

APPENDIX A

Mathematical supplement

A.1. Matrices and matrix norms

We consider $n \times n$ matrices with complex entries. Such a matrix a is *hermitian* if it equals its conjugate transpose, $a = a^*$; it is *non-negative* if $\langle \psi, a\psi \rangle \geq 0$ for all vectors ψ . A hermitian matrix a is diagonalizable: there is a unitary matrix U (meaning $U^* = U^{-1}$) such that U^*aU is diagonal. The eigenvalues of a hermitian matrix a are all real. A hermitian matrix is non-negative if and only if all its eigenvalues are non-negative.

If $D = \text{diag}(d_1, \dots, d_n)$ is a diagonal matrix, and f is a real-valued function whose domain contains all the entries d_i of D , then we define

$$f(D) = \text{diag}(f(d_1), \dots, f(d_n)). \quad (\text{A.1})$$

This definition extends to hermitian matrices via diagonalization: If a is an hermitian matrix and f is a real-valued function whose domain contains the spectrum of a , and a is diagonalized as $a = U \text{diag}(\lambda_1, \dots, \lambda_n) U^*$, then we define

$$f(a) = U \text{diag}(f(\lambda_1), \dots, f(\lambda_n)) U^*. \quad (\text{A.2})$$

For an arbitrary $n \times n$ matrix a , the matrix a^*a is hermitian. We define the *absolute value* by $|a| = (a^*a)^{\frac{1}{2}}$, where the square root of a^*a is defined as above. Then $|a|$ is hermitian. The p -norm of a matrix is then defined as

$$\|a\|_p = (\text{Tr } |a|^p)^{1/p}. \quad (\text{A.3})$$

We also define the *operator norm* of a by

$$\|a\| = \sup_{\psi \neq 0} \frac{\|a\psi\|}{\|\psi\|} \quad (\text{A.4})$$

where $\|\psi\| = (\sum_i |\psi_i|^2)^{1/2}$ is the usual vector norm. Note that $\|a\|_\infty = \lim_{p \rightarrow \infty} \|a\|_p = \|a\|$, the operator norm of a .

A.2. Hölder inequality for matrices

PROPOSITION A.1 (Hölder inequality for matrices). *If $1 \leq p, q, r \leq \infty$ with $\frac{1}{p} + \frac{1}{q} = \frac{1}{r}$, we have*

$$\|ab\|_r \leq \|a\|_p \|b\|_q.$$

It follows from a simple induction that

$$\left\| \prod_{j=1}^n a_j \right\|_r \leq \prod_{j=1}^n \|a_j\|_{p_j} \quad (\text{A.5})$$

whenever $1 \leq r, p_1, \dots, p_n$ with $\sum_{j=1}^n \frac{1}{p_j} = \frac{1}{r}$.

There are no short proofs in the case of matrices. The proof here is due to Fröhlich [1978] and it uses chessboard estimates. The proof of Proposition A.1 can be found after that of Lemma A.4.

LEMMA A.2 (Chessboard estimate). *For any $n \in \mathbb{N}$ and any matrices a_1, \dots, a_{2n} , we have*

$$|\text{Tr } a_1 \dots a_{2n}| \leq \prod_{i=1}^{2n} \left(\text{Tr } (a_i a_i^*)^n \right)^{1/2n}.$$

PROOF. Since $(a, b) \mapsto \text{Tr } a^* b$ is an inner product, we have the Cauchy–Schwarz inequality: $|\text{Tr } a b|^2 \leq \text{Tr } a^* a \text{Tr } b^* b$. The following inequality follows:

$$|\text{Tr } a_1 \dots a_{2n}|^2 \leq \text{Tr } (a_1 \dots a_n a_n^* \dots a_1^*) \text{Tr } (a_{2n}^* \dots a_{n+1}^* a_{n+1} \dots a_{2n}). \quad (\text{A.6})$$

This allows to use a reflection positivity argument. By replacing a_i with $a_i / \sqrt{\text{Tr } (a_i a_i^*)^n}$ it is enough to prove the inequality for matrices that satisfy $\text{Tr } (a_i a_i^*)^n = 1$; the general result follows from scaling. Note that the set of such matrices is compact.

Let a_1, \dots, a_{2n} be matrices that maximise $|\text{Tr } a_1 \dots a_{2n}|$, with maximum number of matching neighbours $a_{i+1} = a_i^*$. Suppose there exists an index j such that $a_{j+1} \neq a_j^*$. Using cyclicity, we can assume that $j = n$. By the inequality (A.6), $a_1, \dots, a_n, a_n^*, \dots, a_1^*$ and $a_{2n}^*, \dots, a_{n+1}^*, a_{n+1}, \dots, a_{2n}$ are also maximisers. At least one has strictly more matching neighbours, hence a contradiction. The maximum is then $\text{Tr } (a a^*)^n$ for some matrix $a \in \{a_1, \dots, a_n\}$, which is equal to 1. \square

Chessboard estimates allow to prove what is essentially the case $r = 1$ of Hölder's inequality.

COROLLARY A.3. *We have*

$$|\text{Tr } a_1 \dots a_n| \leq \prod_{i=1}^n \|a_i\|_{p_i}$$

for all n and all $p_i \geq 1$ such that $\sum_{i=1}^n \frac{1}{p_i} = 1$.

PROOF. It suffices to consider rational p_i , by continuity. Let ℓ be a positive integer such that $2\ell/p_i$ is integer for all i . Let $a_i = U_i |a_i|$ be the polar decomposition of a_i , and let

$$b_i = |a_i|^{p_i/2\ell}, \quad \hat{b}_i = U_i |a_i|^{p_i/2\ell}. \quad (\text{A.7})$$

Then $a_i = \hat{b}_i b_i^{(2\ell/p_i)-1}$, and we have

$$\begin{aligned} |\mathrm{Tr} a_1 \dots a_n| &= |\mathrm{Tr} \hat{b}_1 \underbrace{b_1 \dots b_1}_{(2\ell/p_1)-1} \dots \hat{b}_n \underbrace{b_n \dots b_n}_{(2\ell/p_n)-1}| \\ &\leq \prod_{i=1}^n (\mathrm{Tr} |a_i|^{p_i})^{1/p_i} \\ &= \prod_{i=1}^n \|a_i\|_{p_i}. \end{aligned} \tag{A.8}$$

The inequality follows from Lemma A.2 and from the identities

$$\mathrm{Tr} (b_i b_i^*)^\ell = \mathrm{Tr} (\hat{b}_i \hat{b}_i^*)^\ell = \mathrm{Tr} |a_i|^{p_i}. \tag{A.9}$$

□

LEMMA A.4. *Let $r, r' \in [1, \infty]$ such that $\frac{1}{r} + \frac{1}{r'} = 1$. Then for any square matrix a , we have*

$$\|a\|_r = \max_{\|c\|_{r'}=1} \mathrm{Tr} c^* a.$$

PROOF. The right side is smaller by Corollary A.3:

$$|\mathrm{Tr} c^* a| \leq \|c\|_{r'} \|a\|_r = \|a\|_r. \tag{A.10}$$

In order to check that this inequality is saturated, let $a = U|a|$ be the polar decomposition of a , and choose $c = \|a\|_r^{1-r} U|a|^{r-1}$. Then $\|c\|_{r'} = 1$ and $\mathrm{Tr} c^* a = \|a\|_r$. □

PROOF OF PROPOSITION A.1. Starting with Lemma A.4 and then using Corollary A.3 with $a_1 = c^*$, $a_2 = a$, $a_3 = b$ and $p_1 = r$, $p_2 = p$, $p_3 = q$, we have

$$\begin{aligned} \|ab\|_r &= \sup_{\|c\|_{r'}=1} \mathrm{Tr} c^* ab \\ &\leq \sup_{\|c\|_{r'}=1} \|c\|_{r'} \|a\|_p \|b\|_q. \end{aligned} \tag{A.11}$$

□

A.3. Trotter and Duhamel

We now review a useful expansion for the exponential of a sum of two non-commuting operators, namely the Duhamel formula.

PROPOSITION A.5 (Lie–Trotter formula). *Let a, b be $n \times n$ matrices. Then*

$$e^{a+b} = \lim_{N \rightarrow \infty} \left(e^{\frac{1}{N}a} e^{\frac{1}{N}b} \right)^N = \lim_{N \rightarrow \infty} \left[e^{\frac{1}{N}a} \left(1 + \frac{1}{N}b \right) \right]^N.$$

PROOF. We prove the second formula — the mild changes for the other formula are straightforward. Let K_N be the matrix such that

$$e^{\frac{1}{N}a} \left(1 + \frac{1}{N}b\right) = 1 + \frac{1}{N}(a+b) + K_N. \quad (\text{A.12})$$

It is clear that $\|K_N\| = O(\frac{1}{N^2})$. We have

$$\left[e^{\frac{1}{N}a} \left(1 + \frac{1}{N}b\right) \right]^N = \left(1 + \frac{1}{N}(a+b)\right)^N + R_N, \quad (\text{A.13})$$

where R_N is a matrix whose norm satisfies

$$\|R_N\| \leq \sum_{k=0}^{N-1} \binom{N}{k} \left\|1 + \frac{1}{N}(a+b)\right\|^k \|K_N\|^{N-k} = O\left(\frac{1}{N}\right). \quad (\text{A.14})$$

The first term in the right side of (A.13) converges to e^{a+b} . \square

PROPOSITION A.6 (Duhamel formula). *Let a, b be $n \times n$ matrices. Then*

$$\begin{aligned} e^{a+b} &= e^a + \int_0^1 e^{ta} b e^{(1-t)(a+b)} dt \\ &= \sum_{k \geq 0} \int_{0 < t_1 < \dots < t_k < 1} dt_1 \dots dt_k e^{t_1 a} b e^{(t_2 - t_1)a} b \dots b e^{(1 - t_k)a}. \end{aligned}$$

PROOF. Let $F(s)$ be the matrix-valued function

$$F(s) = e^{sa} + \int_0^s e^{ta} b e^{(s-t)(a+b)} dt. \quad (\text{A.15})$$

We show that, for all s ,

$$e^{s(a+b)} = F(s). \quad (\text{A.16})$$

The derivative of $F(s)$ is

$$F'(s) = e^{sa} a + e^{sa} b + \int_0^s e^{ta} b e^{(s-t)(a+b)} (a+b) dt = F(s)(a+b). \quad (\text{A.17})$$

On the other hand, the derivative of $e^{s(a+b)}$ is $e^{s(a+b)}(a+b)$. The identity (A.16) clearly holds for $s = 0$ and, since both sides satisfy the same differential equation, they must be equal for all s .

We can iterate Duhamel's formula N times so as to get

$$\begin{aligned} e^{a+b} &= \sum_{k=0}^N \int_{0 < t_1 < \dots < t_k < 1} dt_1 \dots dt_k e^{t_1 a} b e^{(t_2 - t_1)a} b \dots b e^{(1 - t_k)a} \\ &\quad + \int_{0 < t_1 < \dots < t_N < 1} dt_1 \dots dt_N e^{t_1 a} b e^{(t_2 - t_1)a} b \dots b \left[e^{(1 - t_N)(a+b)} - e^{(1 - t_N)a} \right]. \end{aligned} \quad (\text{A.18})$$

Using $\|e^{ta}\| \leq e^{t\|a\|}$, the last line is less than $2e^{\|a\| + \|b\|} \frac{\|b\|^N}{N!}$ and so it vanishes in the limit $N \rightarrow \infty$. The summand is less than $e^{\|a\|} \frac{\|b\|^k}{k!}$, so that the sum is absolutely convergent. \square

A.4. Further matrix inequalities

PROPOSITION A.7 (Golden–Thompson inequality). *Let a, b be hermitian matrices. Then*

$$\mathrm{Tr} (e^{a+b}) \leq \mathrm{Tr} (e^a e^b).$$

PROOF. Hölder’s inequality, in the form (A.5) with $r = 1$, $p_j = n$ and $a_j = ab$, implies that $|\mathrm{Tr} (ab)^n| \leq \|ab\|_n^n$. The latter is equal to $\mathrm{Tr} (a^2 b^2)^{n/2}$ since a, b are hermitian. Letting n be a power of 2, we can iterate and we get

$$\mathrm{Tr} (ab)^n \leq \mathrm{Tr} a^n b^n. \quad (\text{A.19})$$

We use this inequality with $a \mapsto e^{\frac{1}{n}a}$ and $b \mapsto e^{\frac{1}{n}b}$, which gives

$$\mathrm{Tr} \left(e^{\frac{1}{n}a} e^{\frac{1}{n}b} \right)^n \leq \mathrm{Tr} e^a e^b. \quad (\text{A.20})$$

The left side converges to $\mathrm{Tr} e^{a+b}$ as $n \rightarrow \infty$ by the Trotter formula (Proposition A.5). \square

PROPOSITION A.8 (Klein inequality). *Let f be a convex differentiable function, and a, b be hermitian matrices with eigenvalues in the domain of f . Then*

$$\mathrm{Tr} [f(a) - f(b) - (a - b)f'(b)] \geq 0.$$

With $f(s) = e^s$, exchanging a and b , we get

$$\mathrm{Tr} (e^a - e^b) \leq \mathrm{Tr} (a - b) e^a. \quad (\text{A.21})$$

PROOF. Let (ϕ_i) and (ψ_i) be orthonormal bases of eigenvectors of a and b , and let (α_i) and (β_i) the eigenvalues. Let $c_{ij} = \langle \phi_i, \psi_j \rangle$. Then

$$\begin{aligned} & \langle \phi_i, [f(a) - f(b) - (a - b)f'(b)] \phi_i \rangle \\ &= f(\alpha_i) - \sum_j |c_{ij}|^2 f(\beta_j) - \sum_j |c_{ij}|^2 (\alpha_i - \beta_j) f'(\beta_j) \\ &= \sum_j |c_{ij}|^2 [f(\alpha_i) - f(\beta_j) - (\alpha_i - \beta_j) f'(\beta_j)] \\ &\geq 0. \end{aligned} \quad (\text{A.22})$$

\square

PROPOSITION A.9 (Peierls–Bogolubov inequality). *Let f be convex on \mathbb{R} and a, h be hermitian matrices such that $\mathrm{Tr} e^{-h} = 1$. Then*

$$f(\mathrm{Tr} a e^{-h}) \leq \mathrm{Tr} f(a) e^{-h}.$$

PROOF. Let (ϕ_i) and (η_i) be the eigenvectors and eigenvalues of h . Using Jensen's inequality twice,

$$\begin{aligned} f(\operatorname{Tr} a e^{-h}) &= f\left(\sum_i \langle \phi_i, a \phi_i \rangle e^{-\eta_i}\right) \leq \sum_i f(\langle \phi_i, a \phi_i \rangle) e^{-\eta_i} \\ &\leq \sum_i \langle \phi_i, f(a) \phi_i \rangle e^{-\eta_i} = \operatorname{Tr} f(a) e^{-h}. \end{aligned} \tag{A.23}$$

□

PROPOSITION A.10 (Peierls inequality). *Let a be a hermitian matrix and (ϕ_i) an orthonormal set of vectors. Then*

$$\sum_i e^{\langle \phi_i, a \phi_i \rangle} \leq \operatorname{Tr} e^a.$$

PROOF. Let α_j be the eigenvalues of a with corresponding orthonormal eigenvectors ψ_j . Then

$$e^{\langle \phi_i, a \phi_i \rangle} = \exp\left\{\sum_j \alpha_j |\langle \phi_i, \psi_j \rangle|^2\right\} \leq \sum_j |\langle \phi_i, \psi_j \rangle|^2 e^{\alpha_j}. \tag{A.24}$$

We used Jensen's inequality and

$$\sum_j |\langle \phi_i, \psi_j \rangle|^2 = \operatorname{Tr} |\phi_i\rangle\langle \phi_i| = 1. \tag{A.25}$$

The claim follows by summing over i , using $\sum_i |\langle \phi_i, \psi_j \rangle|^2 = \operatorname{Tr} |\psi_j\rangle\langle \psi_j| = 1$. □

A.5. A discrete Poincaré inequality

THEOREM A.11 (Discrete Poincaré Inequality on \mathbb{Z}^d). *Let*

$$\Omega = \{1, \dots, n\}^d \subset \mathbb{Z}^d,$$

and let $u : \Omega \rightarrow \mathbb{R}$ be a function with zero mean,

$$\sum_{x \in \Omega} u(x) = 0.$$

Define the discrete gradient by

$$(\nabla_i u)(x) = u(x + e_i) - u(x),$$

whenever x and $x + e_i$ both lie in Ω . Then there exists a constant C_d depending only on the dimension d such that

$$\sum_{x \in \Omega} u(x)^2 \leq C_d n^2 \sum_{x \in \Omega} \sum_{i=1}^d (u(x + e_i) - u(x))^2. \tag{A.26}$$

PROOF. Extend u to the discrete torus $\mathbb{T}_n^d = (\mathbb{Z}/n\mathbb{Z})^d$ by setting it to be zero outside Ω and then making it periodic. This does not increase the discrete Dirichlet energy and preserves the zero-mean property:

$$\widehat{u}(k) = \frac{1}{n^d} \sum_{x \in \mathbb{T}_n^d} u(x) e^{-2\pi i(k \cdot x)/n}, \quad k \in \mathbb{T}_n^d,$$

so that $\widehat{u}(0) = 0$.

Using Parseval's identity,

$$\sum_{x \in \mathbb{T}_n^d} |u(x)|^2 = n^d \sum_{k \in \mathbb{T}_n^d} |\widehat{u}(k)|^2.$$

The discrete gradient in Fourier variables satisfies

$$\widehat{\nabla_i u}(k) = \left(e^{2\pi i k_i/n} - 1 \right) \widehat{u}(k),$$

and hence

$$\sum_{x \in \mathbb{T}_n^d} |\nabla_i u(x)|^2 = n^d \sum_{k \neq 0} |e^{2\pi i k_i/n} - 1|^2 |\widehat{u}(k)|^2.$$

Note that

$$|e^{2\pi i k_i/n} - 1|^2 = 4 \sin^2\left(\frac{\pi k_i}{n}\right) \geq \frac{4}{\pi^2} \frac{\pi^2 k_i^2}{n^2} = \frac{4k_i^2}{n^2},$$

using $\sin t \geq \frac{2}{\pi}t$ for $t \in [0, \pi/2]$.

Summing in $i = 1, \dots, d$ gives the lower bound for the discrete Laplacian symbol:

$$\sum_{i=1}^d |e^{2\pi i k_i/n} - 1|^2 \geq \frac{4}{n^2} |k|^2, \quad k \neq 0.$$

Therefore,

$$\sum_{x \in \mathbb{T}_n^d} |\nabla u(x)|^2 \geq \frac{4}{n^2} n^d \sum_{k \neq 0} |k|^2 |\widehat{u}(k)|^2.$$

Since $|k|^2 \geq 1$ for $k \neq 0$, we obtain

$$n^d \sum_{k \neq 0} |\widehat{u}(k)|^2 \leq \frac{n^2}{4} \sum_{x \in \mathbb{T}_n^d} |\nabla u(x)|^2.$$

Using Parseval again gives

$$\sum_{x \in \mathbb{T}_n^d} u(x)^2 \leq C_d n^2 \sum_{x \in \mathbb{T}_n^d} |\nabla u(x)|^2,$$

with C_d depending only on the dimension. Since extension to the torus does not increase the right-hand side and the left-hand side restricts back to Ω , the same inequality holds on Ω itself. \square

COROLLARY A.12 (Another Poincaré inequality). *Let $X \Subset \mathbb{Z}^d$. Any function u with zero mean satisfies*

$$\|u\|_1 \leq \sqrt{C_d |X|} \operatorname{diam} X \|\nabla u\|_2.$$

PROOF. Use Cauchy-Schwarz to get $\|u\|_1 \leq \sqrt{|X|} \|u\|_2$. The result follows from Theorem A.11. (If X is not a box, it can be embedded in a box of size $\operatorname{diam} X$.) \square

A.6. About convex functions

Here is a simple result about convex functions on \mathbb{R}^k . It is useful to approximate an infinite-volume Gibbs state by a finite-volume one.

PROPOSITION A.13 (Exact matching of subgradients under a unique supporting hyperplane). *Let $f_n, f : \mathbb{R}^d \rightarrow (-\infty, +\infty]$ be proper, lower semicontinuous, convex and equi-coercive. Assume that $f_n \rightarrow f$ pointwise on a dense subset of \mathbb{R}^d (equivalently, $f_n \xrightarrow{\Gamma} f$).*

Let $x \in \mathbb{R}^d$ and suppose that there exists $p \in \partial f(x)$ such that

$$(*) \quad \operatorname{argmin}_{y \in \mathbb{R}^d} f(y) - \langle p, y \rangle = \{x\}.$$

Then there exists a sequence $x_n \rightarrow x$ such that $p \in \partial f_n(x_n)$ for all $n \in \mathbb{N}$. In particular, $p \in \partial f(x) \cap \partial f_n(x_n)$ for all n , $x_n \rightarrow x$.

PROOF. Fix $p \in \partial f(x)$ satisfying $(*)$ and define the tilted functionals

$$g_n(y) := f_n(y) - \langle p, y \rangle, \quad g(y) := f(y) - \langle p, y \rangle. \quad (\text{A.27})$$

Since subtraction of a continuous linear functional preserves Γ -convergence, we have

$$g_n \xrightarrow{\Gamma} g. \quad (\text{A.28})$$

Because $p \in \partial f(x)$, we have $0 \in \partial g(x)$, so x is a minimiser of g . By assumption $(*)$, this minimiser is in fact unique:

$$\operatorname{argmin} g = \{x\}. \quad (\text{A.29})$$

Equi-coercivity of (f_n) implies equi-coercivity of (g_n) , hence each g_n admits at least one minimiser. By the fundamental stability result for minimisers under Γ -convergence, there exist minimisers $x_n \in \operatorname{argmin} g_n$ such that $x_n \rightarrow x$.

For each n , the minimality of x_n yields

$$0 \in \partial g_n(x_n) = \partial f_n(x_n) - p, \quad (\text{A.30})$$

so $p \in \partial f_n(x_n)$ for all n .

Since $p \in \partial f(x)$ by construction, we obtain

$$p \in \partial f(x) \cap \partial f_n(x_n) \quad \text{for all } n, x_n \rightarrow x, \quad (\text{A.31})$$

which is the desired conclusion. \square

A.7. Infinite-volume limits

The following lemma is key to proving the existence of the infinite-volume limit of the free energy. Think of a_Λ as the free energy in domain Λ .

LEMMA A.14. *Assume that $(a_\Lambda)_{\Lambda \in \mathbb{Z}^d}$ is a set-indexed sequence of real numbers that satisfies*

- *translation invariance: $a_{\Lambda+x} = a_\Lambda$ for all $\Lambda \in \mathbb{Z}^d$ and all $x \in \mathbb{Z}^d$;*
- *bulk property: there exists a constant c such that for any k and any mutually disjoint $\Lambda_1, \dots, \Lambda_k \in \mathbb{Z}^d$, we have*

$$\left| a_{\cup_{i=1}^k \Lambda_i} - \sum_{i=1}^k a_{\Lambda_i} \right| \leq c \sum_{i=1}^k |\partial_1 \Lambda_i|.$$

Then there exists $\lambda \in \mathbb{R}$ such that

$$\lambda = \lim_{n \rightarrow \infty} \frac{a_{\Lambda_n}}{|\Lambda_n|}$$

along any van Hove sequence (Λ_n) of domains in \mathbb{Z}^d .

PROOF. Let $B_k = \{1, \dots, 2^k\}^d$ the box of size 2^k . We have for any $k < m$ that

$$\left| a_{B_m} - 2^{(m-k)d} a_{B_k} \right| \leq c 2^{(m-k)d} \cdot 2d \cdot 2^{k(d-1)}. \quad (\text{A.32})$$

Indeed, $2^{(m-k)d}$ is the number of translates of B_k whose union is B_m , and $2d \cdot 2^{k(d-1)}$ is the cardinality of $\partial_1 B_k$. Then

$$\left| \frac{a_{B_m}}{|B_m|} - \frac{a_{B_k}}{|B_k|} \right| \leq c 2d \cdot 2^{-k}. \quad (\text{A.33})$$

It follows that $(\frac{a_{B_k}}{|B_k|})$ is a Cauchy sequence; let λ denote its limit. There remains to verify that all van Hove sequences converge to the same limit.

For given k , consider the tiling of \mathbb{Z}^d with boxes of the form $B_k + z$, $z \in (2^k \mathbb{Z})^d$. Given $\Lambda \in \mathbb{Z}^d$, let $\mathfrak{z}(\Lambda) = \{z \in (2^k \mathbb{Z})^d : B_k + z \subset \Lambda\}$, and let

$$\Lambda' = \Lambda \setminus \cup_{z \in \mathfrak{z}(\Lambda)} (B_k + z). \quad (\text{A.34})$$

We have the properties

$$\begin{aligned} 2^{kd} |\mathfrak{z}(\Lambda)| + |\Lambda'| &= |\Lambda|; \\ |\Lambda'| &\leq 2^{kd} |\partial_1 \Lambda|; \\ \left| a_\Lambda - a_{\Lambda'} - |\mathfrak{z}(\Lambda)| a_{B_k} \right| &\leq c |\partial_1 \Lambda'| + c |\mathfrak{z}(\Lambda)| 2d \cdot 2^{k(d-1)}. \end{aligned} \quad (\text{A.35})$$

Consequently,

$$\left| \frac{a_\Lambda}{|\Lambda|} - \frac{a_{B_k}}{|B_k|} \right| \leq \frac{|a_{\Lambda'}|}{|\Lambda|} + \underbrace{\left| \frac{|\mathfrak{z}(\Lambda)| |B_k|}{|\Lambda|} - 1 \right|}_{=|\Lambda'|/|\Lambda|} \frac{|a_{B_k}|}{|B_k|} + c \frac{|\Lambda'|}{|\Lambda|} + c 2d \cdot 2^{-k} \underbrace{\frac{|\mathfrak{z}(\Lambda)| 2^{kd}}{|\Lambda|}}_{\leq 1}. \quad (\text{A.36})$$

We also have $|a_{\Lambda'}| \leq c' |\Lambda'|$ (one can take $c' = |a_{\{0\}}|(1 + 2dc)$).

We apply the above inequality to a van Hove sequence (Λ_n) . The right side is as small as we want by first taking k large enough and then n large enough. \square

A.8. Proof of the quantum Pinsker inequality

This section is devoted to the proof of Lemma 6.3 (quantum Pinsker's inequality). It relies on the convexity of the relative entropy, which can be proved using Lieb's concavity theorem, and on a clever path that allows to use the classical Pinsker inequality for Bernoulli variables.

LEMMA A.15. *Let $a, b, h \in \mathcal{M}_n$ such that $a, b \geq 0$, $[a, b] = 0$, and $h = h^*$. Assume that $\begin{pmatrix} a & h \\ h & b \end{pmatrix} \geq 0$. Then $h \leq a^{1/2}b^{1/2}$.*

PROOF. Let $u, v \in \mathbb{C}^n$ such that $\|u\| = \|v\| = 1$ and define $x = a^{-1/2}u$, $y = b^{-1/2}v$. Then, with $\mathbf{0}$ the zero vector in \mathbb{C}^n , we have

$$0 \leq \begin{pmatrix} x^* & \mathbf{0}^* \\ \mathbf{0}^* & y^* \end{pmatrix} \begin{pmatrix} a & h \\ h & b \end{pmatrix} \begin{pmatrix} x & \mathbf{0} \\ \mathbf{0} & y \end{pmatrix} = \begin{pmatrix} x^*ax & x^*hy \\ y^*hx & y^*hy \end{pmatrix} = \begin{pmatrix} 1 & u^*a^{-1/2}hb^{-1/2}v \\ v^*b^{-1/2}ha^{-1/2}u & 1 \end{pmatrix}. \quad (\text{A.37})$$

The latter is a 2×2 matrix; its determinant is nonnegative so that $|u^*a^{-1/2}hb^{-1/2}v| \leq 1$. This implies that $\|a^{-1/2}hb^{-1/2}\| \leq 1$. Next one can check that $a^{-1/4}b^{-1/4}ha^{-1/4}b^{-1/4}$ has the same eigenvalues as $a^{-1/2}hb^{-1/2}$ (if v is eigenvector of the first matrix, then $a^{-1/4}b^{1/4}v$ is eigenvector of the second matrix with the same eigenvalue). Then

$$\|a^{-1/4}b^{-1/4}ha^{-1/4}b^{-1/4}\| \leq 1, \quad (\text{A.38})$$

and, since this matrix is hermitian, we have

$$a^{-1/4}b^{-1/4}ha^{-1/4}b^{-1/4} \leq \mathbb{1} \iff h \leq a^{-1/2}b^{-1/2}. \quad (\text{A.39})$$

\square

THEOREM A.16 (Lieb's concavity). *Let $a_1, a_2, b_1, b_2 \geq 0$ be $n \times n$ complex matrices and let $\alpha \in [0, 1]$. Then*

$$(a_1 + a_2)^\alpha \otimes (b_1 + b_2)^{1-\alpha} \geq a_1^\alpha \otimes b_1^{1-\alpha} + a_2^\alpha \otimes b_2^{1-\alpha}.$$

PROOF. Let $x(\alpha) = a_1^\alpha \otimes b_1^{1-\alpha}$, $y(\alpha) = a_2^\alpha \otimes b_2^{1-\alpha}$, $z(\alpha) = (a_1 + a_2)^\alpha \otimes (b_1 + b_2)^{1-\alpha}$. We need to show that $z(\alpha) \geq x(\alpha) + y(\alpha)$ for all $\alpha \in [0, 1]$. It is actually enough to show this in a dense subset since all expressions are continuous in α . This clearly holds for $\alpha \in \{0, 1\}$. We now show that if it holds for α and β , then it also holds for $\frac{\alpha+\beta}{2}$.

We have $x(\frac{\alpha+\beta}{2}) = x(\alpha)^{1/2}x(\beta)^{1/2}$, and the same relations for y and z . Then

$$\begin{pmatrix} x(\alpha) & x(\frac{\alpha+\beta}{2}) \\ x(\frac{\alpha+\beta}{2}) & x(\beta) \end{pmatrix} = \begin{pmatrix} x(\alpha)^{1/2} \\ x(\beta)^{1/2} \end{pmatrix} \begin{pmatrix} x(\alpha)^{1/2} & x(\beta)^{1/2} \end{pmatrix} \geq 0. \quad (\text{A.40})$$

The latter inequality holds quite generally, only using that $x(\alpha)$ is hermitian. We have a similar inequality for y . Then

$$0 \leq \begin{pmatrix} x(\alpha) & x(\frac{\alpha+\beta}{2}) \\ x(\frac{\alpha+\beta}{2}) & x(\beta) \end{pmatrix} + \begin{pmatrix} y(\alpha) & y(\frac{\alpha+\beta}{2}) \\ y(\frac{\alpha+\beta}{2}) & y(\beta) \end{pmatrix} \leq \begin{pmatrix} z(\alpha) & x(\frac{\alpha+\beta}{2}) + y(\frac{\alpha+\beta}{2}) \\ x(\frac{\alpha+\beta}{2}) + y(\frac{\alpha+\beta}{2}) & z(\beta) \end{pmatrix}. \quad (\text{A.41})$$

The second holds because the difference is equal to $\begin{pmatrix} z(\alpha) - x(\alpha) - y(\alpha) & 0 \\ 0 & z(\beta) - x(\beta) - y(\beta) \end{pmatrix}$, which is nonnegative by assumption. We now use Lemma A.15 and we get

$$x(\frac{\alpha+\beta}{2}) + y(\frac{\alpha+\beta}{2}) \leq z(\alpha)^{1/2} z(\beta)^{1/2} = z(\frac{\alpha+\beta}{2}). \quad (\text{A.42})$$

We can start with $\alpha = 0$ and $\beta = 1$ and iterate the inequality, so it applies to all multiples of 2^{-k} for arbitrary k ; this set is dense in $[0, 1]$. \square

COROLLARY A.17. *Let $a_1, a_2, b_1, b_2 \geq 0$ be $n \times n$ complex matrices and let $\alpha \in [0, 1]$. Then*

$$\text{Tr}((a_1 + a_2)^\alpha (b_1 + b_2)^{1-\alpha}) \geq \text{Tr}(a_1^\alpha b_1^{1-\alpha}) + \text{Tr}(a_2^\alpha b_2^{1-\alpha}).$$

PROOF. We use the following correspondence between \mathcal{M}_n and $\mathbb{C}^n \otimes \mathbb{C}^n$:

$$\text{Tr} a^T b = \sum_{i,j=1}^n a_{i,j} b_{i,j} = \sum_{i,j=1}^n \langle i | \otimes \langle i | (a \otimes b) | j \rangle \otimes | j \rangle. \quad (\text{A.43})$$

This allows to use Theorem A.16. \square

We now introduce the relative entropy. It is usually applied to density matrices but it helps to define it more generally on all positive-definite matrices. One can check that it is always nonnegative; we allow the value $+\infty$.

DEFINITION A.18. *The **relative entropy** $S(\cdot||\cdot)$ is the following function of two positive-definite matrices $a, b \in \mathcal{M}_n$:*

$$S(a||b) = \text{Tr} a (\log a - \log b).$$

LEMMA A.19. *We have for all $a, b \geq 0$ that*

$$S(a||b) = \lim_{\varepsilon \rightarrow 0^+} \frac{1}{\varepsilon} (\text{Tr} a - \text{Tr} a^{1-\varepsilon} b^\varepsilon).$$

PROOF. Let $f(\varepsilon) = \text{Tr} a^{1-\varepsilon} b^\varepsilon$. The right side is the derivative $-f'(0)$ which is equal to $S(a||b)$. \square

THEOREM A.20 (Joint convexity of the relative entropy). *If $a_1, a_2, b_1, b_2 \geq 0$ are complex matrices in \mathcal{M}_n , then*

$$S(a_1 + a_2 || b_1 + b_2) \leq S(a_1 || b_1) + S(a_2 || b_2).$$

Since $S(\lambda a || \lambda b) = \lambda S(a || b)$, the joint convexity of the relative entropy follows immediately. And since the entropy is equal to $S(a) = S(a || \mathbb{1})$, it is convex too.

PROOF. Starting with Lemma A.19, and using Corollary A.17, we have

$$\begin{aligned} S(a_1 + a_2 \| b_1 + b_2) &= \lim_{\varepsilon \rightarrow 0^+} \frac{1}{\varepsilon} \left(\text{Tr}(a_1 + a_2) - \text{Tr}(a_1 + a_2)^{1-\varepsilon} (b_1 + b_2)^\varepsilon \right) \\ &\leq \lim_{\varepsilon \rightarrow 0^+} \frac{1}{\varepsilon} \left(\text{Tr} a_1 - \text{Tr} a_1^{1-\varepsilon} b_1^\varepsilon + \text{Tr} a_2 - \text{Tr} a_2^{1-\varepsilon} b_2^\varepsilon \right) \\ &= S(a_1 \| b_1) + S(a_2 \| b_2). \end{aligned} \tag{A.44}$$

□

LEMMA A.21. Let $U = \text{diag}(1, e^{\frac{2\pi i}{n}}, \dots, e^{\frac{2\pi i}{n}(n-1)})$. Then for any matrix $a \in \mathcal{M}_n$ we have

$$\text{diag } a = \frac{1}{n} \sum_{k=0}^{n-1} U^k a U^{-k}.$$

PROOF. The element (ℓ, m) of the left side is equal to

$$\frac{1}{n} \sum_{k=0}^{n-1} e^{\frac{2\pi i}{n} k \ell} a_{\ell, m} e^{-\frac{2\pi i}{n} k m} = \frac{1}{n} a_{\ell, m} \sum_{k=0}^{n-1} e^{\frac{2\pi i}{n} k(\ell-m)} = \begin{cases} a_{\ell, \ell} & \text{if } \ell = m, \\ 0 & \text{if } \ell \neq m. \end{cases} \tag{A.45}$$

□

Let $P = \text{diag}(1, \dots, 1, 0, \dots, 0)$ with k elements equal to 1, where $k \in \{1, \dots, n-1\}$. We consider the map $\Phi : \mathcal{M}_n \rightarrow \mathcal{M}_n$ defined by

$$\Phi(a) = \text{diag} \left(\underbrace{\frac{1}{k} \text{Tr } P a \dots \frac{1}{k} \text{Tr } P a}_{k \text{ elements}}, \underbrace{\frac{1}{n-k} \text{Tr}(1-P)a \dots \frac{1}{n-1} \text{Tr}(1-P)a}_{n-k \text{ elements}} \right). \tag{A.46}$$

The image $\Phi(a)$ is a simple matrix with just two values.

LEMMA A.22. There exists a finite number L and unitary matrices U_1, \dots, U_L such that

$$\Phi(a) = \frac{1}{L} \sum_{\ell=1}^L U_\ell a U_\ell^{-1}.$$

PROOF. If a is diagonal we can consider the permutation $(1, \dots, k)(k+1, \dots, n)$ and its permutation matrix V . Then $\Phi(a) = \frac{1}{k(n-k)} \sum_{\ell=1}^{k(n-k)} V^\ell a V^{-\ell}$. Indeed, this amounts to average over diagonal matrices where the first k elements of a have been rotated, as well as the last $n-k$ elements.

For the general case we combine this with Lemma A.21 to get

$$\Phi(a) = \frac{1}{nk(n-k)} \sum_{\ell=1}^{k(n-k)} \sum_{m=0}^{n-1} V^\ell U^m a U^{-m} V^{-\ell}. \tag{A.47}$$

□

We can now prove the quantum Pinsker inequality.

PROOF OF LEMMA 6.3. Let P be the projector onto the subspace of the eigenvectors of $\rho - \sigma$ with nonnegative eigenvalues and let Φ be the map defined in Eq. (A.46). By Lemma A.22 and the convexity of the relative entropy (Theorem A.20), we obtain that

$$S(\rho\|\sigma) \geq S(\Phi(\rho)\|\Phi(\sigma)). \quad (\text{A.48})$$

The latter is equal to the classical relative entropy of two Bernoulli random variables with parameters $\text{Tr } P\rho$ and $\text{Tr } P\sigma$. Using the classical Pinsker inequality (see Exercise A.1), we get that it is greater than

$$\frac{1}{2}(|\text{Tr } P\rho - \text{Tr } P\sigma| + |\text{Tr } (1 - P)\rho - \text{Tr } (1 - P)\sigma|)^2 = \frac{1}{2}\|\rho - \sigma\|_1^2. \quad (\text{A.49})$$

The last identity uses the fact that P is the projector onto the suitable eigensubspace of $\rho - \sigma$. \square

We consider one more proposition, that gives the partial trace as a convex combination of unitary operations. It can be combined with Theorem A.20 to give inequalities involving the (relative) entropy in different domains.

PROPOSITION A.23. *Let Tr_1 denote the partial trace with respect to the first space of $\mathcal{H}_1 \otimes \mathcal{H}_2$; on operators of the form $a \otimes b$, we have that $\text{Tr}_1 a \otimes b = (\text{Tr } a)b$ is an operator on \mathcal{H}_2 . Then there exist unitary matrices U_1, \dots, U_L on $\mathcal{H}_1 \otimes \mathcal{H}_2$ such that*

$$\text{Tr}_1 a = \frac{1}{L} \sum_{\ell=1}^L U_\ell a U_\ell^{-1}.$$

This can be proved as in Lemma A.22 with unitary matrices on \mathcal{H}_1 tensored with the identity.

EXERCISE A.1. *Prove the classical Pinsker inequality in the case of Bernoulli random variables.*

BIBLIOGRAPHICAL REFERENCES

For the proof of the quantum Pinsker's inequality we mainly rely on Watrous [2018] and on Carlen [2010].