ENTROPY AND THE UNCERTAINTY PRINCIPLE

RUPERT L. FRANK AND ELLIOTT H. LIEB

ABSTRACT. We generalize, improve and unify theorems of Rumin, and Maassen–Uffink about classical entropies associated to quantum density matrices. These theorems refer to the classical entropies of the diagonals of a density matrix in two different bases. Thus they provide a kind of uncertainty principle. Our inequalities are sharp because they are exact in the high-temperature or semi-classical limit.

1. INTRODUCTION

The von Neumann entropy of a quantum state (density matrix) can be calculated either in momentum space or in configuration space and the two are equal. They can even be zero. Nevertheless, the corresponding classical entropies, determined by the diagonals of the two representations of the density matrix, can be different, and they can even be negative, but their sum cannot be arbitrarily small. This sum of the classical entropies can thus serve as a measure of the quantum mechanical uncertainty principle.

This point of view was advocated by Deutsch [De], who, among other things, proved a lower bound on this sum, which was later improved by Maassen and Uffink [MaUf], following a conjecture of Kraus [Kr]. These inequalities were obtained for a general pair of bases, not just momentum and configuration space. In the momentum–configuration basis an improvement on these previous inequalities was made by Rumin [Ru], who was able to add a term to the inequality involving the largest eigenvalue of the density matrix. He raised the question whether this additional term could be further improved by using a larger quantity, namely, the von Neumann entropy of the density matrix. In this paper we prove that this surmise is correct.

We prove even more by combining the Maaseen-Uffink investigation with the Rumin surmise. Rumin was concerned with the momentum–configuration space duality, whereas Maassen-Uffink were concerned with arbitrary pairs of bases of the Hilbert space. For this they introduced a parameter c which somehow quantifies the disparity between the two bases. As one might expect, the k, x pair has the largest c-value, i.e., $c = 1$. We show how our theorem applies to any pair with the corresponding c-dependent improvement found in [MaUf].

Work partially supported by NSF grants PHY–1068285 (R.L.F.) and PHY–0965859 (E.H.L.).

Our theorem and simple proof are supported by a semi-classical intuition, as evidenced by our use of the Golden-Thompson inequality. The only other ingredient in our proof is the Gibbs variational principle. Because our constant in Theorem 2.1 agrees with the semi-classical limit it is the best possible.

2. Rumin's conjecture and its generalizations

For any trace class operator $\gamma \geq 0$ on $L^2(\mathbb{R}^d)$ we denote by $\rho_\gamma(x) = \gamma(x,x)$ its density; see (2.2) for a precise definition. Moreover,

$$\hat{\gamma}(k,k') = \iint_{\mathbb{R}^d\times\mathbb{R}^d} e^{2\pi i(k\cdot x - k'\cdot x')}\gamma(x,x')\,dx\,dx'$$

and

$$\rho_{\widehat{\gamma}}(k) = \hat{\gamma}(k,k) = \iint_{\mathbb{R}^d\times\mathbb{R}^d} e^{2\pi i k\cdot(x-x')}\gamma(x,x')\,dx\,dx'\,.$$

We note that if $\operatorname{Tr}\gamma = 1$, then

$$\int_{\mathbb{R}^d} \rho_\gamma(x)\,dx = \int_{\mathbb{R}^d} \rho_{\widehat{\gamma}}(k)\,dk = 1\,.$$

Our main result is

Theorem 2.1. *For any* $\gamma \geq 0$ *with* $\operatorname{Tr}\gamma = 1$ *and*

$$\int_{\mathbb{R}^d} \rho_\gamma(x)\ln_+\rho_\gamma(x)\,dx < \infty \quad \text{and} \quad \int_{\mathbb{R}^d} \rho_{\widehat{\gamma}}(k)\ln_+\rho_{\widehat{\gamma}}(k)\,dk < \infty\,,$$

where $\ln_+\rho = \max\{\ln\rho, 0\}$*, one has*

$$-\int_{\mathbb{R}^d} \rho_\gamma(x)\ln\rho_\gamma(x)\,dx - \int_{\mathbb{R}^d} \rho_{\widehat{\gamma}}(k)\ln\rho_{\widehat{\gamma}}(k)\,dk \geq -\operatorname{Tr}\gamma\ln\gamma\,. \tag{2.1}$$

Remarks. (1) While the entropy on the right side of (2.1) is necessarily non-negative, those on the left side can have either sign.
(2) Inequality (2.1) is saturated in the semi-classical limit. This can be verified by taking $\gamma = Z_\beta^{-1}\exp(-\beta(-\Delta + x^2))$ and letting $\beta \to 0$; see [Ru].
(3) For γ of rank one, this is Hirschman's inequality [Hi]. This was improved by Beckner [Be]. However, because of (1) this improvement is not possible if one allows for mixed states (i.e., γ of higher rank).
(4) The inequality for γ equal to a multiple of a projection was proved in [Ru]. More generally, Rumin proves (2.1) with $\ln\|\gamma\|_\infty$ instead of $\operatorname{Tr}\gamma\ln\gamma$.
(5) The inequality shares the following tensorization property: If $d = n + m$, we can think of $L^2(\mathbb{R}^d)$ as $L^2(\mathbb{R}^n)\otimes L^2(\mathbb{R}^m)$. Then the main inequality for $\gamma = \gamma_n\otimes\gamma_m$ equals the sum of the inequalities for γ_n and γ_m.
(6) If, instead, we define the Fourier transform by

$$\widetilde{\gamma}(p,q) = \iint e^{i(p\cdot x - q\cdot y)}\gamma(x,y)\,\frac{dx\,dy}{(2\pi)^d} \quad \text{and} \quad \rho_{\widetilde{\gamma}}(p) = \iint e^{ip\cdot(x-y)}\gamma(x,y)\frac{dx\,dy}{(2\pi)^d}\,,$$

then (2.1) becomes

$$-\int \rho_\gamma(x)\ln\rho_\gamma(x)\,dx - \int \rho_{\widehat{\gamma}}(p)\ln\rho_{\widehat{\gamma}}(p)\,dp \geq -\operatorname{Tr}\gamma\ln\gamma + d\ln(2\pi)\,.$$

Theorem 2.1 is a special case of a more general Theorem 2.2 below. We listed Theorem 2.1 separately because it was the starting point of our investigation and was conjectured by Rumin.

The more general theorem includes the discrete case as well as the continuous case in Theorem 2.1. It is not entirely a triviality that the discrete and continuous cases are contained in one theorem because, as is well known, many entropy inequalities are true in one case and not in the other. For example, the discrete entropy is always positive while the continuous entropy can be, and often is, negative.

The general set-up consists of two sigma-finite measure spaces (X,μ) and (Y,ν). We denote by $L^2(X)$ and $L^2(Y)$ the corresponding spaces of square-integrable functions. Let γ be a non-negative operator on $L^2(X)$ with $\operatorname{Tr}\gamma = 1$. Then we have $\gamma = \sum_j \lambda_j |f_j\rangle\langle f_j|$ with orthonormal functions (f_j) and numbers $\lambda_j \in [0,1]$ satisfying $\sum_j \lambda_j = 1$. We define the density ρ_γ of γ, a function on X, by

$$\rho_\gamma(x) = \sum_j \lambda_j |f_j(x)|^2\,. \tag{2.2}$$

By monotone convergence, we have

$$\int_X \rho_\gamma(x)\,d\mu(x) = \sum_j \lambda_j = \operatorname{Tr}\gamma = 1\,. \tag{2.3}$$

Assume now that there is a unitary operator $\mathcal{U} : L^2(X) \to L^2(Y)$. For γ as before, we define an operator $\hat{\gamma}$ on $L^2(Y)$ by

$$\hat{\gamma} = \mathcal{U}\,\gamma\,\mathcal{U}^*\,.$$

This operator is non-negative and has $\operatorname{Tr}\hat{\gamma} = 1$. Its density $\rho_{\widehat{\gamma}}$ is defined similarly to that of ρ_γ, namely,

$$\rho_{\widehat{\gamma}}(y) = \sum_j \lambda_j |g_j(y)|^2\,,$$

where $\hat{\gamma} = \sum_j \lambda_j |g_j\rangle\langle g_j|$ and $g_j = \mathcal{U} f_j$. As in (2.3),

$$\int_Y \rho_{\widehat{\gamma}}(y)\,d\nu(y) = 1\,. \tag{2.4}$$

Our final assumption is that $\mathcal{U}$ is bounded from $L^1(X)$ to $L^\infty(Y)$. This property guarantees that $\mathcal{U}$ has an integral kernel $\mathcal{U}(y,x)$ with

$$\infty > \|\mathcal{U}\|_{L^1\to L^\infty} = \text{ess-sup}_{x,y}|\mathcal{U}(y,x)| := \sup\{t : (\mu\times\nu)(\{(x,y) : |\mathcal{U}(x,y)| > t\}) > 0\}\,.$$

Theorem 2.2. *Under the above assumptions, let $\gamma \geq 0$ be an operator in $L^2(X)$ with $\operatorname{Tr}\gamma = 1$ and such that*

$$\int_X \rho_\gamma(x) \ln_+ \rho_\gamma(x)\, d\mu(x) < \infty \quad \text{and} \quad \int_Y \rho_{\widehat{\gamma}}(y) \ln_+ \rho_{\widehat{\gamma}}(y)\, d\nu(y) < \infty\,,$$

where $\ln_+ \rho = \max\{\ln \rho, 0\}$. Then

$$- \int_X \rho_\gamma(x) \ln \rho_\gamma(x)\, d\mu(x) - \int_Y \rho_{\widehat{\gamma}}(y) \ln \rho_{\widehat{\gamma}}(y)\, d\nu(y) \geq - \operatorname{Tr}\gamma \ln \gamma - 2 \ln \|\mathcal{U}\|_{L^1 \to L^\infty}\,. \tag{2.5}$$

We illustrate this theorem by some **examples**.

(1) If $X = Y = \mathbb{R}^d$ with Lebesgue measure and $\mathcal{U}$ the Fourier transform (i.e., $\mathcal{U}(k,x) = e^{-2\pi i k\cdot x}$), then we recover Theorem 2.1. In this case, $-2\ln\|\mathcal{U}\|_{L^1\to L^\infty} = 0$.

(2) Let $X = (-L/2, L/2)$ with Lebesgue measure, $Y = L^{-1}\mathbb{Z}$ with L^{-1} times counting measure and let $\mathcal{U}$ be the discrete Fourier transform, that is, $\mathcal{U}(k,x) = e^{-2\pi i k x}$. Then (2.5) holds with $-2\ln\|\mathcal{U}\|_{L^1\to L^\infty} = 0$.

(3) Let $X = Y = \mathbb{Z}/N\mathbb{Z} = \{0, 1, \ldots, N-1\}$ for some $N \in \mathbb{N}$ with counting measure and let $\mathcal{U}(k,n) = N^{-1/2} e^{-i2\pi k n/N}$. Then (2.5) holds with $-2\ln\|\mathcal{U}\|_{L^1\to L^\infty} = \ln N$.

(4) The following is a generalization of Example (3) and is related to [De, Kr, MaUf]. Let $(|a_j\rangle)_j$ and $(|b_k\rangle)_k$ two orthonormal bases in a separable Hilbert space $\mathcal{H}$ and put

$$c = \sup_{j,k} |\langle a_j | b_k \rangle|\,.$$

By the Schwarz inequality, $0 < c \leq 1$. Let $\gamma \geq 0$ be an operator on $\mathcal{H}$ with $\operatorname{Tr}\gamma = 1$. Define

$$p_j := \langle a_j|\gamma|a_j\rangle\,, \qquad q_k := \langle b_k|\gamma|b_k\rangle\,.$$

Then

$$- \sum_j p_j \ln p_j - \sum_k q_k \ln q_k \geq - \operatorname{Tr}\gamma\ln\gamma - 2\ln c\,, \tag{2.6}$$

which follows from Theorem 2.2 by noting that, if the change of bases is denoted by $\mathcal{U}$, then $\|\mathcal{U}\|_{L^1\to L^\infty} = c$. The weaker inequality without the term $\operatorname{Tr}\gamma\ln\gamma$ on the right side was shown in [MaUf] with a different proof.

In passing, we note that each of the entropies on the left side of (2.6) is greater than or equal to $-\operatorname{Tr}\gamma\ln\gamma$. This follows from the concavity of $-p\ln p$, the fact (derived from the variational principle) that the sequence (p_j) is majorized by the sequence of eigenvalues of γ, and Karamata's theorem (see, e.g., [HaLiPo] or [LiSe, Rem. 4.7]).

3. Proof of Theorem 2.2

Our proof is based on the following two well known lemmas in quantum statistical mechanics; see, e.g., [Ca, Si].

Lemma 3.1 (Gibbs variational principle). *Let H be a self-adjoint operator such that e^{-H} is trace class. Then for any $\gamma \geq 0$ with $\operatorname{Tr}\gamma = 1$,*

$$\operatorname{Tr}\gamma H + \operatorname{Tr}\gamma \ln \gamma \geq -\ln \operatorname{Tr} e^{-H}$$

with equality iff $\gamma = \exp(-H)/\operatorname{Tr}\exp(-H)$.

Lemma 3.2 (Golden-Thompson inequality). *For self-adjoint operators A and B,*

$$\operatorname{Tr} e^{A+B} \leq \operatorname{Tr} e^{A/2} e^{B} e^{A/2} .$$

Proof of Theorem 2.2. We first note that

$$-\int \rho_\gamma(x) \ln \rho_\gamma(x)\, d\mu(x) - \int \rho_{\widehat{\gamma}}(y) \ln \rho_{\widehat{\gamma}}(y)\, d\nu(y) = \operatorname{Tr}\gamma H$$

with the operator $H = -\ln \rho_\gamma - \mathcal{U}^* \ln \rho_{\widehat{\gamma}} \mathcal{U}$ in $L^2(X)$. Here, $\ln \rho_\gamma$ and $\ln \rho_{\widehat{\gamma}}$ are considered as multiplication operators, and we used the fact that $\operatorname{Tr}_{L^2(X)} \mathcal{U}^* A \mathcal{U} = \operatorname{Tr}_{L^2(Y)} A$. By Lemmas 3.1 and 3.2,

$$\begin{aligned} -\int \rho_\gamma(x) \ln \rho_\gamma(x)\, d\mu(x) - \int \rho_{\widehat{\gamma}}(y) \ln \rho_{\widehat{\gamma}}(y)\, d\nu(y) + \operatorname{Tr}\gamma \ln \gamma &\geq -\ln \operatorname{Tr} e^{-H} \\ &\geq -\ln \operatorname{Tr} \rho_\gamma^{1/2} \mathcal{U}^* \rho_{\widehat{\gamma}} \mathcal{U} \rho_\gamma^{1/2} . \end{aligned}$$

The trace on the right side is the square of the Hilbert-Schmidt norm of the operator $\rho_{\widehat{\gamma}}^{1/2} \mathcal{U} \rho_\gamma^{1/2}$, which has kernel

$$\rho_{\widehat{\gamma}}(y)^{1/2}\, \mathcal{U}(y,x) \rho_\gamma(x)^{1/2} .$$

Thus,

$$\begin{aligned} \operatorname{Tr} \rho_\gamma^{1/2} \mathcal{U}^* \rho_{\widehat{\gamma}} \mathcal{U} \rho_\gamma^{1/2} &= \iint_{X\times Y} \rho_{\widehat{\gamma}}(y) |\mathcal{U}(y,x)|^2 \rho_\gamma(x)\, d\mu(x)\, d\nu(y) \\ &\leq \|\mathcal{U}\|^2_{L^1\to L^\infty} \int_Y \rho_{\widehat{\gamma}}(y)\, d\nu(y) \int_X \rho_\gamma(x)\, d\mu(x) . \end{aligned}$$

By (2.3) and (2.4), this equals $\|\mathcal{U}\|^2_{L^1\to L^\infty}$, and the proof of the theorem is complete. □

References

[Be] W. Beckner, *Inequalities in Fourier analysis.* Ann. of Math. (2) **102** (1975), no. 1, 159–182.

[Ca] E. A. Carlen, *Trace inequalities and quantum entropy. An introductory course.* In: Entropy and the quantum, R. Sims and D. Ueltschi (eds.), 73–140, Contemp. Math. **529**, Amer. Math. Soc., Providence, RI, 2010.

[De] D. Deutsch, *Uncertainty in quantum measurements.* Phys. Rev. Lett. **50** (1983), 631–633.

[HaLiPo] G. H. Hardy, J. E. Littlewood, G. Pólya, *Inequalities.* Cambridge University Press, Cambridge, 1952.

[Hi] I. I. Hirschman Jr., *A note on entropy.* Amer. J. Math. **79** (1957), 152–156.

[Kr] K. Kraus, *Complementary observables and uncertainty relations.* Phys. Rev. D **35** (1987), 3070–3075.

[LiSe] E. H. Lieb, R. Seiringer, *The stability of matter in quantum mechanics.* Cambridge University Press, Cambridge, 2010.

[MaUf] H. Maassen, J. B. M. Uffink, *Generalized entropic uncertainty relations.* Phys. Rev. Lett. **60** (1988), no. 12, p. 1103–1106.

[Ru] M. Rumin, *Balanced distribution-energy inequalities and related entropy bounds.* Preprint (2010), arXiv:1008.1674

[Si] B. Simon, *Trace ideals and their applications.* Second edition. Amer. Math. Soc., Providence, RI, 2005.

Rupert L. Frank, Department of Mathematics, Princeton University, Washington Road, Princeton, NJ 08544, USA

E-mail address: rlfrank@math.princeton.edu

Elliott H. Lieb, Departments of Mathematics and Physics, Princeton University, P. O. Box 708, Princeton, NJ 08544, USA

E-mail address: lieb@princeton.edu