclusion of the Proof

We have $\int T_t \phi_N \, d\mu_N = \int \phi_N \, d\mu_N$ by a direct computation (and justification).

Due to the uniform convexity of V_N, we have

$$\|S_t \phi\|_{\mathrm{Lip}} \leq e^{-ct/2} \|\phi\|_{\mathrm{Lip}},$$

which follows from estimating the Lipschitz norm the function $W_N(x, t)$ used to define S_t.

Thus, $T_t \phi_N \to \int \phi_N \, d\mu_N$ as $N \to \infty$ with a dimension-independent rate of convergence. Since $T_t \phi_N$ preserves AATP, we know that the sequence of constant functions $\int \phi_N \, d\mu_N$ has AATP.

But this just means that the sequence of constants has a limit as $N \to \infty$.

Unification of Free Entropy

Classical and Microstates Entropy

Lemma

$\chi(\mu) = \limsup_{N \to \infty} (N^{-2} h(\mu_N) + (m/2) \log N)$.

Classical and Microstates Entropy

Lemma

$\chi(\mu) = \limsup_{N \to \infty} (N^{-2} h(\mu_N) + (m/2) \log N)$.

Sketch of proof:

Fix a large value of R. Because of the tail bounds on μ_N, the limiting behavior of $N^{-2} h(\mu_N)$ will be unchanged if we truncate μ_N to $\|x\| \leq R$. If we choose a neighborhood $\mathcal{U}$ of μ, then μ_N is concentrated on the microstate space $\Gamma_{N,R}(\mathcal{U})$.

Classical and Microstates Entropy

Since DV_N has AATP, so does $V_N - V_N(0)$. Since trace polynomials are continuous with respect to convergence in moments, we see that V_N is approximately constant on the microstate space $\Gamma_{N,R}(\mathcal{U})$ if $\mathcal{U}$ is sufficiently small. So we can approximate μ_N by the uniform distribution on the microstate space, and hence approximate $N^{-2} h(\mu_N) + (m/2) \log N$ by $N^{-2} \log |\Gamma_{N,R}(\mathcal{U})| + (m/2) \log N$.

Convergence of Fisher Information

Lemma

If DV_N has AATP, then the (normalized) classical Fisher information converges to the free Fisher information (and the latter is finite).

Convergence of Fisher Information

Lemma

If DV_N has AATP, then the (normalized) classical Fisher information converges to the free Fisher information (and the latter is finite).

Sketch of proof:

Suppose that f_k is a sequence of trace polynomials which as k increases provide better and better asymptotic approximations for DV_N. Then f_k will converge in $L^2(\mu)$ to some f. Also, f is a free conjugate variables for μ since the f_k's approximately satisfy the integration by parts formula. Then we check that $\|DV_N\|_{L^2(\mu_N)} \to \|f\|_{L^2(\mu)}$.

Convergence of Fisher Information

We know that $h(\mu_N)$ is given by integrating the classical Fisher information, and χ^* is given by integrating the free Fisher information.

Convergence of Fisher Information

We know that $h(\mu_N)$ is given by integrating the classical Fisher information, and χ^* is given by integrating the free Fisher information.

So to prove $N^{-2}h(\mu_N) + (m/2)\log N \to \chi^*(\mu)$, it suffices to show that the classical Fisher information of $\mu_N * \sigma_{N,t}$ converges to the free Fisher information of $\mu \boxplus \sigma_t$ for each t.

Convergence of Fisher Information

We know that $h(\mu_N)$ is given by integrating the classical Fisher information, and χ^* is given by integrating the free Fisher information.

So to prove $N^{-2}h(\mu_N) + (m/2)\log N \to \chi^*(\mu)$, it suffices to show that the classical Fisher information of $\mu_N * \sigma_{N,t}$ converges to the free Fisher information of $\mu \boxplus \sigma_t$ for each t.

By the last lemma, it suffices to show that the conjugate variables for $\mu_N * \sigma_{N,t}$ have AATP for each t.

Evolution of Potentials

So we reduce the proof to the following claim:

Theorem

Let $V_{N,t}$ be the potential corresponding to $\mu_N * \sigma_{N,t}$. Then $V_{N,t}$ is convex and semi-concave, and $DV_{N,t}$ has AATP.

Evolution of Potentials

So we reduce the proof to the following claim:

Theorem

Let $V_{N,t}$ be the potential corresponding to $\mu_N * \sigma_{N,t}$. Then $V_{N,t}$ is convex and semi-concave, and $DV_{N,t}$ has AATP.

Since the density of $\mu_N * \sigma_{N,t}$ evolves according to the heat equation, we can compute that

$$\partial_t V_{N,t} = \frac{1}{2N} \Delta V_{N,t} - \frac{1}{2} \|DV_{N,t}\|_2^2.$$

Evolution of Potentials

We play the same game as before: Consider semigroups that would give the first and the second term on the right hand side individually, and then blend them together using Trotter's formula.

Evolution of Potentials

We play the same game as before: Consider semigroups that would give the first and the second term on the right hand side individually, and then blend them together using Trotter's formula.

We already know that the heat semigroup P_t corresponds to the Laplacian operator $(1/2N)\Delta V_{N,t}$.

Evolution of Potentials

We play the same game as before: Consider semigroups that would give the first and the second term on the right hand side individually, and then blend them together using Trotter's formula.

We already know that the heat semigroup P_t corresponds to the Laplacian operator $(1/2N)\Delta V_{N,t}$.

Meanwhile, the equation $\partial_t u = -(1/2)\|u\|_2^2$ can be solved using the Hopf-Lax inf-convolution semigroup

$$Q_t u(x) = \inf_y \left[u(y) - \frac{1}{2t}\|x - y\|_2^2 \right]$$

Trotter's formula

The operators P_t and Q_t both preserve the class of functions that are convex and C-semiconcave.

Trotter's formula

The operators P_t and Q_t both preserve the class of functions that are convex and C-semiconcave.

Also, for all such functions $DV_{N,t}$ is C-Lipschitz.

Trotter's formula

The operators P_t and Q_t both preserve the class of functions that are convex and C-semiconcave.

Also, for all such functions $DV_{N,t}$ is C-Lipschitz.

Based on these facts, we can do a bunch of explicit estimates and show that

$$(P_{1/2^k} Q_{1/2^k})^{2^k t} u$$

converges as $k \to \infty$ to some $R_t u$ for dyadic t.

Trotter's formula

Also, we can show that $D(P_{1/2^k} Q_{1/2^k})^{2^k t} u$ converges as $k \to \infty$.

Trotter's formula

Also, we can show that $D(P_{1/2^k} Q_{1/2^k})^{2^k t} u$ converges as $k \to \infty$.

The semigroup R_t extends to all real t. Also, we can check that it gives a viscosity solution to the PDE, and hence gives the unique smooth solution.

Trotter's formula

Also, we can show that $D(P_{1/2^k} Q_{1/2^k})^{2^k t} u$ converges as $k \to \infty$.

The semigroup R_t extends to all real t. Also, we can check that it gives a viscosity solution to the PDE, and hence gives the unique smooth solution.

Some explicit estimates for the gradients show that $R_t u$ is Hölder-1/2 continuous in t and this implies that the classical Fisher information is Lipschitz in t (on compact time intervals).

AATP

To show that $DV_{N,t}$ has AATP, it only remains to check that that if Du is AATP, then so is $D(Q_t u)$.

AATP

To show that $DV_{N,t}$ has AATP, it only remains to check that that if Du is AATP, then so is $D(Q_t u)$.

To do this, we use the magical fact that

$$D(Q_t u)(x) = Du(x - tD(Q_t u)(x)).$$

AATP

To show that $DV_{N,t}$ has AATP, it only remains to check that that if Du is AATP, then so is $D(Q_t u)$.

To do this, we use the magical fact that

$$D(Q_t u)(x) = Du(x - tD(Q_t u)(x)).$$

This says that $D(Q_t u)(x)$ is the fixed point of the function $y \mapsto Du(x - ty)$. SInce Du is C-Lipschitz, this will be a contraction for small values of t.

AATP

To show that $DV_{N,t}$ has AATP, it only remains to check that that if Du is AATP, then so is $D(Q_t u)$.

To do this, we use the magical fact that

$$D(Q_t u)(x) = Du(x - tD(Q_t u)(x)).$$

This says that $D(Q_t u)(x)$ is the fixed point of the function $y \mapsto Du(x - ty)$. SInce Du is C-Lipschitz, this will be a contraction for small values of t.

The fixed point can thus be evaluated by iterating this function. Since AATP is preserved by composition and limits, if u_N has AATP, then $D(Q_t u_N)$ has AATP for small t.

AATP

To show that $DV_{N,t}$ has AATP, it only remains to check that that if Du is AATP, then so is $D(Q_t u)$.

To do this, we use the magical fact that

$$D(Q_t u)(x) = Du(x - tD(Q_t u)(x)).$$

This says that $D(Q_t u)(x)$ is the fixed point of the function $y \mapsto Du(x - ty)$. SInce Du is C-Lipschitz, this will be a contraction for small values of t.

The fixed point can thus be evaluated by iterating this function. Since AATP is preserved by composition and limits, if u_N has AATP, then $D(Q_t u_N)$ has AATP for small t.

Since Q_t is a semigroup, this can be extended to all t.