

FREE MONOTONE TRANSPORT

A. GUIONNET[†] AND D. SHLYAKHTENKO[‡]

ABSTRACT. By solving a free analog of the Monge-Ampère equation, we prove a non-commutative analog of Brenier's monotone transport theorem: if an n -tuple of self-adjoint non-commutative random variables Z_1, \dots, Z_n satisfies a regularity condition (its conjugate variables ξ_1, \dots, ξ_n should be analytic in Z_1, \dots, Z_n and ξ_j should be close to Z_j in a certain analytic norm), then there exist invertible non-commutative functions F_j of an n -tuple of semicircular variables S_1, \dots, S_n , so that $Z_j = F_j(S_1, \dots, S_n)$. Moreover, F_j can be chosen to be monotone, in the sense that $F_j = \mathcal{D}_j g$ and g is a non-commutative function with a positive definite Hessian. In particular, we can deduce that $C^*(Z_1, \dots, Z_n) \cong C^*(S_1, \dots, S_n)$ and $W^*(Z_1, \dots, Z_n) \cong L(\mathbb{F}(n))$. Thus our condition is a useful way to recognize when an n -tuple of operators generate a free group factor. We obtain as a consequence that the q -deformed free group factors $\Gamma_q(\mathbb{R}^n)$ are isomorphic (for sufficiently small q , with bound depending on n) to free group factors. We also partially prove a conjecture of Voiculescu by showing that free Gibbs states which are small perturbations of a semicircle law generate free group factors. Lastly, we show that entrywise monotone transport maps for certain Gibbs measure on matrices are well-approximated by the matricial transport maps given by free monotone transport.

1. INTRODUCTION.

1.1. **On a notion of density in free probability.** Let X_1, \dots, X_n be classical random variables. Thus $X = (X_1, \dots, X_n)$ can be viewed as a function defined on a measure space (Ω, ω) with values in (say) \mathbb{R}^n . Of special interest is the law of X , which is the measure on \mathbb{R}^n obtained as the push-forward $\mu = X_*\omega$. Very often, it is assumed that μ is Lebesgue absolutely continuous and the density $\rho = d\mu/(\prod dx_j)$ then plays a key role. For example, the density is involved in the classical definition of entropy $(\int \rho \log \rho \prod dx_j)$, Fisher information $(\int \left| \frac{\rho'}{\rho} \right|^2 \prod dx_j)$ and so on.

In passing to the non-commutative case, one assumes that X_1, \dots, X_n are self-adjoint elements of some finite von Neumann algebra (M, τ) , where $\tau : M \rightarrow \mathbb{C}$ is a normal faithful trace. In other words, X_j are self-adjoint operators on some Hilbert space H containing a vector ξ so that $\tau(T) = \langle T\xi, \xi \rangle$ satisfies $\tau(P(X)Q(X)) = \tau(Q(X)P(X))$ for any non-commutative polynomials P and Q evaluated at $X = (X_1, \dots, X_n)$.

The only (partially) satisfactory extension of the notion of joint law of X_1, \dots, X_n to the non-commutative case uses the moment method. More precisely, one defines the non-commutative law of X_1, \dots, X_n as the functional μ which assigns to a non-commutative monomial P the value $\tau(P(X))$.

[†]guionnet@math.mit.edu, CNRS & École Normale Supérieure de Lyon, France and MIT, Department of mathematics, USA. Research supported by ANR-08-BLAN-0311-01 and Simons foundation.

[‡]shlyakht@math.ucla.edu, UCLA Department of Mathematics. Research supported by NSF grants DMS-0900776 and DMS-1161411 and DARPA HR0011-12-1-0009.

Unfortunately, no satisfactory replacement notion of *density* has been obtained so far with the exception of the case $n = 1$. In this case, spectral theory gives a suitable replacement; however, in absence of any commutation between X_1, \dots, X_n there is no satisfactory theory of “joint spectrum” even if X_j ’s all act on a finite-dimensional vector space. This causes a number of problems in free probability theory. For example, Voiculescu introduced two definitions of free entropy (χ and χ^*), but for neither of them is there a simple formula analogous to the classical case as soon as $n \neq 1$.

1.2. Free Gibbs states and log-concave measures. One hint that gives hope that a satisfactory replacement for the notion of density can be found lies in the existence of free analogs of strictly log-concave measures. Log-concave measures on \mathbb{R}^n are probability laws having particularly nice density: it has the form $\exp(-V(x_1, \dots, x_n))dx_1 \cdots dx_n$, where V is a strictly convex function.

It turns out that these laws have free probability analogs: If V is close to a quadratic potential (that is $V(X_1, \dots, X_n) = \frac{1}{2} \sum X_j^2 + \beta W(X_1, \dots, X_n)$ for a fixed polynomial W and sufficiently small β), there exists a unique non-commutative law τ_V which satisfies the “Schwinger-Dyson” equation [Gui06, GMS06]

$$\tau_V(P\mathcal{D}V) = \tau_V \otimes \tau_V(\text{Tr}(\mathcal{J}P)).$$

This law τ_V is called the *free Gibbs law with potential V* . Here \mathcal{D} and \mathcal{J} are suitable non-commutative replacements for the gradient and Jacobian, respectively.

Free Gibbs laws associated to such convex non-commutative functions are very well-behaved. It is worth mentioning that they arise as limits of random matrix models associated to probability measures with log-concave density $\exp(-N\text{Tr}(V(A_1, \dots, A_n)))$ on the spaces of $N \times N$ self-adjoint matrices. Alternatively, the law τ_V can be characterized as the minimizer of the relative free entropy

$$\chi_V(\tau_V) = \chi(\tau_V) - \tau_V(V),$$

where $\chi(\tau_V)$ is Voiculescu’s (microstates) free entropy [Voi94, Voi02b].

Free Gibbs states have many nice properties with respect to free stochastic calculus, various free differential operators and so on (see e.g. [GMS06, GMS07, GS09]). Thus the fact that, without knowing what a non-commutative density is, we are able to single out a class of non-commutative laws which are similar in property to classical log-concave measures (which by definition are measures with a nice density!) strongly suggests that a non-commutative version of density could be found. Moreover, if it were to be found, such log-concave measures are natural first candidates for a detailed study.

1.3. Monotone transportation maps as replacements for densities. There is an alternative way of talking about densities in the classical case, which is a consequence of the following theorem of Brenier: assume that μ is a measure on \mathbb{R}^n satisfying some technical conditions (Lebesgue absolutely continuous, finite second moment, etc.) Let ν denote the standard Gaussian measure on \mathbb{R}^n . Then *there exists a canonical monotone transport map* $\phi_\mu : \mathbb{R}^n \rightarrow \mathbb{R}^n$ *from ν to μ* . Here by a *transport map from ν to μ* we mean a map ϕ satisfying $\phi_*\nu = \mu$; and ϕ is called *monotone* if $\phi = \nabla\psi$ for some convex function ψ (roughly speaking, this means that the Jacobian of ϕ is positive-definite almost everywhere). Our point is that such a transport map carries all the information contained in the knowledge of the density $d\mu / \prod dx_j$ (indeed, this density is essentially the Jacobian of the map). A sign of the fact that ϕ_μ is a good analytic object is that it continues to exist even when the density does

not (for example, when μ is not Lebesgue absolutely continuous, but does not give mass to “small sets”, cf. [Bre91]).

It is thus very tempting to ask if there is any analog of Brenier’s results in the case of non-commutative probability. Here one immediately faces a major obstacle: unlike in the classical case, there are many more non-commutative probability spaces than classical probability spaces. Indeed, up to isomorphism, the unit interval $[0, 1]$ with Lebesgue measure is the unique non-atomic probability space. At the same time, there are many non-isomorphic non-commutative von Neumann algebras. In fact, there is not even a (separable) von Neumann algebra that contains all others [Oza04], thus there cannot exist a “master law” so that all laws can be obtained as “push-forwards” of that law (as is the case classically). However, as we show in the present paper, one may hope to obtain such a result in certain cases (and, indeed, the Brenier map only exists under certain analytic assumptions about the target measure, even in the classical case).

1.4. The main results. In this paper, we give the first examples of existence of non-commutative monotone transport. We thus take the first steps in the study of this subject.

To state our results, recall that the Gaussian law ν can be characterized by the following integration by parts formula: for any $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$, denoting by x the vector (x_1, \dots, x_n) ,

$$\int x \cdot f(x) d\nu(x) = \int \text{Tr}(Jf(x))d\nu(x),$$

where Jf denotes the Jacobian of f . In other words, $x = J^*I$, where J^* denotes the adjoint of J viewed as an unbounded operator $L^2(\mathbb{R}^n, \nu) \rightarrow L^2(M_{n \times n}, \nu)$, and I denotes the $n \times n$ identity matrix.

Recall that a free semicircular family $S = (S_1, \dots, S_n)$ is a free-probability analog of an iid n -tuple of Gaussian random variables. The semicircle law can also be characterized by the formula of the form

$$S = \mathcal{J}^*(I),$$

where \mathcal{J}^* is the adjoint of the map \mathcal{J} , which is a suitable non-commutative replacement for the Jacobian.

Our main theorem shows that monotone transport exists in the non-commutative case, provided that the law of an n -tuple $Z = (Z_1, \dots, Z_n)$ is “close” to the semicircle law, in the sense that it “almost satisfies” the equation $Z = \mathcal{J}^*(I)$, i.e., Z is close to $\mathcal{J}^*(I)$ in a certain analytic function norm $\|\cdot\|_A$ (see §2.1). The conditions on Z automatically imply that the law of Z is a free Gibbs law.

Theorem 1.1 (Existence of monotone transport). *Let $S = (S_1, \dots, S_n)$ be a free semicircular family. If $Z_1, \dots, Z_n \in (M, \tau)$ are non-commutative random variables such that there exists $F = F^*$ in the closure of \mathcal{A} for the norm $\|\cdot\|_A$ so that $\mathcal{J}^*(I) = \mathcal{D}F$ is so that $\|Z - \mathcal{J}^*(I)\|_A$ is sufficiently small, then there exist $Y_1, \dots, Y_n \in W^*(S_1, \dots, S_n)$ so that:*

- the law of $Y = (Y_1, \dots, Y_n)$ is the same as the law of $Z = (Z_1, \dots, Z_n)$, so that $W^*(Z) \cong W^*(Y)$;
- $W^*(S) = W^*(Y)$ ($\cong W^*(Z)$), so that Y_j can be viewed as “non-commutative measurable functions” of S_1, \dots, S_n (we in fact show that these functions can be taken to be non-commutative power series);
- $Y_j = \mathcal{D}_j G$ for some $G \in W^*(S_1, \dots, S_n)$ and $\mathcal{J}Y$ is positive-definite.

Here \mathcal{D} and \mathcal{J} are suitable non-commutative replacements for the gradient and Jacobian of a function; these differential operators play key roles in free probability theory [Voi98, Voi02b, Voi02a, Gui09, Gui06, AGZ10]). In particular, the last condition is a kind of convexity requirement on G (since $\mathcal{J}Y = \mathcal{J}\mathcal{D}G$ is a kind of Hessian of G).

By analogy with Brenier's theorem, we call $F = \mathcal{D}G$ "free monotone transport" from the semicircle law to the law of Z . We can view the function F as playing the role of density; the fact that we are able to construct F in the space of certain analytic maps is a reflection of the fact that the law of Z_1, \dots, Z_n has a "nice density" relative to the free semicircle law.

By Voiculescu's change of variables formula, one can claim that for laws μ satisfying the assumptions of our theorem,

$$\chi(\mu) = \tau \otimes \tau \otimes \text{Tr}(\log \mathcal{J}F_\mu) + \text{universal constant}.$$

We thus see that free entropy can be expressed by a simple formula in terms of the monotone transport map, and is concave in this map.

Our theorem yields a number of isomorphism results, since it gives a rather general condition of when an n -tuple of operators generates a free group factor. In particular, the theorem applies to free log-concave Gibbs laws with potential of the form $\frac{1}{2} \sum X_j^2 + \beta W(X_1, \dots, X_n)$ if β is sufficiently small:

Theorem 1.2. *Let W be a fixed self-adjoint polynomial in n variables and set $V_\beta = \frac{1}{2} \sum_{j=1}^n X_j^2 + \beta W$. Let τ_β be the free Gibbs state with potential V_β . Then for sufficiently small β , $W^*(\tau_\beta) \cong L(\mathbb{F}_n)$.*

This theorem partially answers in the affirmative a conjecture of Voiculescu [Voi06, p. 240] (the full conjecture involves arbitrary values of β).

We also prove:

Theorem 1.3. *Let $\Gamma_q(\mathbb{R}^n)$ be the von Neumann algebra generated by n q -deformed semicircular elements of Bozejko and Speicher. Then there are numbers $q_0(n) > 0$ so that for all $|q| < q_0(n)$, $\Gamma_q(\mathbb{R}^n) \cong L(\mathbb{F}_n)$.*

This settles (for small values of q) the 20-year old question of the isomorphism class of such q -deformed von Neumann algebras. Furthermore, since $L(\mathbb{F}_n)$ is a factor, is non-hyperfinite, strongly solid [OP10] (thus is solid and has no Cartan subalgebras), has the Haagerup property and the complete metric approximation property, (again, for the small values of q) our result can be viewed as the culmination of a number of results about these algebras, including [BS91, BKS97, Nou04, Śni01, Śni04, Shl04, Ric05, Shl09, KN11, Dab10, Avs11].

It would be very interesting to extend our main theorem (which is limited by its hypothesis to laws "close" to the semicircle law), and to study the analytical properties of the map F_μ (these would be non-commutative analogs of properties of the density associated to a classical law μ). For example, F_μ could be used to at least formulate multi-variables analogs of various regularity results in single-variable free probability. Similarly, F_μ could play a prominent role in the study of multi-variable random matrix ensembles.

1.5. Free monotone transport as a limit of classical monotone transport. As we mentioned before, the free Gibbs law τ_V can be obtained as a limit of classical Gibbs measures

in the following sense: for any non-commutative polynomial P ,

$$\tau_V(P) = \lim_{N \rightarrow \infty} \mathbb{E}_{\mu_V^{(N)}} \left[\frac{1}{N} \text{Tr}(P(A_1, \dots, A_n)) \right]$$

where $\mu_V^{(N)}$ is a measure on n -tuples of self-adjoint $N \times N$ matrices given by

$$d\mu_V^{(N)} = \exp(-N \text{Tr} V(A_1, \dots, A_n)) \times \text{Lebesgue measure}.$$

Let W be a fixed self-adjoint non-commutative polynomial and set $V = \frac{1}{2} \sum X_j^2 + W$. If we denote by $\mu^{(N)}$ the Gaussian measure, there exists a unique monotone transport map $F^{(N)}$ which pushes forward $\mu^{(N)}$ into $\mu_V^{(N)}$. This map operates on matrices entrywise. On the other hand, let F be the free monotone transport map taking the semicircle law to the free Gibbs state τ_V . Using functional calculus, F induces for each N a map of n -tuples of $N \times N$ matrices (a kind of “matricial” transport). In Theorem 4.7, we show that $F^{(N)} - F \rightarrow 0$ in L^2 . In other words, the “entrywise” transport is well-approximated by “matricial” transport given by functional calculus on matrices. This is somewhat parallel to Biane’s result [Bia97] which shows that in the Gaussian case, the “entrywise” Segal–Bargmann transform is asymptotically “matricial” and yields the limiting transform in free probability theory.

1.6. The one-dimensional case. In the case that $n = 1$, so that we are dealing with a single random variable, non-commutativity goes away. In this section, we explain the proof of our main result in the case $n = 1$; indeed, many of the formulae that hold in the general multi-variable case have simple ad hoc explanations in the one-variable case. The arguments in this section are subsumed by the arguments for a general n ; however, we feel that their inclusion is justified as they serve to clarify our proof of the general case.

We thus start with a semicircular variable X viewed as the operator of multiplication by x on $L^2([-2, 2], \eta)$, where η is the semicircle law: $d\eta(x) = \chi_{[-2, 2]}(x) \cdot \frac{1}{2\pi} \sqrt{4 - x^2} dx$. We write $\tau(h) = \int h(x) d\eta(x)$. Thus η is the unique maximizer of the functional

$$(1.1) \quad \chi(\mu) = \iint \log |s - t| d\mu(s) d\mu(t) - \int \frac{1}{2} t^2 d\mu(t).$$

We fix a function W which is analytic on the disk of radius A and let $V(x) = \frac{1}{2} x^2 + W(x)$. We now consider measure η_V which is the unique maximizer of

$$\chi_V(\mu) = \iint \log |s - t| d\mu(s) d\mu(t) - \int V(t) d\mu(t).$$

If we write τ_V for the linear functional $\tau_V(h) = \int h(x) d\eta_V(x)$, then τ_V is the free Gibbs state associated to V .

Our goal is to construct a function F which is analytic on a disk of radius A' for some $4 < A' < A$, so that $Y = F(X)$ has law η_V when X has law η . In other words,

$$(1.2) \quad \int h(F(x)) d\eta(x) = \int h(y) d\eta_V(y),$$

or, equivalently, $F_* \eta = \eta_V$.

By definition $\chi_V((x + \delta f)_* \eta_V) \leq \chi_V(\eta_V)$ for all real numbers δ which implies that the unique minimizer η_V satisfies the Schwinger–Dyson equation

$$\int V'(t) f(t) d\eta_V(t) = \iint \frac{f(s) - f(t)}{s - t} d\eta_V(s) d\eta_V(t).$$

If we write $\mathcal{J}f(s, t) = \frac{f(s)-f(t)}{s-t}$, $\mathcal{D}V(x) = V'(x)$ and identify $L^\infty([-2, 2]^2, d\eta_V(s)d\eta_V(t))$ with $L^\infty([-2, 2], d\eta_V) \bar{\otimes} L^\infty([-2, 2], d\eta_V)$, then the Schwinger-Dyson equation reads

$$(1.3) \quad \tau_V(f\mathcal{D}V) = \tau_V \otimes \tau_V(\mathcal{J}f).$$

In the case that τ_V has a connected support (e.g., V is strictly convex on a sufficiently large interval or if W is sufficiently small), the Schwinger-Dyson equation determines τ_V (or, equivalently, η_V) uniquely, see e.g. [BG13].

Note that (1.3) is equivalent to saying that $\mathcal{J}^*(1 \otimes 1) = \mathcal{D}V$. It is not hard to see that $\mathcal{J}^*(1 \otimes 1) = 2 \int \frac{1}{x-y} d\eta_V(y)$ (here and for the remainder of the section, the improper integrals are taken in the sense of principal value). Thus (1.2) is equivalent to

$$2 \int \frac{1}{x-y} d\eta_V(y) = V'(x) \quad \forall x \in \text{supp}(\eta_V).$$

Replacing x by $F(x)$ and y by $F(y)$ and remembering that $F_*\eta = \eta_V$ gives us the following equation for F :

$$2 \int \frac{1}{F(x) - F(y)} d\eta(y) = V'(F(x)) = F'(x) + W'(F(x)), \quad \forall x \in [-2, 2].$$

To solve this equation, we try to find a solution of the form $F(x) = x + f(x)$ with f small if W is. Then the equation becomes

$$(1.4) \quad 2 \int \frac{1}{x-y} \frac{1}{1 + \mathcal{J}f(x, y)} d\eta(y) = x + f(x) + W'(x + f(x)),$$

where, as before, we use the notation $\mathcal{J}f(x, y) = \frac{f(x)-f(y)}{x-y}$.

Using the fact that η is the semicircle law, so that $2 \int \frac{1}{x-y} d\eta(x) = x$, we find that (1.4) is equivalent to

$$-2 \int \frac{1}{x-y} \frac{\mathcal{J}f(x, y)}{1 + \mathcal{J}f(x, y)} d\eta(y) = f(x) + W'(x + f(x)).$$

To deal with the difference quotient $\mathcal{J}f(x, y) = \frac{f(x)-f(y)}{x-y}$, we add and subtract the term $f'(x)$:

$$\begin{aligned} & 2 \int \frac{1}{x-y} (f'(x) - \mathcal{J}f(x, y)) \frac{1}{1 + \mathcal{J}f(x, y)} d\eta(y) \\ &= f'(x) \left[2 \int \frac{1}{x-y} \frac{1}{1 + \mathcal{J}f(x, y)} d\eta(x) \right] + f(x) + W'(x + f(x)). \end{aligned}$$

We can substitute (1.4) into the right hand side, to get

$$(1.5) \quad \begin{aligned} & 2 \int \frac{1}{x-y} (f'(x) - \mathcal{J}f(x, y)) \frac{1}{1 + \mathcal{J}f(x, y)} d\eta(y) \\ &= f'(x) [x + f(x) + W'(x + f(x))] + f(x) + W'(x + f(x)) \\ &= (xf)'(x) + \frac{1}{2}(f(x)^2)' + [W(x + f(x))]' \end{aligned}$$

The resulting equation is equivalent to the previous one under the assumption that $f'(x) \neq -1$ almost surely.

Since the right-hand side is a full derivative, it is natural to try to express the left-hand side as a full derivative, too. Simple algebra shows that

$$\int \frac{1}{x-y} (f'(x) - \mathcal{J} f(x, y)) \frac{1}{1 + \mathcal{J} f(x, y)} d\eta(y) = \partial_x \int \log(1 + \mathcal{J} f(x, y)) d\eta(y).$$

Substituting this into (1.5) and removing derivatives from both sides finally gives us

$$2 \int \log(1 + \mathcal{J} f(x, y)) d\eta(y) = (xf)(x) + \frac{1}{2}(f(x))^2 + W(x + f(x)) + \text{const.}$$

Finally, we seek g so that $f = g'$ solves the previous equation. This gives us the following equation for g :

$$(1.6) \quad xg'(x) = -W(x + g'(x)) - \frac{1}{2}(g'(x))^2 + 2 \int \log \left(1 + \frac{g'(x) - g'(y)}{x - y} \right) d\eta(y) + \text{const.}$$

The constant can be fixed by requiring that both sides of the equation vanish at $x = 0$. Note that the operator $\mathcal{N}g = xg'$ multiplies monomials of degree n by n . Let Σ be the inverse of \mathcal{N} (defined on polynomials with zero constant term), $\Sigma(\hat{g})(x) = \int_0^x y^{-1} \hat{g}(y) dy$, and let $\hat{g} = \mathcal{N}g$ so that $g = \Sigma\hat{g}$. With this notation, our equation becomes

$$\hat{g}(x) = -W(x + (\Sigma\hat{g})'(x)) - \frac{1}{2}((\Sigma\hat{g})(x))^2 + 2 \int \log \left(1 + \frac{(\Sigma\hat{g})'(x) - (\Sigma\hat{g})'(y)}{x - y} \right) d\eta(y) + \text{const.}$$

We can now rewrite this equation as a fixed point equation

$$\hat{g}(x) = \Psi_x(\hat{g}(x)).$$

Note that $\Psi_x(0) = W(x)$. For W small, this equation can be solved by iteration in the space of analytic functions converging on a large enough disk. The essential fact here is that $\hat{g} \mapsto \Psi_x(\hat{g})$ is Lipschitz for a certain analytic norm for any x . This gives us a function \hat{g} , which is sufficiently small. Retracing our steps (and noting that for sufficiently small W , f is small, so that $|f'(x)| < 1$ almost surely), we get back the desired map F given by $F = x + f = G'$ if $G = \frac{1}{2}x^2 + g(x)$.

It is rather straightforward that if we consider $V_\beta = \frac{1}{2}x^2 + W_\beta(x)$ for some family W_β analytically depending on β and sufficiently small, then our transport map $F = F_{V_\beta}$ also depends on β analytically.

It is worth noting that for W sufficiently small, the map $F(x) = x + f(x)$ has a positive derivative and is thus monotone. It is therefore the unique monotone map satisfying $F_*\eta = \eta_V$. In particular, the map F gives optimal transport between η and η_V .

It was pointed out to us by Y. Dabrowski that equation (1.6) can be regarded as a free analog of the classical Monge-Ampère equation. To see this, put $V(x) = \frac{1}{2}x^2$ and $\tilde{V} = V + W$, and consider the classical Monge-Ampère equation for the monotone transport map H from the Gibbs measure $d\mu = \exp(-V(x))dx$ to $d\tilde{\mu} = \exp(-\tilde{V}(x))dx$. Let JH stand for the Jacobian (derivative) of H . Then the Monge-Ampère equation reads:

$$\det(JH(x)) = \frac{\exp(-V(x))}{\exp(-\tilde{V}(H(x)))} = \exp(\tilde{V}(H(x)) - V(x)).$$

(We write $\det JH$ even though JH is a 1×1 matrix to make the analogy clearer). Taking logarithm of both sides and using that $\log \det A = \text{Tr}(\log A)$ if $A \geq 0$ gives

$$\text{Tr}(\log JH) = \tilde{V}(H(x)) - V(x).$$

This is quite reminiscent of (1.6), if we rewrite it by setting $f = g'$, $F(x) = x + f(x)$ and $\mathcal{J}F = \frac{F(x)-F(y)}{x-y} = 1 + \mathcal{J}f$:

$$2 \int \log(\mathcal{J}F) d\eta(y) = \left\{ \frac{1}{2}(F(x))^2 + W(F(x)) \right\} - \frac{1}{2}x^2 = \tilde{V}(F(x)) - V(x).$$

Later in the paper, we also consider the n -variable version of (1.6), which has the form

$$(1 \otimes \tau + \tau \otimes 1)Tr \log \mathcal{J}F = \mathcal{S} \left[\left\{ \frac{1}{2} \sum F(X)_j^2 + W(F(X)) \right\} - \frac{1}{2} \sum X_j^2 \right]$$

where \mathcal{S} is a certain symmetrization operator.

1.7. Organization of paper. To prove our theorem, we consider in section §2 a certain space of analytic functions in several non-commuting variables. This is the space in which our construction of the monotone transport map takes place. We also discuss the various differential operators which are the non-commutative replacements for gradients and Jacobians. We finish the section with a kind of implicit function theorem for non-commutative analytic functions.

The next section is devoted to the construction of a non-commutative monotone transport map from the semicircle law τ to the law τ_V satisfying the Schwinger-Dyson equation with $V = \frac{1}{2} \sum X_j^2 + W$. We assume that W is a given non-commutative analytic function (which we assume to be sufficiently small in a certain norm).

The next section, §4, is devoted to applications. Finally, the last section collects some open questions and problems.

Acknowledgments. The authors are grateful to Yoann Dabrowski for many useful comments and discussions. We also wish to mention that the idea of looking for a free analog of *optimal* transport (although via some duality arguments) was considered some 10 years ago by Cédric Villani and the authors; although the precise connection between *optimal* and *monotone* transport is still missing for the moment, such considerations have been an inspiration for the present work. We also thank the anonymous referees for their numerous comments which helped us to greatly improve our article.

2. NOTATION AND AN IMPLICIT FUNCTION THEOREM.

2.1. Non-commutative polynomials, power series and norms $\|\cdot\|_A$. We will denote by $\mathcal{A} = \mathbb{C}\langle X_1, \dots, X_n \rangle$ the algebra of non-commutative polynomials in n variables. Let $\mathcal{A}_0 \subset \mathcal{A}$ be the linear span of polynomials with zero constant term.

Following [GMS07], we consider on this space the family of norms $\|\cdot\|_A$, $A > 1$, defined as follows. For a monomial q and arbitrary $P \in \mathbb{C}\langle X_1, \dots, X_n \rangle$, let $\lambda_q(P)$ be the coefficient of q in the decomposition of P as monomials; thus $P = \sum_q \lambda_q(P)q$. Then we set

$$\|P\|_A = \sum_{q: \deg q \geq 0} |\lambda_q(P)| A^{\deg q}.$$

We'll denote by $\mathcal{A}^{(A)}$ (resp., $\mathcal{A}_0^{(A)}$) the completion of \mathcal{A} (resp., \mathcal{A}_0) with respect to the norm $\|\cdot\|_A$. This is a Banach algebra, and can be viewed as the algebra of absolutely convergent power series with radius of convergence at least A .

Note that the norm $\|\cdot\|_A$ has the following property: whenever T_1, \dots, T_n are elements of some Banach algebra Q and $\|T_j\|_Q \leq A$, there exists a contractive map $\mathcal{A}^{(A)} \rightarrow Q$

sending X_j to T_j . Hence, in the following we will see variables as elements of the set of polynomials $\mathcal{A}^{(A)}$. However, $\mathcal{A}^{(A)}$ will still also denote sets of maps, namely the completion of polynomials in the norm $\|\cdot\|_A$.

\mathcal{A}^{op} is the opposite algebra, with multiplication defined by $a^{op} \cdot b^{op} = (ba)^{op}$. It is equipped with the same norm $\|\cdot\|_A$.

2.2. The operators \mathcal{N} , Σ , Π , \mathcal{J} , ∂ and \mathcal{D} . Let us denote by \mathcal{N} the linear operator on $\mathcal{A} = \mathbb{C}\langle X_1, \dots, X_n \rangle$ that multiplies a degree k monomial by k . Let us also denote by

$$\Sigma : \mathcal{A}_0 \rightarrow \mathcal{A}_0$$

the inverse of \mathcal{N} pre-composed by the projection Π onto \mathcal{A}_0 , given by

$$\Pi : P \mapsto P - P(0, 0, \dots, 0).$$

If $g \in \mathcal{A}$ we write $\mathcal{D}_j g$ for the j -th cyclic derivative of g [Voi02a]. For a monomial q , $\mathcal{D}_j q = \sum_{q=AX_j B} BA$.

Note that $\mathcal{D}_j g = \mathcal{D}_j \Pi g$. We'll denote by $\mathcal{D}g$ the cyclic gradient (viewed as a vector):

$$\mathcal{D}g = (\mathcal{D}_j g)_{j=1}^n.$$

If $f = (f_1, \dots, f_n)$ with $f_j \in \mathcal{A}$ then we'll write $\mathcal{J}f \in M_{n \times n}(\mathcal{A} \otimes \mathcal{A}^{op})$ for the matrix given by

$$\mathcal{J}f = (\partial_j f_i)_{i,j=1}^n,$$

where ∂_j is the j -th free difference quotient [Voi98] defined as the derivation from \mathcal{A} to $\mathcal{A} \otimes \mathcal{A}^{op}$ satisfying $\partial_j X_i = \delta_{j=i}$, that is if q is a monomial $\partial_j q = \sum_{q=AX_j B} A \otimes B$.

2.3. Notation: Tr , \mathcal{J}^* and $\#$. We will often assume that τ is a positive trace and also that $1 \otimes 1 \in \text{dom } \partial_j^*$, when $\partial_j : L^2(\mathcal{A}, \tau) \rightarrow L^2(\mathcal{A}, \tau) \bar{\otimes} L^2(\mathcal{A}, \tau)$ is viewed as a densely defined operator. Under this assumption, $\mathcal{A} \otimes \mathcal{A}^{op}$ belongs to the domain of ∂_j^* [Voi98]. We set for $q = \sum a_i \otimes b_i \in \mathcal{A} \otimes \mathcal{A}^{op}$ and $g \in \mathcal{A}$,

$$q \# g = \sum_i a_i g b_i.$$

For $q \in M_{n \times n}(\mathcal{A} \otimes \mathcal{A}^{op})$, and $g = (g_j)_{j=1}^n \in \mathcal{A}^n$, we'll write

$$\begin{aligned} Tr(q) &= \sum_i q_{ii} \in \mathcal{A} \otimes \mathcal{A}^{op}, \\ \mathcal{J}^* q &= \left(\sum_i \partial_i^*(q_{ji}) \right)_{j=1}^n \in L^2(\mathcal{A}, \tau)^n, \\ q \# g &= \left(\sum_i q_{ji} \# g_i \right)_{j=1}^n. \end{aligned}$$

We will also denote by $\#$ multiplication in $M_{n \times n}(\mathcal{A} \otimes \mathcal{A}^{op})$ and write $g \# h = \sum g_i h_i$ if $g, h \in \mathcal{A}^n$.

2.4. Non-commutative notions of transport. Let (M, ϕ) be a von Neumann algebra with trace ϕ , and let $X_1, \dots, X_n \in M$ be self-adjoint elements generating M ; thus M can be regarded as a completion of the algebra \mathcal{A} , and the trace ϕ induces on \mathcal{A} a trace τ_X (called the non-commutative law of X_1, \dots, X_n). We write $M \cong W^*(\tau_X)$. Let $Y_1, \dots, Y_n \in N$ be self-adjoint elements generating another von Neumann algebra N with trace ψ ; we denote by τ_Y the corresponding linear functional on \mathcal{A} so that $N \cong W^*(\tau_Y)$.

Definition 2.1. By *transport* from τ_X to τ_Y we mean an n -tuple of self-adjoint elements $\hat{Y}_1, \dots, \hat{Y}_n \in M$ having the same law as Y_1, \dots, Y_n . We call such transport *monotone* if $\hat{Y} = (\hat{Y}_j)_j$ belongs to the L^2 -closure of the set

$$\{\mathcal{D}g : g \in \text{Alg}(X_1, \dots, X_n) \text{ and } \mathcal{J} \mathcal{D}g \geq 0\}.$$

(as part of the definition of monotone transport, we are making the assumption that the variables X_1, \dots, X_n are algebraically free, and so $\mathcal{J} \mathcal{D}g$ is well-defined for any g in the algebra these variables generate).

When M is abelian, an element of M is an (essentially bounded) function of X_1, \dots, X_n (and $n = 1$.) So in the case that M is abelian, our definition reduces to the statement that the random variables $\hat{Y}_1, \dots, \hat{Y}_n$ are expressed as (bounded measurable) functions $\hat{Y}_j = f_j(X_1, \dots, X_n)$. Since the law of Y is the same as the law of \hat{Y} , the push-forward of the measure τ_X via the function $f = (f_j)_j$ is exactly the measure τ_Y . So our notion of transport coincides with the classical one in the commutative case.

Our definition of monotone transport is analogous to the classical case where one requires that \hat{Y} belongs to the closure of the space of gradients of convex functions. However, note that a different notion of derivative is being used. Nonetheless, when $n = 1$ our requirement reduces to asking that $\hat{Y} = f(X)$ for some $f : \mathbb{R} \rightarrow \mathbb{R}$ so that $f = g'$ for some g and $\frac{f(x)-f(y)}{x-y} \geq 0$, i.e., f is monotone and g convex.

Note that if a transport from τ_X to τ_Y exists, then there exists a trace-preserving embedding $N = W^*(Y)$ into $M = W^*(X)$ given by $Y_j \mapsto \hat{Y}_j$.

2.5. The norm $\|\cdot\|_{A \otimes_{\pi} A}$. For $F = (F_j)_{j=1}^n \in (\mathcal{A}^{(A)})^n$, we will use the notation $\|F\|_A = \sup_j \|F_j\|_A$.

Let us denote by $\|\cdot\|_{A \otimes_{\pi} A}$ the projective tensor product norm on $\mathcal{A}^{(A)} \otimes (\mathcal{A}^{(A)})^{op}$. In other words

$$\left\| \sum a_i \otimes b_i \right\|_{A \otimes_{\pi} A} = \sup_{PI} \|PI(\sum a_i \otimes b_i)\|$$

where the supremum is taken over all maps PI valued in a Banach algebra so that $PI(1 \otimes b)$ and $PI(a \otimes 1)$ commute and have norm bounded by the norm of b and a respectively. In particular, taking PI to be given by the right and left multiplication respectively, we see that for all $q \in \mathcal{A} \otimes \mathcal{A}^{op}$ and $g \in \mathcal{A}$, we have

$$\|q \# g\|_A \leq \|q\|_{A \otimes_{\pi} A} \|g\|_A.$$

$\|\cdot\|_{A \otimes_{\pi} A}$ also extends to $(\mathcal{A} \otimes \mathcal{A}^{op})^n$ by putting for $F = (F_1, \dots, F_n) \in (\mathcal{A} \otimes \mathcal{A}^{op})^n$

$$\|F\|_{A \otimes_{\pi} A} = \max_{1 \leq i \leq n} \|F_i\|_{A \otimes_{\pi} A}.$$

We will also denote by the same symbol the norm on $M_{n \times n}(\mathcal{A} \otimes \mathcal{A}^{op})$ (or its completion) given by identifying that space with the Banach space of (left) multiplication operators on

$(\mathcal{A} \otimes \mathcal{A}^{op})^n$. Explicitly, this norm is given by

$$\|X_{ij}\|_{A \otimes \pi A} = \max_i \sum_j \|X_{ij}\|_{A \otimes \pi A}.$$

In consequence of its definition, we see that the maps

$$\begin{aligned} \# : M_{n \times n}(\mathcal{A} \otimes \mathcal{A}^{op}) \times M_{n \times n}(\mathcal{A} \otimes \mathcal{A}^{op}) &\rightarrow M_{n \times n}(\mathcal{A} \otimes \mathcal{A}^{op}) \\ M_{n \times n}(\mathcal{A} \otimes \mathcal{A}^{op}) \times \mathcal{A}^n &\rightarrow \mathcal{A}^n \end{aligned}$$

are contractive for the norm $\|\cdot\|_{A \otimes \pi A}$ on $M_{n \times n}(\mathcal{A} \otimes \mathcal{A}^{op})$ and the norm $\|\cdot\|_A$ on \mathcal{A}^n .

2.6. Cyclically symmetric functions.

Definition 2.2. We say that a polynomial q in \mathcal{A} is *cyclically symmetric* if for any i_1, \dots, i_k , the coefficient of the monomial $X_{i_1} \cdots X_{i_k}$ in q is the same as the coefficient of the monomial $X_{i_k} X_{i_1} \cdots X_{i_{k-1}}$.

Let \mathcal{S} denote the operator on \mathcal{A}_0 defined on monomials by

$$\mathcal{S} X_{i_1} \cdots X_{i_p} = \frac{1}{p} \sum_{r=1}^p X_{i_{r+1}} \cdots X_{i_p} X_{i_1} \cdots X_{i_r}.$$

Clearly, \mathcal{S} sends \mathcal{A}_0 to the space of cyclically symmetric operators. Note also that $\mathcal{S}g = \Sigma \mathcal{D}g \# X = \Sigma \sum_i \mathcal{D}_i g X_i$. Moreover, one clearly has that $\|\mathcal{S}g\|_A \leq \|g\|_A$. Finally, note that $\mathcal{D}(\mathcal{S}g) = \mathcal{D}g$.

2.7. Implicit function theorem. We next establish a kind of implicit function theorem for elements of the space $(\mathcal{A}^{(A)})^n$. For later purpose note that if $F \in \mathcal{A}^{(A)}$, F is infinitely differentiable in $\mathcal{A}^{(A')}$ for any $A' < A$. These results are most likely folklore, but we were not able to find a precise reference. The proofs follow very closely the classical proof of the implicit function theorem using a contraction mapping principle. We have denoted by \mathcal{A} the algebra $\mathbb{C}\langle X_1, \dots, X_n \rangle$ of non-commutative polynomials in n variables. Let us write $\mathcal{A} * \mathcal{A}$ for the algebra of non-commutative polynomials in $2n$ variables, $\mathbb{C}\langle X_1, \dots, X_n, Y_1, \dots, Y_n \rangle$. Also, remember that variables and maps are all seen as converging power series.

Theorem 2.3 (Implicit Function Theorem). *Assume that $F = (F_j)_{j=1}^n$ where*

$$F_j(X_1, \dots, X_n, Y_1, \dots, Y_n) \in (\mathcal{A} * \mathcal{A})^{(A)}$$

is a non-commutative power series in $2n$ variables, and assume that

$$F(0, \dots, 0, 0, \dots, 0) = 0.$$

Let $Q = \mathcal{J}_2 F(0, \dots, 0)$ (here \mathcal{J}_2 refers to differentiation in the variables Y_1, \dots, Y_n , with X_j 's held fixed), and assume that Q is invertible. Then there exists an $a_0 \in (0, A)$ and a finite constant C so that for all $a \in (0, a_0)$, any $U \in \mathcal{A}^{(A)}$ satisfying $\|U\|_A < a$, there exists a unique $V \in (\mathcal{A}^{(A)})^n$ satisfying $\|V\|_A < Ca$ which solves the equation

$$F(U, V) = 0.$$

Moreover, there exists a function $f \in (\mathcal{A}^{(a)})^n$ so that $V = f(U)$.

Proof. Put $X = (X_1, \dots, X_n)$ and $Y = (Y_1, \dots, Y_n)$ and denote

$$G(X, Y) = Y - Q^{-1} \# F(X, Y).$$

Then, since Q is invertible,

$$G(U, V) = V \iff F(U, V) = 0.$$

Moreover, since $F(0, 0) = 0$, we have

$$G(X, Y) = R(X) + H(X, Y)$$

with $R(X) = -Q^{-1} \# \mathcal{J}_1 F(0, \dots, 0) \# X$ and $H(X, Y) = G(X, Y) - R(X)$ a non-commutative power series in $[(\mathcal{A} * \mathcal{A})^{(A)}]^n$ whose decomposition in monomials contains only monomials of degree two or greater. Therefore, for any X, X' so that $\|X\|_A, \|X'\|_A \leq a$ and any Y, Y' so that $\|Y\|_A \leq b, \|Y'\|_A \leq b$, with $a, b \leq A' < A$, we have

$$(2.1) \quad \|H(X, Y)\|_A \leq C(A')(a + b)^2$$

$$(2.2) \quad \|H(X, Y) - H(X', Y')\|_A \leq C(A')(a + b)(\|X - X'\|_A + \|Y - Y'\|_A)$$

with a bounded constant $C(A')$ increasing in $A' < A$.

Choose now $0 < b_0 < A'$ so that $C(A')b_0^2 < (1/2)b_0$, and then choose $0 < a_0 < b_0/3$ so that $C(A')a_0 < 1/2$, $\|Q^{-1} \# \mathcal{J}_1 F(0, \dots, 0)\|_{A \otimes \pi A} a_0 + C(A')a_0^2 < (1/2)b_0$. These choices imply that $C(A')(a_0^2 + 2a_0b_0) < (1/2)b_0$.

We assume that U is given and satisfies $\|U\|_A \leq a < a_0$ and construct a solution by putting $V_{k+1} = G(U, V_k)$ and $V_0 = 0$.

We claim that $\|V_{k+1}\|_A < b_0$ for all k . Indeed, $\|V_0\|_A = 0 \leq b_0$ and we have by (2.1) and our choices of a_0, b_0 that if $\|V_k\|_A \leq b_0$,

$$\|V_{k+1}\|_A \leq \|R(U)\|_A + \|H(U, V_k)\|_A \leq \|Q^{-1} \# \mathcal{J}_1 F(0, \dots, 0)\|_{A \otimes \pi A} a + C(A')(a + b_0)^2 < b_0.$$

Now by (2.2), we get for all $k \geq 1$,

$$\|V_{k+1} - V_k\|_A \leq C(A')(a + b)\|V_k - V_{k-1}\|_A$$

and hence V_k is a Cauchy sequence since our choice implies that $C(A')(a + b) \leq C(A')(a_0 + b_0) < 1$. Therefore it converges whenever $a \leq a_0$. The limit V_* satisfies $F(U, V_*) = 0$. Furthermore, $\|V_*\|_A$ is bounded above by b_0 . This proves the stated existence.

This solution is unique: indeed, (2.2) guarantees that any two solutions V_* and V'_* would satisfy

$$\|V_* - V'_*\|_A \leq C(A')(a + b)\|V_* - V'_*\|_A$$

implying that $V_* = V'_*$, since $C(A')(a + b) < 1$.

Finally, by induction we see that for all k , $V_k = f_k(U)$ where $f_k(X) = G(X, f_{k-1}(X))$ and $f_0(X) = 0$. As $G \in [(\mathcal{A} * \mathcal{A})^{(A)}]^n$, f_k belongs to $(\mathcal{A}^{(A)})^n$ for all k so that $\|f_{k-1}\|_A \leq A$. Moreover, we see as above that for $a < a_0$, for all $k \geq 0$,

$$\|f_k\|_a \leq b_0 < A$$

so that

$$\begin{aligned} \|f_k - f_{k-1}\|_a &= \|G(X, f_{k-1}(X)) - G(X, f_{k-2}(X))\|_a \\ &= \|H(X, f_{k-1}(X)) - H(X, f_{k-2}(X))\|_a \leq C(A')(a + b)\|f_{k-1} - f_{k-2}\|_a \end{aligned}$$

and (because $C(A')(a + b) < 1$) we conclude that f_k converges to a limit f in $(\mathcal{A}^{(a)})^n$. Hence, $V = f(U)$ with $f \in (\mathcal{A}^{(a)})^n$. \square

2.8. Inverses to non-commutative power series. The following corollary shows that an “absolutely convergent non-commutative power series” has an “absolutely convergent inverse”:

Corollary 2.4. *Let $A' < A < B$ and consider the equation $Y = X + f(X)$ with $f \in (\mathcal{A}^{(B)})^n$ and $\|Y\|_{A'} \leq A$. Then there exists a constant $C > 0$, depending only on A, A' and B , so that whenever $\|f\|_B < C$, then there exists $G \in (\mathcal{A}^{(A)})^n$ so that $X = G(Y)$.*

Proof. We repeat the arguments of the previous proof. We define a sequence of maps

$$G_k = \text{id} - f \circ G_{k-1}$$

with $G_0 = \text{id}$. By definition $\|G_0\|_A \leq A$ and for all k so that $\max_{\ell \leq k} \|G_\ell\|_A \leq B' < B$, there exists a finite constant $C(B')$ depending only on B', B and $\|f\|_B$ (for example, $C(B') \leq \|f\|_B \max_{k \geq 0} k(B')^{k-1} B^{-k}$), so that

$$\|G_{k+1} - G_k\|_A \leq C(B') \|G_k - G_{k-1}\|_A.$$

Since $\|G_1 - G_0\|_A \leq \|f\|_A < C$, we may iterate to deduce that that

$$\|G_{k+1} - G_k\|_A \leq (C(B'))^k \|f\|_A,$$

which implies that if C is small enough so that $C(B') < 1$

$$\|G_{k+1} - G_0\|_A \leq \frac{1}{1 - C(B')} \|f\|_A \leq \frac{C}{1 - C(B')}.$$

Hence, since $\|G_0\|_A \leq A$, if C is small enough so that we can find $B' \in (A, B)$ so that

$$A + \frac{C}{1 - C(B')} \leq B'$$

we are guaranteed that for all k , $\|G_k\|_A \leq B'$. We hereafter choose C so that this is satisfied and conclude that G_k converges in $(\mathcal{A}^{(A)})^n$. Its limit G satisfies $\|G\|_A \leq B'$ and

$$G = \text{id} - f \circ G.$$

Moreover since $\|Y\|_{A'} \leq A$ and $\|X\|_{A'} \leq \|X\|_A \leq A$, we deduce from

$$X + f(X) = Y = G(Y) + f(G(Y))$$

that

$$\|X - G(Y)\|_{A'} \leq \|f(X) - f(G(Y))\|_{A'} \leq C(B') \|X - G(Y)\|_{A'}$$

which ensures that $X = G(Y)$ since we assumed $C(B') < 1$. □

3. CONSTRUCTION OF THE FREE MONOTONE TRANSPORT MAP.

3.1. Outline of the proof of the main result. To prove existence of the transport map, we start with an n -tuple of free semicircular elements X_1, \dots, X_n . We also fix V of the form $V = \frac{1}{2} \sum X_j^2 + W$ where W is a “non-commutative analytic function” of X_1, \dots, X_n which is sufficiently small in $\|\cdot\|_A$.

We then seek to find some elements Y_1, \dots, Y_n in the von Neumann algebra M generated by X_1, \dots, X_n so that Y_1, \dots, Y_n have law τ_V , which is the unique log-concave free Gibbs law satisfying the Schwinger-Dyson equation with potential V , i.e.

$$(3.1) \quad \partial_{Y_j}^*(1 \otimes 1) = Y_j + \mathcal{D}_{Y_j}(W(Y_1, \dots, Y_n))$$

To construct $Y = (Y_1, \dots, Y_n)$ we seek $f = (f_1, \dots, f_n)$ so that $Y = X + f$ (i.e., $Y_j = X_j + f_j$). Assuming we could find such Y with f_j analytic and of small norm, the Schwinger-Dyson equation (3.1) for Y becomes an equation on f , of the form

$$\mathcal{J}^* \left(\frac{1}{1 + \mathcal{J}f} \right) = X + f + (\mathcal{D}W)(X + f).$$

(Here \mathcal{J}^* is the adjoint to \mathcal{J}). Since for the semicircle law, $\mathcal{J}^*(1) = X$, this equation further simplifies as

$$\mathcal{J}^* \left(\frac{\mathcal{J}f}{1 + \mathcal{J}f} \right) + f + (\mathcal{D}W)(X + f) = 0.$$

One is tempted to solve this equation by using the contraction mapping principle; however, one faces the obstacle of bounding the map $f \mapsto \mathcal{J}^*(\mathcal{J}f/(1 + \mathcal{J}f))$. Fortunately, under the assumption that $f = \mathcal{D}g$ and further assumptions on invertibility of $1 + \mathcal{J}f$, one is able to rewrite the left hand side of the equation as the cyclic gradient of a certain expression in g . This gives us an equation for g . It turns out that then one can use a contraction mapping argument to prove the existence of a solution of this equation for g ; this is the content of Proposition 3.13. We then retrace our steps, going from g to $f = \mathcal{D}g$ and then to $Y = X + f$, proving the main result of this section, Theorem 3.15.

To avoid confusion, we will use the following convention: all differential operators (\mathcal{D} , \mathcal{J} , ∂) which either have no indices, or have a numeric index, refer to differentiation with respect to X_1, \dots, X_n . Operators that involve differentiation with respect to Y_1, \dots, Y_n will be labeled as \mathcal{D}_{Y_j} , ∂_{Y_j} , etc.

3.2. A change of variables formula.

Lemma 3.1. *Put $M = W^*(\mathcal{A}, \tau)$. Assume that Y is such that $\mathcal{J}Y = (\partial_{X_j} Y_i)_{ij} \in M_n(M \bar{\otimes} M^{op})$ is bounded and invertible. Assume further that $1 \otimes 1$ belongs to the domain of ∂_j^* for all j .*

(i) *Define*

$$\hat{\partial}_j(Z) = \sum_{i=1}^n \partial_{X_i}(Z) \# ((\mathcal{J}Y)^{-1})_{ij}$$

where $\#$ denotes the multiplication

$$(a \otimes b) \# (A \otimes B) = aA \otimes Bb.$$

Then $\hat{\partial}_j = \partial_{Y_j}$ and $\partial_{Y_j} X_i = ((\mathcal{J}Y)^{-1})_{ij}$.

(ii) $\partial_{Y_j}^*(1 \otimes 1) = \sum_{\ell} \partial_{X_\ell}^* ((\mathcal{J}Y)^{-1})_{\ell j}^*$.

(iii) *Assume in addition that $Y_j = \mathcal{D}_j G$ for some G in the completion of \mathcal{A} with respect to $\|\cdot\|_A$. Assume that $G = G^*$. Let $(a \otimes b)^\dagger = b^* \otimes a^*$. Then $(\mathcal{J}Y)_{ij}^\dagger = (\mathcal{J}Y)_{ij}$ and $(\mathcal{J}Y)^{-1}$ is self-adjoint in $M_n(M \bar{\otimes} M^{op})$.*

Proof. Let $Q = \mathcal{J}Y$. (i) We verify that $\hat{\partial}_j Y_k = \sum_i \partial_{X_i} Y_k \# (Q^{-1})_{ij} = \sum_i Q_{ki} \# (Q^{-1})_{ij} = 1_{k=j} 1 \otimes 1$

To see (ii), we compute

$$\begin{aligned}
\langle \partial_{Y_j}^*(1 \otimes 1), X_{i_1} \dots X_{i_p} \rangle &= \langle 1 \otimes 1, \partial_{Y_j}(X_{i_1} \dots X_{i_p}) \rangle \\
&= \sum_{\ell} \langle 1 \otimes 1, \partial_{X_{\ell}}(X_{i_1} \dots X_{i_p}) \# (Q^{-1})_{\ell j} \rangle \\
&= \sum_{\ell} \langle (Q^{-1})_{\ell j}^*, \partial_{X_{\ell}}(X_{i_1} \dots X_{i_p}) \rangle \\
&= \langle \sum_{\ell} \partial_{X_{\ell}}^*(Q^{-1})_{\ell j}^*, X_{i_1} \dots X_{i_p} \rangle
\end{aligned}$$

To verify (iii), we first claim that if $Y = \mathcal{D}G$ for some G (not necessarily self-adjoint), then

$$(3.2) \quad (\mathcal{J}Y)_{ij} = \sigma[(\mathcal{J}Y)_{ji}],$$

where $\sigma(a \otimes b) = b \otimes a$. Indeed, let $G = X_{i_1} \dots X_{i_k}$; in this case

$$\partial_i Y_j = \sum_{G=AX_i BX_j C} CA \otimes B + \sum_{G=AX_j BX_i C} B \otimes CA.$$

Reversing the role of i and j amounts to switching the two sums, i.e., applying σ .

We next consider G of the form

$$G = X_{i_1} \dots X_{i_k} + X_{i_k} \dots X_{i_1} = P + P^*,$$

where $P = X_{i_1} \dots X_{i_k}$. In that case

$$Y_j = \sum_{P=AX_j B} BA + A^* B^*$$

and therefore we have

$$(\mathcal{J}Y)_{ji} = \partial_i Y_j = \sum_{P=AX_j B} \sum_{BA=RX_i S} R \otimes S + S^* \otimes R^*.$$

Thus, if we denote by $*$ the involution $(a \otimes b)^* = a^* \otimes b^*$, we deduce that

$$(3.3) \quad (\mathcal{J}Y)_{ij}^* = \sigma((\mathcal{J}Y)_{ij}) = (\mathcal{J}Y)_{ji}$$

proving that $\mathcal{J}Y$ is self-adjoint. Moreover we find that

$$(3.4) \quad ((\mathcal{J}Y)_{ij})^\dagger = \sigma((\mathcal{J}Y)_{ij}^*) = \sigma((\mathcal{J}Y)_{ji}) = (\mathcal{J}Y)_{ij}.$$

where the last two equalities come from (3.3). □

A direct consequence of Lemma 3.1 (ii) and (iii), is the following.

Corollary 3.2. *Assume that $g \in \mathcal{A}^{(A)}$, $1 \otimes 1 \in \text{dom } \partial_j^*$ for all j , and put $G = \frac{1}{2} \sum X_j^2 + g$. Let $f_j = \mathcal{D}_j g$ and $Y_j = X_j + f_j$ so that $Y = \mathcal{D}G$. Assume that $1 + \mathcal{J}f$ is invertible. Then Equation (3.1) is equivalent to the equation*

$$(3.5) \quad \mathcal{J}^* \left(\frac{1}{1 + \mathcal{J}f} \right) = X + f + (\mathcal{D}W)(X + f).$$

3.3. An equivalent form of Equation (3.5).

Lemma 3.3. *Let τ be the semicircle law. Assume that the map $\xi \mapsto (\mathcal{J}f)\#\xi + \xi$ is invertible on $(\mathcal{A}^{(A)})^n$, and that $f = \mathcal{D}g$ for some $g = g^*$. Let*

$$K = -\mathcal{J}^* \circ \mathcal{J} - \text{id}.$$

Then equation (3.5) is equivalent to

$$(3.6) \quad K(f) = \mathcal{D}(W(X+f)) + \left[(\mathcal{J}f)\#f + (\mathcal{J}f)\#\mathcal{J}^* \left(\frac{\mathcal{J}f}{1+\mathcal{J}f} \right) - \mathcal{J}^* \left(\frac{(\mathcal{J}f)^2}{1+\mathcal{J}f} \right) \right].$$

Proof. Using the formula $\frac{1}{1+x} = 1 - \frac{x}{1+x}$ and the fact that $\mathcal{J}^*(1 \otimes 1) = X$, we see that (3.5) is equivalent to

$$\mathcal{J}^* \left(\frac{\mathcal{J}f}{1+\mathcal{J}f} \right) + f + (\mathcal{D}W)(X+f) = 0.$$

We will now apply $(1 + \mathcal{J}f)\#$ to both sides of the equation (this map is, by assumption, invertible, so the resulting equation is equivalent to the one in the previous line):

$$\begin{aligned} \mathcal{J}^* \left(\frac{\mathcal{J}f}{1+\mathcal{J}f} \right) + f + (\mathcal{D}W)(X+f) + (\mathcal{J}f)\#\mathcal{J}^* \left(\frac{\mathcal{J}f}{1+\mathcal{J}f} \right) \\ + (\mathcal{J}f)\#f + (\mathcal{J}f)\#(\mathcal{D}W)(X+f) = 0. \end{aligned}$$

Using that $\frac{x}{1+x} = x - \frac{x^2}{1+x}$, we get:

$$\begin{aligned} \mathcal{J}^*(\mathcal{J}f) - \mathcal{J}^* \left(\frac{(\mathcal{J}f)^2}{1+\mathcal{J}f} \right) + f + (\mathcal{D}W)(X+f) \\ + (\mathcal{J}f)\#\mathcal{J}^* \left(\frac{\mathcal{J}f}{1+\mathcal{J}f} \right) + (\mathcal{J}f)\#f + (\mathcal{J}f)\#(\mathcal{D}W)(X+f) = 0. \end{aligned}$$

Thus we have:

$$\begin{aligned} K(f) = (\mathcal{D}W)(X+f) + (\mathcal{J}f)\#(\mathcal{D}W)(X+f) \\ + \left[(\mathcal{J}f)\#f + (\mathcal{J}f)\#\mathcal{J}^* \left(\frac{\mathcal{J}f}{1+\mathcal{J}f} \right) - \mathcal{J}^* \left(\frac{(\mathcal{J}f)^2}{1+\mathcal{J}f} \right) \right]. \end{aligned}$$

We now note that, because of cyclic symmetry and the fact that $f = \mathcal{D}g$ with $g = g^*$, we may use (3.4) to deduce:

$$\mathcal{D}(W(X+f)) = (\mathcal{D}W)(X+f) + (\mathcal{J}f)\#(\mathcal{D}W)(X+f)$$

Hence (3.5) is equivalent to

$$K(f) = \mathcal{D}(W(X+f)) + \left[(\mathcal{J}f)\#f + (\mathcal{J}f)\#\mathcal{J}^* \left(\frac{\mathcal{J}f}{1+\mathcal{J}f} \right) - \mathcal{J}^* \left(\frac{(\mathcal{J}f)^2}{1+\mathcal{J}f} \right) \right],$$

as claimed. \square

3.4. **Some identities involving \mathcal{J} and \mathcal{D} .** We start with the following identity.

Lemma 3.4. *Put $M = W^*(\mathcal{A}, \tau)$. Assume that $1 \otimes 1$ belongs to the domain of ∂_j^* for all j and that τ is a faithful trace, i.e., $\tau(P^*P) = 0$ iff $P = 0$. Let $g \in \mathcal{A}^{(A)}$ and let $f = \mathcal{D}g$. Then for any $m \geq -1$ we have:*

$$(3.7) \quad \frac{1}{m+2} \mathcal{D} \left[(1 \otimes \tau + \tau \otimes 1) \left\{ \sum_{i=1}^n [(\mathcal{J}f)^{m+2}]_{ii} \right\} \right] \\ = -\mathcal{J}^* ((\mathcal{J}f)^{m+2}) + \mathcal{J}f \# \mathcal{J}^* ((\mathcal{J}f)^{m+1}) .$$

Proof. Let us take $f = \mathcal{D}g$. We shall prove that (3.7) is true by showing that it holds weakly. We therefore let P be a test function in $(\mathcal{A}^{(A)})^n$. Then, with the notations of the proof of Lemma 3.1, see (3.2), and with $\#$ the multiplication in $(\mathcal{A}^{(A)})^n$, we have

$$\begin{aligned} \Lambda &:= \tau \left([-\mathcal{J}^* ((\mathcal{J}f)^{m+2}) + \mathcal{J}f \# \mathcal{J}^* ((\mathcal{J}f)^{m+1})] \# P \right) \\ &= -\tau \left(\mathcal{J}^* ((\mathcal{J}f)^{m+2}) \# P \right) + \tau \left(\sum_{ij} (\mathcal{J}f)_{ij} \# (\mathcal{J}^* ((\mathcal{J}f)^{m+1}))_j \# P_i \right) \\ &= -\tau \left(\mathcal{J}^* ((\mathcal{J}f)^{m+2}) \# P \right) + \tau \left(\sum_{ij} (\mathcal{J}^* ((\mathcal{J}f)^{m+1}))_j \# \sigma(\mathcal{J}f)_{ij} \# P_i \right) \end{aligned}$$

where we just used that τ is tracial. Since $f = \mathcal{D}g$, it follows that $\sigma(\mathcal{J}f)_{ij} = (\mathcal{J}f)_{ji}$ by (3.2). Therefore we deduce that

$$\begin{aligned} \Lambda &= -\tau \left(\mathcal{J}^* ((\mathcal{J}f)^{m+2}) \# P \right) + \tau \left(\sum_{ij} (\mathcal{J}^* ((\mathcal{J}f)^{m+1}))_j \# (\mathcal{J}f)_{ji} \# P_i \right) \\ &= -\sum_{ij} \tau \otimes \tau \left((\mathcal{J}f)_{ij}^{m+2} \# \sigma[\mathcal{J}P]_{ij} \right) + \sum_{ij} \tau \otimes \tau \left((\mathcal{J}f)_{ij}^{m+1} \# \sigma[\mathcal{J}((\mathcal{J}f)\#P)]_{ij} \right) \end{aligned}$$

where in the last equality we used the definition of \mathcal{J}^* :

$$\tau(\mathcal{J}^*(U)\#W) = \sum_{ij} \tau \otimes \tau \left(U_{ij} \# \sigma[\mathcal{J}W]_{ij} \right).$$

Let us consider the term $\sigma[\mathcal{J}((\mathcal{J}f)\#P)]$. Using the Leibniz rule, we obtain

$$\begin{aligned} [\mathcal{J}((\mathcal{J}f)\#P)]_{ij} &= \partial_j \sum_k (\partial_k f_i) \# P_k \\ &= \sum_k (\partial_k f_i) \# \partial_j P_k + \sum_k ((\partial_j \otimes 1)(\partial_k f_i)) \# P_k + \sum_k ((1 \otimes \partial_j)(\partial_k f_i)) \# P_k \end{aligned}$$

where we use the notation $(a \otimes b \otimes c) \#_1 \xi = a \xi b \otimes c$, $(a \otimes b \otimes c) \#_2 \xi = a \otimes b \xi c$.

Thus

$$\begin{aligned}
\sigma[\mathcal{J}((\mathcal{J}f)\#P)]_{ij} &= \sum_k \sigma((\partial_k f_i)\# \partial_j P_k) \\
&\quad + \sum_k \sigma[((\partial_j \otimes 1)(\partial_k f_i))\#_2 P_k] + \sum_k \sigma[((1 \otimes \partial_j)(\partial_k f_i))\#_1 P_k] \\
&= \sum_k \sigma(\mathcal{J}P)_{kj} \# \sigma(\mathcal{J}f)_{ik} \\
&\quad + \sum_k \eta[((\partial_j \otimes 1)(\partial_k f_i))\#_1 P_k] + \sum_k \eta[((1 \otimes \partial_j)(\partial_k f_i))\#_2 P_k]
\end{aligned}$$

where $\eta[a \otimes b \otimes c] = b \otimes c \otimes a$. Since $f = \mathcal{D}g$, we know that $\sigma(\mathcal{J}f)_{ik} = (\mathcal{J}f)_{ki}$, so that

$$\begin{aligned}
\sigma[\mathcal{J}((\mathcal{J}f)\#P)]_{ij} &= \sum_k \sigma(\mathcal{J}P)_{kj} \# (\mathcal{J}f)_{ki} \\
&\quad + \sum_k \eta[((\partial_j \otimes 1)(\partial_k f_i))\#_1 P_k] + \sum_k \eta[((1 \otimes \partial_j)(\partial_k f_i))\#_2 P_k]
\end{aligned}$$

We now apply the identity

$$\eta[(\partial_j \otimes 1)\partial_k \mathcal{D}_i g] = (\partial_k \otimes 1)\partial_i \mathcal{D}_j g = (1 \otimes \partial_i)\partial_k \mathcal{D}_j g$$

(and a similar one involving $\eta[((1 \otimes \partial_j)(\partial_k f_i))]$) to deduce that

$$\begin{aligned}
\sigma[\mathcal{J}((\mathcal{J}f)\#P)]_{ij} &= \sum_k \sigma(\mathcal{J}P)_{kj} \# \sigma(\mathcal{J}f)_{ik} \\
&\quad + \sum_k ((1 \otimes \partial_i)(\partial_k f_j))\#_1 P_k + \sum_k ((\partial_i \otimes 1)(\partial_k f_j))\#_2 P_k \\
&= \sigma(\mathcal{J}P)_{kj} \# \sigma(\mathcal{J}f)_{ik} + Q_{ji}^P
\end{aligned}$$

where

$$Q_{ji}^P = \sum_k (1 \otimes \partial_i)\partial_k f_j \#_1 P_k + \sum_k (\partial_i \otimes 1)\partial_k f_j \#_2 P_k.$$

Thus if we set $Q^P = ((Q^P)_{ij})_{ij}$ and $R = (R_{ij})$ with $R_{ij} = \sigma(\mathcal{J}P)_{ji}$, then

$$\begin{aligned}
\Lambda &= -\tau \otimes \tau \otimes Tr((\mathcal{J}f)^{m+2}\#R) + \tau \otimes \tau \otimes Tr((\mathcal{J}f)^{m+1}\#R\#(\mathcal{J}f)) \\
&\quad + \tau \otimes \tau \otimes Tr((\mathcal{J}f)^{m+1}\#Q^P) \\
&= \tau \otimes \tau \otimes Tr((\mathcal{J}f)^{m+1}\#Q^P) \\
(3.8) \quad &= \frac{1}{m+2} \sum_{h=0}^{m+1} \tau \otimes \tau \otimes Tr((\mathcal{J}f)^h\#Q^P\#(\mathcal{J}f)^{m+1-h}).
\end{aligned}$$

We therefore need to prove that

$$\begin{aligned}
(3.9) \quad &\sum_{h=0}^{m+1} \tau \otimes \tau \otimes Tr((\mathcal{J}f)^h\#Q^P\#(\mathcal{J}f)^{m+1-h}) \\
&= \tau(P\#\mathcal{D}[(1 \otimes \tau + \tau \otimes 1)\{Tr(\mathcal{J}f)^{m+2}\}])
\end{aligned}$$

holds for all test function P . We prove this equality by showing that both sides are equal to the derivative of the same quantity. Let $X_j^t = X_j + tP_j$, and consider

$$\begin{aligned} \frac{d}{dt} \Big|_{t=0} (\mathcal{J}f(X^t)_{ij}) &= \frac{d}{dt} \Big|_{t=0} \partial_j f_i(X^t) = \sum_k (1 \otimes \partial_k) \partial_j f_i \#_2 P_k + \sum_k (\partial_k \otimes 1) \partial_j f_i \#_1 P_k \\ &= \sum_k (1 \otimes \partial_j) \partial_k f_i \#_1 P_k + \sum_k (\partial_j \otimes 1) \partial_k f_i \#_2 P_k \end{aligned}$$

where we have used the identity $(\partial_j \otimes 1) \partial_k = (1 \otimes \partial_k) \partial_j$. It follows that

$$Q_{ij}^P = \frac{d}{dt} \Big|_{t=0} \left([\mathcal{J}f(X^t)]_{ij} \right).$$

Thus, we deduce

$$\sum_{t=0}^{m+1} \tau \otimes \tau \otimes Tr \left((\mathcal{J}f)^h \# Q^P \# (\mathcal{J}f)^{m+1-h} \right) = \frac{d}{dt} \Big|_{t=0} \tau \otimes \tau \otimes Tr \left((\mathcal{J}f(X^t))^{m+2} \right).$$

On the other hand, if $R \in M_{n \times n}(\mathcal{A} \otimes \mathcal{A}^{op})$ is arbitrary, then we claim that

$$(3.10) \quad \frac{d}{dt} \Big|_{t=0} \tau \otimes \tau \otimes Tr (R(X^t)) = \sum_i \tau(\mathcal{D}(1 \otimes \tau + \tau \otimes 1)(R_{ii}) \# P).$$

Indeed, this is sufficient to check (3.10) for R a matrix with a single nonzero component $A \otimes B$ in the i, j -th position. For this choice of R , we have

$$\begin{aligned} \frac{d}{dt} \Big|_{t=0} \tau \otimes \tau \otimes Tr(R(X^t)) &= \frac{d}{dt} \Big|_{t=0} \delta_{i=j} \tau(A(X^t)) \tau(B(X^t)) \\ &= \delta_{i=j} \left(\frac{d}{dt} \Big|_{t=0} \tau(A(X^t)) \tau(B) \right) + \left(\frac{d}{dt} \Big|_{t=0} \tau(A) \tau(B(X^t)) \right) \\ &= \delta_{i=j} (\tau(\mathcal{D}A \# P) \tau(B) + \tau(A) \tau(\mathcal{D}B \# P)) \\ &= \sum_i \tau(\mathcal{D}((1 \otimes \tau + \tau \otimes 1)(R_{ii}) \# P)), \end{aligned}$$

as claimed. We now combine (3.8) and (3.10) (with $R = (\mathcal{J}f)^{m+2}$) to conclude that (3.9) holds. Since this identity holds for any P , the Lemma follows. \square

Lemma 3.5. *Assume that $f = \mathcal{D}g$, $g = g^* \in \mathcal{A}^{(A)}$, assume that $1 \otimes 1 \in \text{dom } \partial_j^*$ for all j and that τ is a faithful trace, i.e., $\tau(P^*P) = 0$ iff $P = 0$. Assume that $\|\mathcal{J}f\|_{A \otimes \tau A} < 1$ and let*

$$Q(g) = \left[(1 \otimes \tau + \tau \otimes 1) \left\{ \sum_{i=1}^n (\mathcal{J}f)_{ii} - (\log(1 + \mathcal{J}f))_{ii} \right\} \right].$$

Then

$$\mathcal{D}Q = \mathcal{J}f \# \mathcal{J}^* \left(\frac{\mathcal{J}f}{1 + \mathcal{J}f} \right) - \mathcal{J}^* \left(\frac{(\mathcal{J}f)^2}{1 + \mathcal{J}f} \right).$$

Proof. We use Lemma 3.4 and compare the (converging) Taylor series expansions of both sides term by term. \square

Lemma 3.6. *Assume that τ is the semicircle law. Let*

$$K(f) = -\mathcal{J}^*(\mathcal{J}f) - f.$$

Assume that $f = \mathcal{D}g$, and that $g \in \mathcal{A}$ is cyclically symmetric. Then

$$\begin{aligned} K(f) &= \mathcal{D} \left\{ (1 \otimes \tau + \tau \otimes 1) \left(\sum_i [\mathcal{J}f]_{ii} \right) - f \# X \right\} \\ &= \mathcal{D} \{ (1 \otimes \tau + \tau \otimes 1) (\text{Tr} [\mathcal{J} \mathcal{D}g]) - \mathcal{N}g \}. \end{aligned}$$

Proof. When $m = -1$, the equality in Lemma 3.4 becomes

$$\mathcal{D} \{ (1 \otimes \tau + \tau \otimes 1) (\text{Tr} [\mathcal{J}f]) \} = -\mathcal{J}^*(\mathcal{J}f) + \mathcal{J}f \# \mathcal{J}^*(I),$$

where $I \in M_{n \times n}(\mathcal{A} \otimes \mathcal{A}^{op})$ denotes the identity matrix. Since τ is the semicircle law, $\mathcal{J}^*(I) = X$, and $\mathcal{J}f \# X = \mathcal{N}f$ (here \mathcal{N} is applied entrywise to the vector f).

On the other hand, since \mathcal{D} reduces the degree of polynomials by 1, $\mathcal{N} \mathcal{D}g = \mathcal{D} \mathcal{N}g - \mathcal{D}g = \mathcal{D} \mathcal{N}g - f$. Thus we deduce

$$\begin{aligned} \mathcal{D} \{ (1 \otimes \tau + \tau \otimes 1) (\text{Tr} [\mathcal{J} \mathcal{D}g]) - \mathcal{N}g \} &= -\mathcal{J}^*(\mathcal{J}f) + \mathcal{N} \mathcal{D}g - \mathcal{D} \mathcal{N}g \\ &= K(f). \end{aligned}$$

Using the equality $f \# X = \mathcal{D}g \# X = \mathcal{N}g$ due to the fact that g is cyclically symmetric, we get the statement of the Lemma. \square

Lemma 3.7. *Let τ be the semicircle law. Assume that $f = \mathcal{D}g$ for some $g = g^*$ cyclically symmetric and that $\|\mathcal{J}f\|_{A \otimes \pi A} < 1$. Let*

$$Q(g) = [(1 \otimes \tau + \tau \otimes 1) \text{Tr} \{ (\mathcal{J} \mathcal{D}g) - (\log(1 + \mathcal{J} \mathcal{D}g)) \}].$$

Then Equation (3.5) is equivalent to the equation

$$\mathcal{D} \{ (1 \otimes \tau + \tau \otimes 1) (\text{Tr} [\mathcal{J} \mathcal{D}g]) - \mathcal{N}g \} = \mathcal{D}(W(X + \mathcal{D}g)) + \mathcal{D}Q(g) + \mathcal{J} \mathcal{D}g \# \mathcal{D}g.$$

Proof. By Lemma 3.6, the left-hand side is precisely $K(f)$. By Lemma 3.3, we obtain

$$K(f) = \mathcal{D}(W(X + f)) + \left(\mathcal{J}f \# \mathcal{J}^* \left(\frac{\mathcal{J}f}{1 + \mathcal{J}f} \right) - \mathcal{J}^* \left(\frac{(\mathcal{J}f)^2}{1 + \mathcal{J}f} \right) + \mathcal{J}f \# f \right)$$

which, according to Lemma 3.3 is equivalent to (3.5). The hypothesis of Lemma 3.3 is satisfied since the map $T : \xi \mapsto \mathcal{J}f \# \xi$ is strictly contractive on $(\mathcal{A}^{(A')})^n$ because of the bound on $\mathcal{J}f$ (and so the map $\text{id} + T$ has no kernel). \square

We now turn to the proof of existence of a g that satisfies the equation above. We will use a fixed-point argument; thus we first give some estimates on differential operators that will be involved.

3.5. Technical estimates on certain differential operators. For the remainder of the section, we will assume that $\tau : \mathcal{A} \rightarrow \mathbb{C}$ satisfies

$$(3.11) \quad |\tau(q)| \leq C_0^{\deg q}$$

for any monomial $q \in \mathcal{A}$. We begin with a few technical estimates on certain differential operators.

Lemma 3.8. *Let $g_1, \dots, g_m \in \mathcal{A}_0$. Set*

$$Q_m(g_1, \dots, g_m) = (1 \otimes \tau + \tau \otimes 1) \left\{ \sum_{i=1}^n [(\mathcal{J} \mathcal{D}g_1) \cdots (\mathcal{J} \mathcal{D}g_m)]_{ii} \right\}.$$

Assume that (3.11) holds and moreover that $C_0/A < 1/2$. Then

$$\|Q_m(\Sigma g_1, \dots, \Sigma g_m)\|_A \leq 2(2A^{-2})^m \prod_{k=1}^m \|g_k\|_A.$$

In particular, Q_m extends to a bounded multilinear operator on $\mathcal{A}^{(A)}$ with values in $\mathcal{A}^{(A)}$.

Proof. Note that by definition, assuming that g is a monomial, we have

$$\begin{aligned} [\mathcal{J} \mathcal{D}\Sigma g]_{ji} &= \frac{1}{\deg g} \partial_i \mathcal{D}_j g \\ &= \frac{1}{\deg g} \sum_{g=AX_j B} \sum_{BA=RX_i Q} R \otimes Q. \end{aligned}$$

Thus if g_1, \dots, g_m are monomials, then

$$[(\mathcal{J} \mathcal{D}\Sigma g_1) \cdots (\mathcal{J} \mathcal{D}\Sigma g_m)]_{j_0 j_m} = \frac{1}{\prod_t \deg g_t} \sum_{\substack{j_1, \dots, j_{m-1} \\ g_k = A_k X_{j_k} B_k}} \sum_{B_k A_k = R_k X_{j_{k+1}} Q_k} R_1 \cdots R_m \otimes Q_1 \cdots Q_m$$

and so, using the fact that for a given k there can be at most $\deg g_k$ decompositions of g_k as $A_k X_{j_{k+1}} B_k$, and the degree of A_k determines such decompositions, we conclude that each such sum has at most $\prod_t \deg g_t$ nonzero terms, so that

$$\begin{aligned} &\|(1 \otimes \tau)[(\mathcal{J} \mathcal{D}\Sigma g_1) \cdots (\mathcal{J} \mathcal{D}\Sigma g_m)]_{j_0 j_m}\|_A \\ &\leq \frac{1}{\prod_t \deg g_t} \sum_{\substack{j_1, \dots, j_{m-1} \\ g_k = A_k X_{j_k} B_k}} \sum_{B_k A_k = R_k X_{j_{k+1}} Q_k} A^{\deg(R_1 \cdots R_m)} \otimes |\tau(Q_1 \cdots Q_m)| \\ &\leq \sum_{l_1, \dots, l_m} A^{l_1 + \cdots + l_m} \cdot C_0^{\deg g_1 - l_1 - 2 + \cdots + \deg g_m - l_m - 2} \\ &= A^{\deg g_1 + \cdots + \deg g_m} A^{-2m} \sum_{l_1 \cdots l_{m-1}} \left(\frac{C_0}{A}\right)^{\deg g_1 - l_1 - 2 + \cdots + \deg g_m - l_m - 2} \\ &\leq \prod_{k=1}^m [A^{-2} \|g_k\|_A \frac{1}{1 - C_0/A}] \leq \prod_{k=1}^m [2A^{-2} \|g_k\|_A]. \end{aligned}$$

A similar estimate holds for $(\tau \otimes 1)[(\mathcal{J} \mathcal{D}\Sigma g_1) \cdots (\mathcal{J} \mathcal{D}\Sigma g_m)]_{j_0 j_m}$.

Since $Q_m(\Sigma g_1, \dots, \Sigma g_m)$ is multi-linear in g_1, \dots, g_m , we find that if $g_j = \sum \lambda(j, q)q$ is the decomposition of g_j in terms of monomials, then we get that

$$Q_m(\Sigma g_1, \dots, \Sigma g_m) = \sum_{q_1, \dots, q_m} \lambda(1, q_1) \cdots \lambda(m, q_m) Q_m(\Sigma q_1, \dots, \Sigma q_m)$$

and so

$$\begin{aligned}
\|Q_m(\Sigma g_1, \dots, \Sigma g_m)\|_A &\leq 2(2A^{-2})^m \sum_{q_1, \dots, q_m} \prod_{j=1}^m |\lambda(j, q_j)| \|q_j\|_A \\
&= 2(2A^{-2})^m \prod_{j=1}^m \sum_{q_j} |\lambda(j, q_j)| \|q_j\|_A \\
&= 2(2A^{-2})^m \prod_{j=1}^m \|g_j\|_A.
\end{aligned}$$

This concludes the proof for arbitrary g_1, \dots, g_m . \square

Lemma 3.9. For $g, f \in \mathcal{A}_0$, set $Q_m(\Sigma g) = Q_m(\Sigma g, \dots, \Sigma g)$. Assume that (3.11) holds and moreover that $C_0/A < 1/2$. Then

$$\|Q_m(\Sigma g) - Q_m(\Sigma f)\|_A \leq 2(2A^{-2})^m \sum_{k=0}^{m-1} \|g\|_A^k \|f\|_A^{m-k-1} \|f - g\|_A$$

In particular, $\|Q_m(\Sigma g)\|_A \leq 2(2A^{-2})^m \|g\|_A^m$.

Proof. We perform a telescopic expansion

$$\begin{aligned}
\|Q_m(\Sigma f) - Q_m(\Sigma g)\|_A &= \left\| \sum_{k=0}^{m-1} Q_m(\underbrace{\Sigma g, \dots, \Sigma g}_k, \underbrace{\Sigma f, \dots, \Sigma f}_{m-k}) - Q_m(\underbrace{\Sigma g, \dots, \Sigma g}_{k+1}, \underbrace{\Sigma f, \dots, \Sigma f}_{m-k-1}) \right\|_A \\
&\leq \sum_{k=0}^{m-1} \|Q_m(\underbrace{\Sigma g, \dots, \Sigma g}_k, \Sigma f - \Sigma g, \underbrace{\Sigma f, \dots, \Sigma f}_{m-k-1})\|_A \\
&\leq 2(2A^{-2})^m \sum_{k=0}^{m-1} \|g\|_A^k \|f\|_A^{m-k-1} \|f - g\|_A.
\end{aligned}$$

\square

Corollary 3.10. Assume that (3.11) holds and moreover that $C_0/A < 1/2$. Then the maps $g \mapsto Q_m(\Sigma g)$ extend by continuity to the completion of \mathcal{A}_0 with respect to $\|\cdot\|_A$.

Lemma 3.11. Assume that (3.11) holds and moreover that $C_0/A < 1/2$. Let $g \in \mathcal{A}_0^{(A)}$ be such that $\|g\|_A < \frac{A^2}{2}$, and set

$$Q(\Sigma g) = \sum_{m \geq 0} \frac{(-1)^m}{m+2} Q_{m+2}(\Sigma g).$$

Then this series converges in $\|\cdot\|_A$. Moreover, in the sense of analytic functional calculus on $M_{n \times n}(W^*(\mathcal{A} \otimes \mathcal{A}^{op}, \tau \otimes \tau^{op}))$, we have equality

$$Q(\Sigma g) = \left[(1 \otimes \tau + \tau \otimes 1) \left\{ \sum_{i=1}^n (\mathcal{J} \mathcal{D} \Sigma g)_{ii} - (\log(1 + \mathcal{J} \mathcal{D} \Sigma g))_{ii} \right\} \right].$$

Furthermore, the function Q satisfies the (local) Lipschitz condition on $\{g : \|g\|_A < A^2/2\}$

$$\|Q(\Sigma g) - Q(\Sigma f)\|_A \leq \|f - g\|_A \frac{2}{A^2} \left(\frac{1}{\left(1 - \frac{2\|f\|_A}{A^2}\right) \left(1 - \frac{2\|g\|_A}{A^2}\right)} - 1 \right),$$

and the bound

$$\|Q(\Sigma g)\|_A \leq \frac{\left(\frac{2\|g\|_A}{A^2}\right)^2}{1 - \frac{2\|g\|_A}{A^2}}.$$

Proof. Let $\kappa = \frac{A^2}{2}$ and $\lambda = \|g\|_A$. By Lemma 3.8, we have that $\|Q_{m+2}(g)\|_A \leq 2(\lambda/\kappa)^{(m+2)}$. Thus, as we assumed $\lambda < \kappa$, the series defining $Q(\Sigma g)$ converges. To see the claimed equality with the expression given by functional calculus, we only need to note that $\log(1+x) = -\sum_{m \geq 1} \frac{(-x)^m}{m}$. Finally, since $m+2 \geq 2$ in our series, we obtain that

$$\begin{aligned} \|Q(\Sigma g) - Q(\Sigma f)\|_A &\leq \sum_{m \geq 0} \frac{1}{m+2} \|Q_{m+2}(\Sigma g) - Q_{m+2}(\Sigma f)\|_A \\ &\leq \|f - g\|_A \sum_{m \geq 0} \sum_{k=0}^{m+1} (2A^{-2})^{m+2} \|f\|_A^{m-k+1} \|g\|_A^k \\ &\leq \|f - g\|_A (2A^{-2}) \left(\sum_{l \geq 0} \sum_{k \geq 0} (2A^{-2})^l \|f\|_A^l (2A^{-2})^k \|g\|_A^k - 1 \right), \end{aligned}$$

where we have written $m = l + k - 1$ which is non-negative precisely when l and k are not both zero. Thus, if $\|f\|_A \leq A^2/2$ and $\|g\|_A \leq A^2/2$, we deduce that

$$\|Q(\Sigma g) - Q(\Sigma f)\|_A \leq \|f - g\|_A \frac{2}{A^2} \left(\frac{1}{\left(1 - \frac{2\|f\|_A}{A^2}\right) \left(1 - \frac{2\|g\|_A}{A^2}\right)} - 1 \right).$$

Setting $f = 0$ gives us the estimate

$$\|Q(\Sigma g)\|_A \leq \frac{2\|g\|_A}{A^2} \left(\frac{1}{1 - \frac{2\|g\|_A}{A^2}} - 1 \right),$$

as claimed. \square

Corollary 3.12. *Let $g \in \mathcal{A}_0^{(A)}$, assume that $\|g\|_A < A^2/2$ and let $B \geq A + \|g\|_A$. Let $W \in \mathcal{A}^{(B)}$. Assume that (3.11) holds and moreover that $C_0/A < 1/2$. Let*

$$\begin{aligned} F(g) &= -W(X) + \left((1 \otimes \tau + \tau \otimes 1) \text{Tr}(\mathcal{J} \mathcal{D}\Sigma g) \right. \\ &\quad \left. - \left\{ W(X + \mathcal{D}\Sigma g) - W(X) + Q(\Sigma g) + \frac{1}{2} \mathcal{D}\Sigma g \# \mathcal{D}\Sigma g \right\} \right) \\ &= -W(X + \mathcal{D}\Sigma g) - \frac{1}{2} \mathcal{D}\Sigma g \# \mathcal{D}\Sigma g + (1 \otimes \tau + \tau \otimes 1) \text{Tr} \log(1 + \mathcal{J} \mathcal{D}\Sigma g). \end{aligned}$$

(here we abbreviate $W(X_1, \dots, X_n)$ by $W(X)$, etc.)

Then $F(g)$ is a well-defined function from $\{g \in \mathcal{A}_0^{(A)} : \|g\|_A < A^2/2\}$ to $\mathcal{A}^{(A)}$. Moreover $g \mapsto F(g)$ is locally Lipschitz on $\{g : \|g\|_A < A^2/2\}$:

$$\|F(g) - F(f)\|_A \leq \|f - g\|_A \left\{ \frac{2}{A^2} \left(\frac{1}{\left(1 - \frac{2\|f\|_A}{A^2}\right) \left(1 - \frac{2\|g\|_A}{A^2}\right)} + 1 \right) + \sum_j \|\partial_j W\|_{B \otimes_\pi B} + \frac{1}{2} (\|f\|_A + \|g\|_A) \right\}$$

and bounded:

$$\|F(g)\|_A \leq \|g\|_A \left\{ \frac{2}{A^2} \left(\frac{1}{1 - \frac{2\|g\|_A}{A^2}} + 1 \right) + \sum_j \|\partial_j W\|_{B \otimes_\pi B} + \frac{1}{2} \|g\|_A \right\} + \|W\|_A$$

In particular, if A , ρ and W are such that

$$(3.12) \quad \begin{cases} A > 4, & 0 < \rho \leq 1 \\ \|W\|_A < \frac{\rho}{12} \\ \sum_j \|\partial_j W\|_{(A+\rho) \otimes_\pi (A+\rho)} < \frac{1}{8} \end{cases}$$

then F takes the ball $\{g : \|g\|_A < \frac{\rho}{4}\}$ to itself and is uniformly contractive with constant $\lambda \leq \frac{7}{8}$ on that ball.

Proof. Note that $\|\mathcal{D}\Sigma g\|_A \leq \|g\|_A$ and so $\|X + \mathcal{D}\Sigma g\| \leq B$. Furthermore, for any h ,

$$\sum_i \|\mathcal{D}_i \Sigma h\|_A \leq \|h\|_A.$$

Indeed, if $h = \sum_r \sum_{i_1, \dots, i_r} \alpha_{i_1, \dots, i_r} X_{i_1} \cdots X_{i_r}$, then since $A > 1$

$$\sum_i \|\mathcal{D}_i \Sigma h\|_A = \sum_r \sum_{i_1, \dots, i_r} \sum_{q=1}^r |\alpha_{i_1, \dots, i_r}| r^{-1} \|X_{i_{q+1}} \cdots X_{i_r} X_1 \cdots X_{i_{q-1}}\|_A \leq \frac{1}{A} \|h\|_A \leq \|h\|_A.$$

The Lipschitz property then follows from Lemmas 3.11 and 3.9, as well as the estimates:

$$\begin{aligned} \|W(X + \mathcal{D}\Sigma g) - W(X + \mathcal{D}\Sigma f)\|_A &\leq \sum_j \|\partial_j W\|_{B \otimes_\pi B} \|\mathcal{D}_j \Sigma g - \mathcal{D}_j \Sigma f\|_A \\ &\leq \sum_j \|\partial_j W\|_{B \otimes_\pi B} \|g - f\|_A \end{aligned}$$

and

$$\begin{aligned}
\left\| \frac{1}{2} \mathcal{D}\Sigma g \# \mathcal{D}\Sigma g - \frac{1}{2} \mathcal{D}\Sigma f \# \mathcal{D}\Sigma f \right\|_A &\leq \frac{1}{2} \|\mathcal{D}\Sigma g \# (\mathcal{D}\Sigma g - \mathcal{D}\Sigma f)\|_A \\
&\quad + \frac{1}{2} \|\mathcal{D}\Sigma f \# (\mathcal{D}\Sigma g - \mathcal{D}\Sigma f)\|_A \\
&\leq \frac{1}{2} \left(\sum_i \|\mathcal{D}_i \Sigma g\|_A \max_i \|(\mathcal{D}_i \Sigma g - \mathcal{D}_i \Sigma f)\|_A \right. \\
&\quad \left. + \sum_i \|\mathcal{D}_i \Sigma f\|_A \max_i \|(\mathcal{D}_i \Sigma g - \mathcal{D}_i \Sigma f)\|_A \right) \\
&\leq \frac{1}{2} \|\mathcal{D}\Sigma g - \mathcal{D}\Sigma f\|_A (\|\mathcal{D}\Sigma g\|_A + \|\mathcal{D}\Sigma f\|_A) \\
&\leq \frac{1}{2} \|g - f\|_A (\|g\|_A + \|f\|_A).
\end{aligned}$$

The estimate on $\|F(g)\|_A$ follows from the identity $F(0) = -W$, Lemmas 3.8 and 3.11 and the Lipschitz property.

Assuming that (3.12) holds, so that $A > 4$ and $\|f\|_A, \|g\|_A < \frac{\rho}{4} < \rho \leq 1$, we have that

$$\frac{2}{A^2} \left(\frac{1}{\left(1 - \frac{2\|f\|_A}{A^2}\right) \left(1 - \frac{2\|g\|_A}{A^2}\right)} + 1 \right) < \frac{1}{8} \left(\frac{64}{49} + 1 \right) < \frac{1}{2}$$

and so the Lipschitz constant of F is bounded by $\frac{1}{2} + \frac{1}{4} + \sum_j \|\partial_j W\|_{(A+\rho) \otimes_{\pi} (A+\rho)} < \frac{1}{2} + \frac{1}{4} + \frac{1}{8} = \frac{7}{8}$. Also,

$$\|F(g)\| < \frac{\rho}{4} \left\{ \frac{1}{8} \left(\frac{8}{7} + 1 \right) + \frac{1}{8} + \frac{1}{8} \right\} + \|W\|_A < \left(\frac{2}{3} + \frac{1}{3} \right) \frac{\rho}{4} = \frac{\rho}{4}$$

so that the image of the ball $\{g : \|g\|_A < \frac{\rho}{4}\}$ is contained in that ball. \square

3.6. Existence of g . We will now consider the case that τ is the semicircle law. Then (3.11) holds with $C_0 = 2$. We remind the reader that Π, Σ , are defined in section 2.2 whereas \mathcal{S} is defined in Definition 2.2.

Proposition 3.13. *Let τ be the semicircle law. Assume that for some A and ρ , $W \in \mathcal{A}_0^{(A)}$ is cyclically symmetric and that conditions (3.12) are satisfied, i.e.,*

$$\begin{cases} A > 4, & 0 < \rho \leq 1 \\ \|W\|_A < \frac{\rho}{12} \\ \sum_j \|\partial_j W\|_{(A+\rho) \otimes_{\pi} (A+\rho)} < \frac{1}{8} \end{cases}$$

Then there exists \hat{g} and $g = \Sigma \hat{g}$ (so that $\hat{g} = \mathcal{N}g$) with the following properties:

(i) Both \hat{g} and g belong to the completion of \mathcal{A}_0 with respect to the norm $\|\cdot\|_A$ and $\|g\|_A \leq \rho/4$, $\|\hat{g}\|_A \leq 3\|W\|_A$;

(ii) \hat{g} satisfies the equation $\hat{g} = \mathcal{S}\Pi F(\hat{g})$

(iii) \hat{g} and g depend analytically on W , in the following sense: if the maps $\beta \mapsto W_{\beta}$ are analytic, then also the maps $\beta \mapsto \hat{g}(\beta)$ and $\beta \mapsto g(\beta)$ are analytic, and $g \rightarrow 0$ if $\|W\|_A \rightarrow 0$

(iv) g satisfies the equation

$$\begin{aligned}
 (3.13) \quad \mathcal{S}\Pi\mathcal{N}g &= -W(X) + \mathcal{S}\Pi \left((1 \otimes \tau + \tau \otimes 1)(\text{Tr}(\mathcal{J}\mathcal{D}g)) - \right. \\
 &\quad \left. \left\{ W(X + \mathcal{D}g) - W(X) + Q(g) + \frac{1}{2}\mathcal{D}g\#\mathcal{D}g \right\} \right) \\
 &= \mathcal{S}\Pi \left[-W(X + \mathcal{D}g) - \frac{1}{2}\mathcal{D}g\#\mathcal{D}g + (1 \otimes \tau + \tau \otimes 1)\text{Tr} \log(1 + \mathcal{J}\mathcal{D}g) \right].
 \end{aligned}$$

or, equivalently,

$$\begin{aligned}
 (3.14) \quad -\mathcal{S}\Pi\mathcal{N}g + \mathcal{S}\Pi(1 \otimes \tau + \tau \otimes 1)(\text{Tr}(\mathcal{J}\mathcal{D}g)) &= W(X) \\
 &\quad + \mathcal{S}\Pi \left\{ W(X + \mathcal{D}g) - W(X) + Q(g) + \frac{1}{2}\mathcal{D}g\#\mathcal{D}g \right\}.
 \end{aligned}$$

Proof. Observe that Equation (3.13) is equivalent to

$$\mathcal{S}\Pi\mathcal{N}g = \mathcal{S}\Pi F(\mathcal{N}g)$$

with F the function defined in Corollary 3.12, where we used that W is cyclically symmetric and in $\mathcal{A}_0^{(A)}$. Existence of g follows, essentially, from the Implicit function Theorem 2.3. We prefer to repeat its proof here for completeness. We set $\hat{g}_0 = W(X_1, \dots, X_n)$, and for each $k > 0$,

$$\hat{g}_k = \mathcal{S}\Pi F(\hat{g}_{k-1}),$$

Since $\mathcal{S}\Pi$ is a linear contraction, the last part of Corollary 3.12 implies that under our hypothesis, $\mathcal{S}\Pi \circ F$ is uniformly contractive with constant $7/8$ on the ball $B = \{\hat{g} : \|\hat{g}\|_A < \frac{\rho}{4}\}$, and takes this ball to itself. It follows that $\hat{g}_k \in B$ for all k , and moreover

$$\|\hat{g}_k - \hat{g}_{k-1}\|_A = \|\mathcal{S}\Pi F(\hat{g}_{k-1}) - \mathcal{S}\Pi F(\hat{g}_{k-2})\| < \frac{7}{8}\|\hat{g}_{k-1} - \hat{g}_{k-2}\|_A,$$

so that \hat{g}_k converges in $\|\cdot\|_A$ to a fixed point \hat{g} . Since $\hat{g} \in \mathcal{A}_0^{(A)}$, so does $g = \Sigma\hat{g}$.

Note that if conditions (3.12) are satisfied for some ρ , and $\|W\|_A < \rho'/12$ for some $\rho' < \rho$, then conditions (3.12) are again satisfied with ρ' in place of ρ (indeed, $\|\partial_j W\|_{(A+\rho') \otimes_\pi (A+\rho')} \leq \|\partial_j W\|_{(A+\rho) \otimes_\pi (A+\rho)}$). Since by construction $\|g\|_A \leq \|\hat{g}\|_A \leq \frac{\rho}{4}$, we conclude that in fact $\|g\|_A \leq \frac{\rho'}{4}$, for any $\rho' > 12\|W\|_A$. It follows that $\|g\|_A \leq 3\|W\|_A$.

This proves (i) and (ii).

Since (assuming that $\beta \mapsto W_\beta$ is analytic) each iterate \hat{g}_k is clearly analytic in β and the convergence of the Banach-space valued function $\hat{g}_k(\beta) \in \mathcal{A}_0^{(A)}$ is uniform on any compact disk inside $|\beta| < \beta_0$, it follows from the Cauchy integral formula that the limit is also an analytic function.

Part (iv) follows from the definition of F and Σ . \square

We leave the following to the reader (note that \hat{g} is clearly cyclically symmetric, and it is not hard to see that this implies that also g is cyclically symmetric):

Proposition 3.14. *Assume that $W = W^*$ and that W satisfies the hypothesis of Proposition 3.13. Let g be the solution to (3.14) constructed in Proposition 3.13. Then g belongs to the closure of cyclically symmetric polynomials and also satisfies $g = g^*$.*

3.7. The map $f = \mathcal{D}g$ satisfies Equation (3.5). From now on, we will assume that τ is the semicircle law.

Theorem 3.15. *Let $A > A' > 4$. Then there exists a constant $C(A, A') > 0$ depending only on A, A' so that whenever $W = W^* \in \mathcal{A}^{(A)}$ satisfies $\|W\|_{A+1} < C(A, A')$, there exists an $f \in (\mathcal{A}^{(A')})^n$ which satisfies (3.5). In addition, $f = \mathcal{D}g$ for some $g \in \mathcal{A}^{(A')}$. The solution $f = f_W$ satisfies $\|f_W\|_{A'} \rightarrow 0$ as $\|W\|_A \rightarrow 0$. Moreover, if W_β is a family which is analytic in β then also the solutions f_{W_β} are analytic in β .*

Proof. Let A_1 be so that $4 < A_1 < A$, and let

$$c(A_1, A) = \frac{1}{eA_1 \log(A/A_1)} = \sup_{\alpha \geq 1} \alpha A_1^{-1} (A/A_1)^{-\alpha}.$$

Then if $g = \sum \lambda(q)q$, we obtain the estimate

$$\begin{aligned} \|\mathcal{D}g\|_{A_1} &\leq \sum |\lambda(q)| \deg(q) A_1^{\deg(q)-1} \\ &\leq \sum |\lambda(q)| \deg(q) A^{\deg(q)} A_1^{-1} (A/A_1)^{-\deg(q)} \\ &\leq \sum |\lambda(q)| A^{\deg(q)} c(A, A_1) \leq c(A_1, A) \|g\|_A \end{aligned}$$

Similarly, if we set $c'(A_1, A) = \sup_{\alpha \geq 1} \alpha^2 A_1^{-2} (A/A_1)^{-\alpha} < \infty$, then

$$\begin{aligned} (3.15) \quad \|\mathcal{J}f\|_{A_1 \otimes_\pi A_1} &= \max_i \sum_j \|\partial_j \mathcal{D}_i g\|_{A_1 \otimes_\pi A_1} \\ &\leq \sum |\lambda(q)| (\deg(q))^2 A_1^{\deg(q)-2} \\ &\leq c'(A_1, A) \|g\|_A. \end{aligned}$$

Finally, by the same computation as in the proof of Lemma 3.12,

$$\sum_i \|\partial_i W\|_{(A_1+1) \otimes_\pi (A_1+1)} \leq c(A_1 + 1, A + 1) \|W\|_{A+1}.$$

Thus we can choose $C(A, A') > 0$ so that if $\|W\|_{A+1} < C(A, A')$, then the hypothesis of Proposition 3.13 is satisfied (with $\rho = 1$), and there exists some g which satisfies

$$\begin{aligned} (3.16) \quad &-\mathcal{S}\Pi\mathcal{N}g + \mathcal{S}\Pi(1 \otimes \tau + \tau \otimes 1)(\text{Tr}(\mathcal{J}\mathcal{D}g)) \\ &= \mathcal{S}\Pi W(X) + \mathcal{S}\Pi \left\{ W(X + \mathcal{D}g) - W(X) + Q(g) + \frac{1}{2} \mathcal{D}g \# \mathcal{D}g \right\}. \end{aligned}$$

Furthermore, we may assume (by choosing a perhaps smaller $C(A, A')$) that $f = \mathcal{D}g$ satisfies $\|\mathcal{J}f\|_{A' \otimes_\pi A'} < 1$. To deduce that g satisfies (3.5), we shall apply \mathcal{D} on both sides of the

above equality. To this end, note that if $p = (p_1, \dots, p_n)$ with p_i monomials, then

$$\begin{aligned}
(\mathcal{D}(p\#p))_i &= \mathcal{D}_i\left(\sum_j p_j p_j\right) \\
&= \sum_j \left[\sum_{p_j p_j = p_j A X_i B} B p_j A + \sum_{p_j p_j = A X_i B p_j} B p_j A \right] \\
&= 2 \sum_j \sum_{p_j = A X_i B} B p_j A \\
&= 2 \sum_j \sigma(\mathcal{J} p)_{ji} \# p_j \quad (\text{here } \sigma(a \otimes b) = b \otimes a)
\end{aligned}$$

Now if we assume that $p = \mathcal{D}w$ for some w , then by (3.2) in the proof of (iii) in Lemma 3.1, we find that

$$(\mathcal{D}(p\#p))_i = 2 \sum_j \sigma(\mathcal{J} p)_{ji} \# p_j = 2 \sum_j (\mathcal{J} p)_{ij} \# p_j = 2(\mathcal{J} p) \# p.$$

Using this identity, applying \mathcal{D} to both sides of (3.16) and noting that for any h , $\mathcal{D}\mathcal{S}\Pi h = \mathcal{D}h$, gives us

$$\mathcal{D}\{(1 \otimes \tau + \tau \otimes 1)(\text{Tr}[\mathcal{J}\mathcal{D}g]) - \mathcal{N}g\} = \mathcal{D}(W(X + \mathcal{D}g)) + \mathcal{D}Q(g) + \mathcal{J}\mathcal{D}g\#\mathcal{D}g,$$

which, according to Lemma 3.7 is equivalent to (3.5). \square

Since $Y_W = X + f_W \rightarrow X$ as $\|W\|_A \rightarrow 0$, it follows that for sufficiently small $\|W\|_A$, $\|Y\|$ is bounded by A and so the law of Y is the unique solution to the Schwinger-Dyson equation with self-adjoint potential $V = \frac{1}{2} \sum X_j^2 + W$, and thus the law of Y is exactly τ_V (by [GMS06, Theorem 2.1] or alternatively since, for $\|W\|_A$ sufficiently small, V is then (c, A) convex and the Schwinger-Dyson equation has a unique solution [GS09]). In particular, $W^*(Y_1, \dots, Y_n) \cong W^*(\tau_V)$. We thus obtain a free analog of Brenier's theorem:

Theorem 3.16. *Let $A > A' > 4$, and let $X_1, \dots, X_n \in (M, \tau)$ be semicircular variables. Then there exists a universal constant $C = C(A, A') > 0$ so that whenever $W \in \mathcal{A}^{(A+1)}$ satisfies $\|W\|_{A+1} < C$, there is an element $G \in \mathcal{A}^{(A')}$ so that*

$$(Y_1, \dots, Y_n) = (\mathcal{D}_1 G, \dots, \mathcal{D}_n G) \in \mathcal{A}^{(A')}$$

has law τ_V , $V = \frac{1}{2} \sum X_j^2 + W$.

Moreover, the Hessian $\mathcal{J}\mathcal{D}G$ is a strictly positive element of $M_{n \times n}(M \bar{\otimes} M^{op})$.

In particular, there are trace-preserving injections $C^*(\tau_V) \subset C_{\text{red}}^*(\mathbb{F}_n)$ and $W^*(\tau_V) \subset L(\mathbb{F}_n)$.

If the map $\beta \mapsto W_\beta$ is analytic, then Y_1, \dots, Y_n are also analytic in β . Furthermore, $\|Y_j - X_j\|_{A'}$ vanishes as $\|W\|_{A+1}$ goes to zero.

Proof. We use Theorem 3.15 and set $G = \frac{1}{2} \sum X_j^2 + g$ and the discussion before the statement of the present theorem. The only thing left to prove is that the Hessian of G is strictly positive. But the Hessian of G is given by

$$\mathcal{J}\mathcal{D}G = 1 + \mathcal{J}\mathcal{D}g.$$

In the proof of Theorem 3.15, we have chosen C in such a way that $\|\mathcal{J}\mathcal{D}g\|_{A' \otimes_\pi A'} < 1$, which means that $\mathcal{J}\mathcal{D}G$ is strictly positive. \square

4. SOME APPLICATIONS.

4.1. Analyticity of the solution to the Schwinger-Dyson equation. Under the hypothesis of Theorem 3.16, we deduce that if $\beta \rightarrow W_\beta$ is an analytic family of potentials and $V_\beta = \frac{1}{2} \sum X_j^2 + W_\beta$, then there exists an analytic family of elements Y_β whose law τ_β satisfies the Schwinger-Dyson equation for V_β . We deduce the following corollary, which was already proved in [GS09], by a different method:

Corollary 4.1. *Let $P \in \mathcal{A}$ be a fixed polynomial. Then $\tau_\beta(P)$ is analytic in β in a neighborhood of the origin.*

4.2. Isomorphism results. Applying the implicit function theorem, we can improve Theorem 3.16 at the expense of possibly choosing a smaller bound on W . Indeed, we let $(Y_1, \dots, Y_n) = f = \mathcal{D}g$ be as in Theorem 3.16, we can always assume that $Y_j(0, \dots, 0) = 0$ by replacing each Y_j with $Y_j - Y_j(0, \dots, 0)$. Thus we may assume that $f(0) = 0$.

Theorem 4.2. *Let $A > A' > 4$, and let $X_1, \dots, X_n \in (M, \tau)$ be semicircular variables. Then there exists a universal constant $C = C(A, A') > 0$ so that whenever $W \in \mathcal{A}^{(A+1)}$ satisfies $\|W\|_{A+1} < C$, there is an element $G \in \mathcal{A}^{(A')}$ so that:*

- (1) *If we set $Y_j = \mathcal{D}_j G$, then $Y_1, \dots, Y_n \in \mathcal{A}^{(A')}$ has law τ_V , with $V = \frac{1}{2} \sum X_j^2 + W$;*
- (2) *$X_j = H_j(Y_1, \dots, Y_n)$ for some $H \in \mathcal{A}^{(A')}$;*
- (3) *the Hessian $\mathcal{J} \mathcal{D}G$ is a strictly positive element of $M_{n \times n}(M \bar{\otimes} M^{op})$.*

In particular, there are trace-preserving isomorphisms

$$C^*(\tau_V) \cong C^*(X_1, \dots, X_n), \quad W^*(\tau_V) \cong L(\mathbb{F}_n).$$

Proof. Fix $B \in (A', A)$. By Theorem 3.16, we can write

$$Y = X + f(X)$$

with $\beta = \|f\|_B$ going to zero as $C = \|W\|_{A+1}$ goes to zero. Moreover, as X are semicircular variables, their norms are bounded by 2 and hence for C (and thus β) small enough, Y is bounded by A' so that we can apply Corollary 2.4 with $A' < B$ to find a function $H \in (\mathcal{A}^{(A')})^n$ so that $X = H(Y)$. □

4.2.1. Isomorphism class of algebras with analytic conjugate variables that are close to the generators.

Corollary 4.3. *Let $X_1, \dots, X_n \in (M, \tau)$ be generators of a von Neumann algebra M and assume that the Fisher information $\Phi^*(X_1, \dots, X_n)$ is finite. Assume moreover that $\xi_j^* = \xi_j = \mathcal{D}_j^*(X_j)$ belongs to $\mathcal{A}^{(A+1)}$ for some $A > 4$.*

Then there exists a universal constant $C = C(A)$ so that if $\|\xi_j - X_j\|_{A+1} < C$, then $M \cong L(\mathbb{F}_n)$ and also $C^(X_1, \dots, X_n) \cong C^*(S_1, \dots, S_n)$ where S_1, \dots, S_n are free semicircular variables.*

Proof. Letting $V = \frac{1}{2} \Sigma \left(\sum_{j=1}^n X_j \xi_j + \xi_j X_j \right)$, one gets that $\xi_j = \mathcal{D}_j V$ (cf. [Voi00]). Thus we can write $V = \frac{1}{2} \sum X_j^2 + W$; then for small enough $\|\xi_j - X_j\|_{A+1}$, $W = W^*$ satisfies the hypothesis of Theorem 4.2. □

4.2.2. *Voiculescu's conjecture with polynomial potentials.* The following corollary is a partial answer to a conjecture of Voiculescu [Voi06, p. 240] (the full statement of Voiculescu's conjecture is that the isomorphism should hold for arbitrary polynomials W and arbitrary β ; however, even uniqueness of a free Gibbs state is unknown in that generality!)

Corollary 4.4. *Let $A > 4$ and let $W \in \mathcal{A}^{(A)}$ be a self-adjoint power series. Let $V_\beta = \sum X_j^2 + \beta W$, and let τ_β be a trace satisfying the Schwinger-Dyson equation with potential V_β . Then there exists a $\beta_0 > 0$ so that $W^*(\tau_\beta) \cong W^*(\tau_0) \cong L(\mathbb{F}_n)$ and $C^*(\tau_\beta) \cong C^*(\tau_0) = C^*(S_1, \dots, S_n)$ for all $-\beta_0 < \beta < \beta_0$ (here S_1, \dots, S_n is a free semicircular system).*

4.2.3. *q -deformed free group factors.* In [BS91], Bozejko and Speicher have introduced a family of von Neumann algebras $\Gamma_q(\mathbb{R}^n)$, which are “ q -deformations” of free group factors. For $q = 0$, $\Gamma_q(\mathbb{R}^n) \cong L(\mathbb{F}_n)$, but the question of whether an isomorphism like this holds for $q \neq 0$ remained open. Despite being much-studied ([BS91, BKS97, Nou04, Śni01, Śni04, Shl04, Ric05, Shl09, KN11, Dab10, Avs11] is an incomplete list of results about these factors), the question of the isomorphism class of these algebras for all values of q remains elusive. Nonetheless, we are able to settle it for small q (depending on $n < \infty$):

Corollary 4.5. *Let $\Gamma_q(\mathbb{R}^n)$ be the von Neumann algebra generated by a q -semicircular system $S_1^{(q)}, \dots, S_n^{(q)}$, $n \in \{2, 3, \dots\}$. Then there exists a $0 < q_0 = q_0(n)$ depending on n , so that $\Gamma_q(\mathbb{R}^n) \cong \Gamma_0(\mathbb{R}^n) \cong L(\mathbb{F}_n)$ and $C^*(S_1^{(q)}, \dots, S_n^{(q)}) = C^*(S_1^{(0)}, \dots, S_n^{(0)})$ for all $|q| < q_0$.*

Proof. This follows from the fact [Dab10, Theorem 34] that the conjugate variables $\xi_j = \xi_j^* = \mathcal{D}V$ to $(S_1^{(q)}, \dots, S_n^{(q)})$ exist and that for some $A > 5$, $\|\xi_j - X_j\|_A \rightarrow 0$ as $q \rightarrow 0$. \square

4.2.4. *Free entropy.* Note that under the assumptions of Theorem 4.2, we end up expressing X_j 's as a convergent power series in n semicircular variables S_1, \dots, S_n . By the change of variables formula of Voiculescu [Voi94], we get:

Corollary 4.6. *Let $V = \frac{1}{2} \sum X_j^2 + W$, and assume the hypothesis of Theorem 4.2. Let τ_V be the free Gibbs state with potential V . Then*

$$\chi(\tau_V) = \chi(\tau) + \tau \otimes \tau[\mathrm{Tr}(\log(\mathcal{J} \mathcal{D}G))]$$

where τ is the semicircle law and G is as in Theorem 4.2.

4.3. **Monotone transport for random matrix models.** Let $V = \frac{1}{2} \sum X_j^2 + W$ and τ_V be the free Gibbs state with potential V . Assume that W satisfies the hypothesis of Theorem 4.2 for some $A > A' > 4$ and let $F = (F_1, \dots, F_n) = \mathcal{D}G$ be the map in $(\mathcal{A}^{(A')})^n$ constructed in that theorem. Let us assume that W is small enough so that $\|F\|_{A'} \leq A$. We let

$$V(X_1, \dots, X_n) = \frac{1}{2} \sum X_j^2 + W(X_1, \dots, X_n).$$

on $\{\max_i \|X_i\| \leq A'\}$ and $V = +\infty$ otherwise. Note that V is strictly convex if $\|W\|_{A+1}$ is small enough.

Let $\mu_V^{(N)}$ be the measure on $(M_{N \times N}^{sa})^n$ given by

$$d\mu_V^{(N)} = \frac{1}{Z_N^V} \exp(-N \mathrm{Tr}(V(A_1, \dots, A_n))) dA_1 \dots dA_n$$

where we set $dA_1 \dots dA_n =$ Lebesgue measure and $M_{N \times N}^{sa}$ the set of $N \times N$ self-adjoint matrices. Consider also the Gaussian measure

$$d\mu^{(N)} = \frac{1}{Z_N} \exp \left[-N \text{Tr} \left(\frac{1}{2} \sum A_j^2 \right) \right] dA_1 \dots dA_n.$$

We denote by $\chi : \mathbb{R} \rightarrow [-A', A']$ a smooth cutoff function so that $\chi(x) = x$ for $x \in [-4, 4]$, $|\chi(x)| \leq A'$ and χ has a uniformly bounded derivative. On the space of n -tuples of $N \times N$ self-adjoint matrices $(M_{N \times N}^{sa})^n$ with norm bounded by 4, consider the functional calculus map

$$\tilde{F}^{(N)} : (A_1, \dots, A_n) \mapsto (F_j(\chi(A_1), \dots, \chi(A_n)))_{j=1}^n.$$

Finally, let $F^{(N)} = \nabla \Psi^{(N)}$ be the unique monotone transport map on $(M_{N \times N}^{sa})^n$ so that

$$F_*^{(N)} \mu^{(N)} = \mu_V^{(N)}.$$

Theorem 4.7. *Let $A > A' > 4$ as above. Then there exists a constant c (smaller than that of Theorem 4.2) so that for any $\|W\|_{A+1} < c$, the following statement hold:*

(i) *With the above notation,*

$$\lim_{N \rightarrow \infty} \int \frac{1}{N} \text{Tr} \left(\left[F^{(N)} - \tilde{F}^{(N)} \right]^2 \right) d\mu^{(N)} = 0.$$

In other words, as $N \rightarrow \infty$, the “entrywise” monotone transport $F^{(N)}$ is well-approximated as $N \rightarrow \infty$ by the “matricial” functional calculus map $\tilde{F}^{(N)}$.

(ii) *Assume in addition that W is (c, A) -convex for some $c > 0$. Then $\mathcal{J} F \leq 1$. Furthermore, $\frac{1}{N} \text{Tr} \left(\left[F^{(N)} - \tilde{F}^{(N)} \right]^2 \right)$ vanishes almost surely as N goes to infinity.*

Proof. Let us put, for a probability measure ν on \mathbb{R}^M with density $p(x_1, \dots, x_M) dx_1 \dots dx_M$, and $U : \mathbb{R}^M \rightarrow \mathbb{R} \cup \{+\infty\}$ a convex function,

$$\begin{aligned} H(\nu) &= \int p(x_1, \dots, x_M) \log p(x_1, \dots, x_M) dx_1 \dots dx_M \\ H_U(\nu) &= \int p(x_1, \dots, x_M) \log p(x_1, \dots, x_M) dx_1 \dots dx_M + \int U d\nu. \end{aligned}$$

If h is an invertible transformation with positive-definite Jacobian $\text{Jac } h$, then

$$\begin{aligned} H(h_*\nu) &= H(\nu) - \int \log \det \text{Jac } h \, d\nu \\ &= H(\nu) - \int \text{Tr} \log \text{Jac } h \, d\nu. \end{aligned}$$

Both entropies $H_U(\nu)$ and $H(\nu)$ are convex functions of the density of ν . Moreover, $H_U(h_*\nu)$ and $H(h_*\nu)$ are convex in h if U is convex.

It is well-known that $H_U(\nu)$ is minimized precisely by the Gibbs measure with potential U :

$$H_U(\nu) = \inf_{\nu'} H_U(\nu') \iff d\nu(x_1, \dots, x_M) = \frac{1}{Z} \exp(-U(x_1, \dots, x_M)) dx_1 \dots dx_M.$$

Similarly, for $V \in \mathcal{A}$, let $\chi(\tau)$ be Voiculescu’s microstates free entropy, and put $\chi_V(\tau) = \chi(\tau) - \tau(V)$.

Then using [Voi02b, GS09, Gui09] one has

$$\chi_V(\tau) = \sup_{\tau'} \chi_V(\tau') \iff \tau \text{ is the free Gibbs law with potential } V.$$

Moreover, if we set $V^{(N)}(A_1, \dots, A_n) = N \text{Tr}(V(A_1, \dots, A_n))$, then following [GS09, proof of Theorem 5.1] or [GMS06, Section 3.3] we see that $\mu_V^{(N)}(\frac{1}{N} \text{Tr}(P))$ converges towards $\tau_V(P)$ for all polynomials P , in particular the limit does not depend on the cutoff provided $\|W\|_{A+1}$ is small enough. In fact, as $\mu_V^{(N)}$ has a strictly log-concave density, Brascamp-Lieb inequalities allow one to show that the matrices under $\mu_V^{(N)}$ are bounded by 4 with overwhelming probability. Moreover, by definition of the entropy, we find that

$$\begin{aligned} \chi_V(\tau_V) &= \lim_{N \rightarrow \infty} \left(\frac{1}{N^2} \log Z_N^V + \frac{n}{2} \log N \right) \\ &= \lim_{N \rightarrow \infty} \left(\frac{n}{2} \log N - \frac{1}{N^2} H_{V^{(N)}}(\mu_V^{(N)}) \right) \end{aligned}$$

Furthermore, we claim that, if $\tilde{\mu}_V^N = \tilde{F}_*^{(N)} \mu^{(N)}$, we have

$$(4.1) \quad \lim_{N \rightarrow \infty} \frac{1}{N^2} H_{V^{(N)}}(\tilde{\mu}_V^{(N)}) - \frac{n}{2} \log N = -\chi_V(\tau_V).$$

To see this, we first note that

$$\lim_{N \rightarrow \infty} \frac{1}{N^2} \int V^{(N)} d\tilde{\mu}_V^{(N)} = \tau_V(V).$$

Moreover,

$$\frac{1}{N^2} H(\tilde{\mu}_V^{(N)}) = \frac{1}{N^2} H(\mu^{(N)}) - \mathbb{E}_{\mu^{(N)}} \left[\frac{1}{N^2} \text{Tr} \log \text{Jac} \tilde{F}^{(N)} \right].$$

On the other hand (see e.g. [Gui09]), because of concentration phenomena, as $\mu^{(N)}$ has a strictly log-concave density, if we take τ to be the semicircle law, then

$$\mathbb{E}_{\mu^{(N)}} \left[\frac{1}{N^2} \text{Tr} \log \text{Jac} \tilde{F}^{(N)} \right] \rightarrow \tau \otimes \tau(\text{Tr} \log \mathcal{J} F).$$

Thus

$$\frac{1}{N^2} H(\tilde{\mu}_V^{(N)}) - \frac{n}{2} \log N \rightarrow -\chi(\tau) - \tau \otimes \tau(\text{Tr} \log \mathcal{J} F) = -\chi(\tau_V)$$

so that using Corollary 4.6 gives (4.1). As a consequence, we have proved that

$$\lim_{N \rightarrow \infty} \frac{1}{N^2} (H_{V^{(N)}}(\tilde{\mu}_V^{(N)}) - H_{V^{(N)}}(\mu_V^{(N)})) = 0.$$

By convexity of entropy as a function of the transport maps, for $\varepsilon \in [0, 1]$,

$$\begin{aligned} & H_{V^{(N)}} \left(\left[(1 - \varepsilon) F^{(N)} + \varepsilon \tilde{F}^{(N)} \right]_{\#} \mu^{(N)} \right) - H_{V^{(N)}}(\mu_V^{(N)}) \\ & \leq (1 - \varepsilon) H_{V^{(N)}}(F_*^{(N)} \mu^{(N)}) + \varepsilon H_{V^{(N)}}(\tilde{F}_*^{(N)} \mu^{(N)}) - H_{V^{(N)}}(\mu_V^{(N)}) \\ & = \varepsilon (H_{V^{(N)}}(\tilde{\mu}_V^{(N)}) - H_{V^{(N)}}(\mu_V^{(N)})). \end{aligned}$$

On the other hand, let $\Delta^{(N)} = F^{(N)} - \tilde{F}^{(N)}$ and set

$$D_N = \partial_{\varepsilon} \Big|_{\varepsilon=0} H_{V^{(N)}} \left([F^{(N)} + \varepsilon \Delta^{(N)}]_* \mu^{(N)} \right).$$

Since $H_{V^{(N)}} \left(\left[(1 - \varepsilon)F^{(N)} + \varepsilon\tilde{F}^{(N)} \right]_* \mu^{(N)} \right)$ has an absolute minimum at $\varepsilon = 0$, we deduce that $D_N \geq 0$.

Furthermore, assuming that $\|W\|_{A+1}$ is small enough so that on the set $\{\max_{1 \leq i \leq n} \|A_i\|_\infty \leq A\}$ the Hessian $\text{Hess}(V^{(N)})$ is bounded from below by cN times the identity operator for some $c > 0$, and noting that by definition the image of $\tilde{F}^{(N)}$ is bounded by A whereas the image of $F^{(N)}$ composed with χ is also bounded by A ,

$$\begin{aligned}
& H_{V^{(N)}} \left(\left[(1 - \varepsilon)F^{(N)} + \varepsilon\tilde{F}^{(N)} \right]_* \mu^{(N)} \right) - \varepsilon D_N - H_{V^{(N)}}(\mu_V^{(N)}) \\
&= \int_0^\varepsilon (\varepsilon - t) \partial_t^2 H_{V^{(N)}} \left([F^{(N)} + t\Delta^{(N)}]_* \mu^{(N)} \right) dt \\
(4.2) \quad &= \mathbb{E}_{\mu^{(N)}} \int_0^\varepsilon t \left\{ \text{Tr} \left[\left(\frac{\text{Jac}(\Delta^{(N)})}{\text{Jac}(tF^{(N)} + (1-t)\tilde{F}^{(N)})} \right)^2 \right] \right. \\
&\quad \left. + \text{Tr} \left(\text{Hess}(V^{(N)}(F^{(N)} + t\Delta^{(N)}))(\Delta^{(N)})^2 \right) \right\} dt \\
&\geq \frac{\varepsilon^2}{2} cN \mathbb{E}_{\mu^{(N)}} (\text{Tr}((\Delta^{(N)})^2)).
\end{aligned}$$

We thus get (recalling that $D_N \geq 0$)

$$\begin{aligned}
\frac{c}{2} N \varepsilon^2 \mathbb{E}_{\mu^{(N)}} \text{Tr}((\Delta^{(N)})^2) &\leq \varepsilon (H_{V^{(N)}}(\tilde{\mu}_V^{(N)}) - H_{V^{(N)}}(\mu_V^{(N)})) - \varepsilon D_N \\
&\leq \varepsilon (H_{V^{(N)}}(\tilde{\mu}_V^{(N)}) - H_{V^{(N)}}(\mu_V^{(N)})).
\end{aligned}$$

Since

$$\alpha_N = \frac{1}{N^2} \left| H_{V^{(N)}}(\tilde{\mu}_V^{(N)}) - H_{V^{(N)}}(\mu_V^{(N)}) \right| \rightarrow 0$$

we can now choose $\varepsilon = \sqrt{\alpha_N} \rightarrow 0$ to conclude that

$$(4.3) \quad \frac{c}{2} \mathbb{E}_{\mu^{(N)}} \left[\frac{1}{N} \text{Tr}((\Delta^{(N)})^2) \right] \leq \alpha_N^{1/2}$$

which completes the proof of the first point.

Let us now assume that W is (c, A) -convex for some $c > 0$.

Let $W^{(N)} = N \text{Tr}(W(A_1, \dots, A_n))$ be defined on matrices satisfying $\max_j \|A_j\|_\infty \leq A'$ and be infinite otherwise. Then $W^{(N)}$ is convex. It then follows from Caffarelli's results [Caf00, Vil03] that the optimal transport map $F^{(N)}$ taking the Gaussian measure $\mu^{(N)}$ to the measure with density $Z_N^{-1} \exp(-\frac{1}{2}N \sum \text{Tr}(A_j^2) + W^{(N)}(A_1, \dots, A_n))$ has Jacobian uniformly bounded by 1.

The map $(A_1, \dots, A_n) \rightarrow \frac{1}{N} \text{Tr}((\Delta^{(N)}(A_1, \dots, A_n))^2)$ can be viewed as the composition of the map $R : (A_1, \dots, A_n) \rightarrow \Delta^{(N)}(A_1, \dots, A_n)$ and the map $Q : (A_1, \dots, A_n) \rightarrow \frac{1}{N} \sum \text{Tr}(A_j^2)$. The Jacobian of R is given (on (A_1, \dots, A_n) with $\max_j \|A_j\| < A$) by $\mathcal{J} \mathcal{D}g - \text{Hess} \Psi^{(N)}$, where $\Psi^{(N)}$ is such that $F^{(N)} = \nabla \Psi^{(N)}$. If $\max_j \|A_j\|_\infty < A$, this is bounded (as an operator on Hilbert spaces $M_{N \times N}^n \rightarrow M_{N \times N}^n$ endowed Hilbert Schmidt norms $\|A\| = \sum_j \text{Tr}(A_j^* A_j)$) because g is a power series and because of Caffarelli's bound. Hence the map R is Lipschitz with a uniform Lipschitz constant on the set where $\max_j \|A_j\|_\infty < A$. The map Q is Lipschitz with Lipschitz constant of the form C/\sqrt{N} (see [Gui09, Lemma 6.2]).

Therefore, by concentration inequalities and (4.3), we deduce

$$\lim_{N \rightarrow \infty} \frac{1}{N} \text{Tr}((\Delta^{(N)})^2) = 0 \quad \mu^{(N)\text{-a.s.}}$$

We next come back to (4.2) and observe that $\text{Jac}(\tilde{F}^{(N)})$ is bounded above uniformly by some constant M_0 (as a small smooth perturbation of the identity). Furthermore, $\text{Jac}(\tilde{F}^{(N)})$ is bounded above by 1 by [Caf00]. Thus we get that for all $\varepsilon < 1/2$ and some constant $M = \max(M_0, 1)$ independent of ε ,

$$\mu^{(N)} \left[\frac{1}{N^2} \text{Tr} \{ (\text{Jac}(\Delta^{(N)}))^2 \} \right] \leq M \alpha_N / \varepsilon.$$

Taking once again $\varepsilon = \sqrt{\alpha_N} \rightarrow 0$, we obtain that

$$\mu^{(N)} \left[\frac{1}{N^2} \text{Tr} \left\{ \left(\text{Jac}(F^{(N)}) - \text{Jac}(\tilde{F}^{(N)}) \right)^2 \right\} \right] \rightarrow 0.$$

We now apply [Caf00] to conclude that

$$(4.4) \quad \text{Jac } F^{(N)} \leq 1.$$

Let $Y_N = \text{Jac}(F^{(N)})$, $\tilde{Y}_N = \text{Jac}(\tilde{F}^{(N)})$ be random variables taking values in the space $M_{n \times n}(\text{End}(M_{N \times N}^{sa})) \cong M_{n \times n}(M_{N \times N} \otimes M_{N \times N})$ endowed with the normalized trace $\frac{1}{nN^2} \text{Tr} \otimes \text{Tr} \otimes \text{Tr}$. Both Y_N and \tilde{Y}_N are bounded in operator norm, and consequently define elements Y and \tilde{Y} in the ultraproduct von Neumann algebra $\prod_N^\omega M_{n \times n}(M_{N \times N} \otimes M_{N \times N})$. Since $\|Y_N - \tilde{Y}_N\|_2 \rightarrow 0$, $Y = \tilde{Y}$.

Once again, because of concentration,

$$\mu^{(N)} \left[\frac{1}{N^2} \text{Tr} \left\{ (\text{Jac}(\tilde{F}^{(N)}))^p \right\} \right] \rightarrow \tau \otimes \tau \text{Tr}((\mathcal{J} F)^p)$$

so that the spectrum of $\mathcal{J} F$ is the same as that of \tilde{Y} (and so the same as that of Y).

By (4.4), the spectrum of Y_N lies in the interval $[0, 1]$. But this implies that also the spectrum of the limiting operator Y is contained in the same set. Thus $\mathcal{J} F \leq 1$. \square

5. OPEN QUESTIONS.

We list some open questions that are raised by our results.

- (1) In the classical case, Brenier's theorem [Bre91, Vil03] asserts much more than the statement of our main theorem: the classical analog of the map $\mathcal{D}g$ gives *optimal* transport from τ to τ_V for quadratic Wasserstein distance. It would be nice to understand if the same holds true in the non-commutative case (see [BV01] for the extension of the notion of the Wasserstein distance to non-commutative random variables). Note that the map we construct is optimal in the single-variable case $n = 1$.
- (2) Brenier gave a heuristic derivation of his theorem through a very general "polar factorization" theorem. Does a theorem like that hold in the non-commutative case? There is an infinitesimal analog of his decomposition (related to the classical Helmholtz decomposition of vector fields) which has been extensively studied by Voiculescu in [Voi02a]. Can non-commutative monotone transport be also obtained in the same way?
- (3) Does the positivity condition $\mathcal{J} \mathcal{D}G \in \{F > 0 : F \in M_{n \times n}(\tau \otimes \tau^{\text{op}})\}$ on the "Hessian" of G translate into any kind of convexity properties of G ?

- (4) What happens to our map $F = \mathcal{D}G$ in the case that V is not strictly convex? It can be seen that in the absence of the bounds on W the isomorphism (or even the embedding $C^*(\tau_V) \subset C^*(S_1, \dots, S_n)$) fails to exist at least on C^* -level. This is due to the fact that for certain non-convex polynomials, solutions to the Schwinger-Dyson equations may lead to C^* -algebras with non-trivial projections. But the free semicircular system generates a projectionless C^* -algebra. Thus failure of convexity of $V = \frac{1}{2} \sum X_j^2 + W$ must be “visible” as a defect of regularity of the transport map F .
- (5) Is the monotone transport map unique? More precisely, let X_1, \dots, X_n be a semicircular family, and assume that $Y = (Y_1, \dots, Y_n)$ and $Y' = (Y'_1, \dots, Y'_n)$ both belong to the L^2 -closure of $\{\mathcal{D}g(X_1, \dots, X_n) : \mathcal{J} \mathcal{D}g \geq 0\}$. If the law of Y is the same as the law of Y' , is $Y = Y'$?

We can prove a uniqueness statement for our monotone transport if we assume more on the transport map and the “target” n -tuple Y_1, \dots, Y_n :

Theorem 5.1. *Let X_1, \dots, X_n be a semicircular family, and assume that $Y = (Y_1, \dots, Y_n)$ and $Y' = (Y'_1, \dots, Y'_n)$ both belong to the L^2 -closure of $\{\mathcal{D}g : \mathcal{J} \mathcal{D}g \geq 0\}$. Assume moreover that Y and Y' are both invertible non-commutative power series in X_1, \dots, X_n . If Y and Y' have as their laws the same free Gibbs law τ_V with $V = \frac{1}{2} \sum X_i^2 + W$, W small enough, then $Y = Y'$.*

Proof. The assumption that Y, Y' are invertible non-commutative power series allow us to apply Voiculescu’s change of variable formula for free entropy [Voi94]:

$$\begin{aligned} \chi(Y) &= \chi(X) + \tau \otimes \tau(\text{Tr}(\log \mathcal{J}Y)) \\ \chi(Y') &= \chi(X) + \tau \otimes \tau(\text{Tr}(\log \mathcal{J}Y')). \end{aligned}$$

Since Y and Y' have the same law, they have the same free entropy (which is finite). Thus we conclude that

$$\tau \otimes \tau(\text{Tr}(\log \mathcal{J}Y)) - \tau(V(Y)) = \tau \otimes \tau(\text{Tr}(\log \mathcal{J}Y')) - \tau(V(Y')).$$

Let

$$\psi(Y) = \chi(Y) - \chi(X) - \tau(V(Y)).$$

Then ψ is maximal iff Y has the law τ_V (see e.g. [Voi02b], we sketch the argument for completeness: if one replaces Y by $Y + \varepsilon P(Y)$ for some polynomials P_1, \dots, P_n , then

$$\psi(Y + \varepsilon P(Y)) = \psi(Y) + \varepsilon \{\tau \otimes \tau(\text{Tr}(\mathcal{J}YP)) - \tau(\mathcal{D}V(Y)P)\} + O(\varepsilon^2),$$

and so any maximizer to Ψ satisfies the Schwinger-Dyson equation and has the same law as Y .)

It follows that

$$\tau \otimes \tau(\text{Tr}(\log \mathcal{J}Y)) - \tau(V(Y)) = \chi(Y) - \chi(X) - \tau(V(Y)) = \max_Z \chi(Z) - \chi(X) - \tau(V(Z)),$$

so a fortiori

$$\tau \otimes \tau(\text{Tr}(\log \mathcal{J}Y)) - \tau(V(Y)) = \sup_{Y \in \{\mathcal{D}g : \mathcal{J} \mathcal{D}g \geq 0\}} \tau \otimes \tau(\text{Tr}(\log \mathcal{J}Y)) - \tau(V(Y)).$$

But since $\tau \otimes \tau(\text{Tr}(\log \mathcal{J}Y)) - \tau(V(Y))$ is strictly concave in Y for $\mathcal{J}Y$ in the positive cone of $M_{n \times n}(M \bar{\otimes} M^{op})$ (here $M = W^*(X_1, \dots, X_n)$), it follows that there is at most one

n -tuple Y in the closure of $Y \in \{\mathcal{D}g : \mathcal{J} \mathcal{D}g \geq 0\}$ which gives this maximal value. Thus $Y = Y'$. \square

REFERENCES

- [AGZ10] G. Anderson, A. Guionnet, and O. Zeitouni, *An introduction to random matrices*, Cambridge University Press, 2010.
- [Avs11] S. Avsec, *Strong solidity of the q -Gaussian algebras for all $-1 < q < 1$* , Preprint, 2011.
- [Bia97] P. Biane, *Segal–Bargmann transform, functional calculus on matrix spaces and the theory of semi-circular and circular systems*, J. Funct. Anal. **144** (1997), 232–286.
- [BG13] G. Borot, and A. Guionnet, *Asymptotic expansion of β matrix models in the one-cut regime*, Comm. Math. Phys., **317** (2013), YEAR = 2013, 447–483.
- [BKS97] M. Bożejko, B. Kummerer, and R. Speicher, *q -Gaussian processes: non-commutative and classical aspects*, Comm. Math. Phys **185** (1997), 129–154.
- [Bre91] Y. Brenier, *Polar factorization and monotone rearrangement of vector-valued functions*, Comm. Pure Appl. Math **44** (1991), 375–417.
- [BS91] M. Bożejko and R. Speicher, *An example of a generalized Brownian motion*, Commun. Math. Phys. **137** (1991), 519 – 531.
- [BV01] P. Biane and D.-V. Voiculescu, *A free probability analogue of the Wasserstein metric on the trace-state space*, Geom. Funct. Anal. **11** (2001), no. 6, 1125–1138. MR MR1878316 (2003d:46087)
- [Caf00] L. Caffarelli, *Monotonicity properties of optimal transportation and the FKG and related inequalities*, Comm. Math. Phys **214** (2000), 547–563.
- [Dab10] Y. Dabrowski, *A free stochastic partial differential equation*, Preprint, arXiv.org:1008:4742, 2010.
- [GMS06] A. Guionnet and E. Maurel-Segala, *Combinatorial aspects of matrix models*, ALEA Lat. Am. J. Probab. Math. Stat. **1** (2006), 241–279. MR 2249657 (2007g:05087)
- [GMS07] A. Guionnet and E. Maurel-Segala, *Second order asymptotics for matrix models*, Ann. Probab. **35** (2007), 2160–2212.
- [GS09] A. Guionnet and D. Shlyakhtenko, *Free diffusions and matrix models with strictly convex interaction*, Geom. Funct. Anal. **18** (2009), 1875–1916.
- [Gui06] A. Guionnet, *Random matrices and enumeration of maps*, Proceedings Int. Cong. Math. **3** (2006), 623–636.
- [Gui09] A. Guionnet, *Large random matrices: lectures on macroscopic asymptotics*, Lecture Notes in Mathematics, vol. 1957, Springer-Verlag, Berlin, 2009, Lectures from the 36th Probability Summer School held in Saint-Flour, 2006. MR 2498298 (2010d:60018)
- [KN11] M. Kennedy and A. Nica, *Exactness of the Fock space representation of the q -commutation relations*, Comm. Math. Phys. **308** (2011), 115–132.
- [Nou04] A. Nou, *Non injectivity of the q -deformed von Neumann algebra*, Math. Annalen **330** (2004), 17–38.
- [OP10] N. Ozawa and S. Popa, *On a class of II_1 factors with at most one Cartan subalgebra*, Ann. of Math **172** (2010), 713–749.
- [Oza04] N. Ozawa, *There is no separable universal II_1 -factor*, Proc. Amer. Math. Soc. **132** (2004), 487–490.
- [Ric05] E. Ricard, *Factoriality of q -Gaussian von Neumann algebras*, Comm. Math. Phys. **257** (2005), 659–665.
- [Shl04] D. Shlyakhtenko, *Some estimates for non-microstates free entropy dimension with applications to q -semicircular families*, Int. Math. Res. Notices **51** (2004), 2757–2772.
- [Shl09] ———, *Lower estimates on microstates free entropy dimension*, Analysis and PDE **2** (2009), 119–146.
- [Śni01] P. Śniady, *Gaussian random matrix models for q -deformed Gaussian variables*, Comm. Math. Phys. **216** (2001), no. 3, 515–537. MR 2003a:81096
- [Śni04] ———, *Factoriality of Bożejko-Speicher von Neumann algebras*, Comm. Math. Phys **246** (2004), 561–567.
- [Vil03] C. Villani, *Topics in optimal transportation*, Graduate Studies in Mathematics, Vol. 58, AMS, Providence, RI, 2003.
- [Voi94] D.-V. Voiculescu, *The analogues of entropy and of Fisher’s information measure in free probability theory II*, Invent. Math. **118** (1994), 411–440.

- [[Voi98 ———, *The analogues of entropy and of Fisher's information measure in free probability*, *V*, *Invent. Math.* **132** (1998), 189–227.
- [[Voi00 ———, *A note on cyclic gradients*, *Indiana Univ. Math. J.* **49** (2000), 837–841.
- [[Voi02a ———, *Cyclomorphy*, *Int. Math. Research Notices* **No. 6** (2002), 299–332.
- [[Voi02b ———, *Free entropy*, *Bull. London Math. Soc.* **34** (2002), no. 3, 257–278. MR 2003c:46077
- [[Voi06 ———, *Symmetries arising from free probability theory*, *Frontiers in Number Theory, Physics, and Geometry I* (Pierre Cartier, Bernard Julia, Pierre Moussa, and Pierre Vanhove, eds.), Springer Berlin Heidelberg, 2006, pp. 231–243.