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

$$f(D) = \operatorname{diag}(f(d_1), \ldots, f(d_n)). \tag{A.1}$$

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 $a = U\operatorname{diag}(\lambda_1, \ldots, \lambda_n)U^*$, then we define

$$f(a) = U\operatorname{diag}(f(\lambda_1), \ldots, f(\lambda_n))U^*. \tag{A.2}$$

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

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

We also define the *operator norm* of a by

$$\|a\| = \sup_{\psi \neq 0} \frac{\|a\psi\|}{\|\psi\|} \tag{A.4}$$

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

$$\Big\|\prod_{j=1}^{n} a_j\Big\|_r \leq \prod_{j=1}^{n} \|a_j\|_{p_j} \tag{A.5}$$

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*

$$\big|\mathrm{Tr}\, a_1 \dots a_{2n}\big| \leq \prod_{i=1}^{2n} \Big(\mathrm{Tr}\,(a_i a_i^*)^n\Big)^{1/2n}.$$

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

$$\big|\mathrm{Tr}\, a_1 \dots a_{2n}\big|^2 \leq \mathrm{Tr}\,\big(a_1 \dots a_n a_n^* \dots a_1^*\big)\ \mathrm{Tr}\,\big(a_{2n}^* \dots a_{n+1}^* a_{n+1} \dots a_{2n}\big). \tag{A.6}$$

This allows to use a reflection positivity argument. By replacing a_i with $a_i / \sqrt{\mathrm{Tr}\,(a_i a_i^*)^n}$ it is enough to prove the inequality for matrices that satisfy $\mathrm{Tr}\,(a_i a_i^*)^n = 1$; 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 $|\mathrm{Tr}\, a_1 \dots a_{2n}|$, 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 $\mathrm{Tr}\,(aa^*)^n$ for some matrix $a \in \{a_1, \dots, a_n\}$, which is equal to 1. □

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

Corollary A.3. *We have*

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

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

$$b_i = |a_i|^{p_i/2\ell}, \qquad \hat{b}_i = U_i|a_i|^{p_i/2\ell}. \tag{A.7}$$

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

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

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

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

□

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*

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

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

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

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 $\operatorname{Tr} c^* a = \|a\|_r$. □

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

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

□

A.3. Trotter and Duhamel

We now review a ueful 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*

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

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

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

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

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

where R_N is a matrix whose norm satisfies

$$\|R_N\| \leq \sum_{k=0}^{N-1} \binom{N}{k} \|1 + \tfrac{1}{N}(a+b)\|^k \, \|K_N\|^{N-k} = O(\tfrac{1}{N}). \tag{A.14}$$

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

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

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

We show that, for all s,

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

The derivative of $F(s)$ is

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

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

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

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

A.4. Further matrix inequalities

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

$$\operatorname{Tr}\big(e^{a+b}\big) \leq \operatorname{Tr}\big(e^{a}\, e^{b}\big).$$

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

$$\operatorname{Tr}(ab)^n \leq \operatorname{Tr} a^n b^n. \tag{A.19}$$

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

$$\operatorname{Tr}\Big(e^{\frac{1}{n}a}\, e^{\frac{1}{n}b}\Big)^n \leq \operatorname{Tr} e^{a}\, e^{b}. \tag{A.20}$$

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

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*

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

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

$$\operatorname{Tr}\big(e^{a} - e^{b}\big) \leq \operatorname{Tr}(a-b)\, e^{a}. \tag{A.21}$$

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

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

□

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

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

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

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

□

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

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

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

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

We used Jensen's inequality and

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

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

A.5. About convex functions

Here is a simple result about convex functions on $\mathbb{R}$. Recall that the right and left derivatives of a convex function $f(s)$ at $s = 0$ are given respectively by $\partial_+ f(0) = \inf_{s>0} \frac{f(s)-f(0)}{s}$ and $\partial_- f(0) = \sup_{s<0} \frac{f(s)-f(0)}{s}$.

PROPOSITION A.11. *Let f_n be a sequence of continuously differentiable, convex functions on $\mathbb{R}$ such that $f_n \to f$ pointwise. For each $m \in [\partial_- f(0), \partial_+ f(0)]$ there is a sequence $s_n \to 0$ such that $m = \lim_{n\to\infty} f_n'(s_n)$.*

PROOF. We claim the following: for any $\varepsilon, \delta > 0$ there is $N = N(\varepsilon, \delta)$ such that for any $n > N$ we have

$$f_n'(\delta) > \partial_+ f(0) - \varepsilon, \qquad f_n'(-\delta) < \partial_- f(0) + \varepsilon. \tag{A.26}$$

The result then follows using the mean value theorem for the continuous function $f_n'(s)$: if it is the case that $m \in [\partial_- f(0) + \varepsilon, \partial_+ f(0) - \varepsilon]$ then there is some $s_n \in [-\delta, \delta]$ satisfying $f_n'(s_n) = m$, otherwise we may take $s_n = \delta$ or $s_n = -\delta$.

We prove the claim for $\partial_+ f(0)$. We have

$$f_n'(\delta) \geq \frac{f_n(\delta) - f_n(\delta/2)}{\delta/2}, \qquad \frac{f(\delta/2) - f(0)}{\delta/2} \geq \partial_+ f(0). \tag{A.27}$$

Moreover,

$$\lim_{n\to\infty} \frac{f_n(\delta) - f_n(\delta/2)}{\delta/2} - \frac{f(\delta/2) - f(0)}{\delta/2} = \frac{f(\delta) - f(\delta/2)}{\delta/2} - \frac{f(\delta/2) - f(0)}{\delta/2} \geq 0. \tag{A.28}$$

So for n large enough we have

$$f_n'(\delta) \geq \frac{f_n(\delta) - f_n(\delta/2)}{\delta/2} \geq \frac{f(\delta/2) - f(0)}{\delta/2} - \varepsilon \geq \partial_+ f(0) - \varepsilon, \tag{A.29}$$

as claimed. □