PROOF OF A WEHRL-TYPE ENTROPY INEQUALITY FOR THE AFFINE AX + B GROUP

ELLIOTT H. LIEB AND JAN PHILIP SOLOVEJ

Abstract. After reviewing the affine group \( AX + B \), its representation on \( L^2(\mathbb{R}_+) \), and its coherent states, we define the Wehrl type conjecture for \( L^p \)-norms of these coherent states (also known as the Renyi entropies). We give the proof of this conjecture in the case that \( p \) is an even integer and we show how the general case reduces to an unsolved problem about analytic functions on the upper half plane.

Contents

1. Introduction
2. The generalized conjecture and our main results
3. An analytic formulation
References

1. Introduction

The starting point for this work, historically, was Wehrl’s definition of semiclassical entropy and his conjecture about its minimum value [12]. Given a density matrix \( \rho \) on \( L^2(\mathbb{R}) \) (a positive operator whose trace is 1) we define its classical probability density (or Husimi function) as follows:

\[
\rho^{cl}(p, q) := \langle p, q | \rho | p, q \rangle.
\] (1.1)

We use the usual Dirac notation in which \( | \cdots \rangle \) is a vector and \( \langle \cdots | \cdots \rangle \) is the inner product conjugate linear in the first variable and linear in the second. Here \( | p, q \rangle \) is the (Schrödinger, Klauder, Glauber) coherent state, which is a normalized function in \( L^2(\mathbb{R}) \) parametrized by \( p, q \in \mathbb{R}^2 \), which is the classical phase space for a particle in one-dimension. It is given (with \( \hbar = 1 \)) by

\[
| p, q \rangle(x) = \pi^{-\frac{1}{4}} \exp[-(x - q)^2 + ipx].
\]

The von Neumann entropy of any quantum state with density matrix \( \rho \) is \( S^Q(\rho) := -\text{Tr} \rho \log \rho \) and the classical entropy of any continuous probability density \( \tilde{\rho}(p, q) \) is \( S^{cl}(\tilde{\rho}) := -\int \tilde{\rho} \log \tilde{\rho} \, dp \, dq \). The \( S^Q \) entropy is non-negative, while \( S^{cl}(\tilde{\rho}) \geq 0 \) for any \( \tilde{\rho} \) that is point-wise \( \leq 1 \), as is the case for \( \rho^{cl} \).

Date: June 1, 2019.

This work was supported by the Villum Centre of Excellence for the Mathematics of Quantum Theory (QMATH) and the ERC Advanced grant 321029.
For the reader’s convenience we recall some basic facts of, and the interest in, the Wehrl entropy. A classical probability does not always lead to a positive entropy and it often leads to an entropy that is $-\infty$. The quantum von Neumann entropy is always non-negative. Wehrl’s contribution was to derive a classical probability distribution from a quantum state that has several very desirable features. One is that it is always non-negative. Another is that it is monotone, meaning that the Wehrl entropy of a quantum state of a tensor product is always greater than or equal to the entropy of each subsystem obtained by taking partial traces. This monotonicity holds for marginals of classical probabilities, but not always for the quantum von Neumann entropy (the entropy of the universe may be smaller than the entropy of a peanut). In short the Wehrl entropy combines some of the desirable properties of the classical and quantum entropies.

The minimum possible von Neumann entropy is zero and occurs when $\rho$ is any pure state, while that of the classical $S^{cl}(\rho^{cl})$ is strictly positive; Wehrl’s conjecture was that the minimum is 1 and occurs when

$$\rho = |p, q\rangle\langle p, q|.$$  

That is, $\rho$ is a pure state projector onto any coherent state. This conjecture was proved by one of us [8]. The uniqueness of this choice of $\rho$ was shown by Carlen [2]. Recently De Palma [4] determined the minimal Wehrl entropy when the von Neumann entropy of $\rho$ is fixed to some positive value.

The Wehrl conjecture was generalized in [8] to the theorem that $||\rho^{cl}||_s$, $s > 1$ was maximized for the same choice of $\rho$. This is equivalent to saying that the classical Renyi entropy of $\rho^{cl}$ is minimized for this same choice of $\rho$. Later, in our joint paper [9], the same upper bound result was shown to hold for the integral of any convex function of $\rho^{cl}$, not just $x \rightarrow x^s$.

The possibility of using coherent states to associate a classical density to a quantum state $\rho$ has a group-theoretic interpretation, which we will explain below. In the case of the classical coherent states discussed above this relates to the Heisenberg group and the coherent states are heighest weight vectors in the unitary irreducible representation. This suggests that we can look at other groups, their representations, and associated coherent states (heighest weight vectors) and ask whether the analog of the Wehrl conjecture holds.

The first question in this direction was raised in [8] for the group $SU(2)$ where the representations are labeled by the quantum spin $J = 0, 1/2, 1, 3/2, \ldots$. A few spin cases were proved by Schupp in [11] and in full generality by us in [9].

The obvious next case is $SU(N)$ for general $N$ which has many kinds of representations. We showed the Wehrl hypothesis for all the symmetric representations in [10]. Here we turn our attention to another group of some physical but also mathematical interest which is the affine group. This group, like the Heisenberg group, is not compact and it is not even unimodular which means that the left and right invariant Haar measures are different. The purely mathematical interest is in signal processing. The coherent states for the affine group are like continuous families of wavelets.

The affine group is $G = \mathbb{R}_+ \times \mathbb{R}$ with the composition rule

$$(a, b) \cdot (a', b') = (aa', ab' + b).$$
This comes from thinking of the group acting on the real line as
\[ x \mapsto ax + b. \]

The group is, therefore, sometimes referred to as the \( ax + b \) group. It is not unimodular and the left Haar measure on the group is given by
\[ a^{-2} \, da \, db, \]
(for reference, the right Haar measure is \( a^{-1} \, da \, db \)). The group has two irreducible faithful unitary representations \([1, 6]\) which may be realized on \( L^2(\mathbb{R}_+, dk) \) either by
\[ (U(a, b)f)(k) = \exp(-2\pi ibk)a^{1/2}f(ak) \]
or with \( e^{-2\pi ibk} \) replaced by \( e^{+2\pi ibk} \). For the representation above the coherent states we will consider are given in terms of fiducial vectors
\[ \eta_\alpha(k) = C(\alpha)k^{\alpha} \exp(-k) \]
where the parameter \( \alpha \) is positive and which we will henceforth keep fixed. These functions are identified as affine coherent states (extremal weight vectors) for the representation of the affine group in \([3, 7]\). The constant \( C(\alpha) \) is chosen such that
\[ \int_0^\infty |\eta_\alpha(k)|^2/k \, dk = 1, \]
i.e.,
\[ C(\alpha) = 2^\alpha \Gamma(2\alpha)^{-1/2}. \]

Then
\[ \int_{-\infty}^\infty \int_0^\infty U(a, b)|\eta_\alpha\rangle\langle \eta_\alpha|U(a, b)^*a^{-2} \, da \, db = I. \] (1.2)

For any function \( f \in L^2(\mathbb{R}_+) \) we may introduce the coherent state transform
\[ h_f(a, b) = \langle U(a, b)\eta_\alpha | f \rangle = C(\alpha)a^{\alpha+1/2} \int_0^\infty \exp(-ka + 2\pi ibk)k^{\alpha} f(k) \, dk. \] (1.3)

As in the case of the classical coherent states, we then have
\[ \int_{-\infty}^\infty \int_0^\infty |h_f(a, b)|^2 a^{-2} \, da \, db = \int_0^\infty |f(k)|^2 \, dk, \]
and if \( f \) is normalized in \( L^2 \) we may consider \( |h_f(a, b)|^2 \) as a probability density relative to the Haar measure \( a^{-2} \, da \, db \). Observe that in contrast to the classical case we do not have \( |h_f(a, b)|^2 \leq 1 \), but rather \( |h_f(a, b)|^2 \leq \int |\eta_\alpha(k)|^2 dk = \alpha \). This is a consequence of the fact that (1.2) requires \( \int_0^\infty |\eta_\alpha(k)|^2/k \, dk = 1 \), which is different from the \( L^2 \)-normalization. This difference is due to the group not being uni-modular. The Wehrl entropy
\[ -\int |h_f(a, b)|^2 \ln(|h_f(a, b)|^2) a^{-2} \, da \, db \]
is, therefore, not necessarily non-negative, but it is, however, still bounded from below by \(-\ln \alpha\). The natural generalization of Wehrl’s conjecture in this case is that the entropy is minimal for \( f = \alpha^{-1/2} \eta_\alpha \).

In the following section we state a more general conjecture and prove it in special cases. In the last section of the paper we show that our conjecture and theorem are equivalent to \( L^p \) estimates for analytic functions in the complex upper half plane.
2. THE GENERALIZED CONJECTURE AND OUR MAIN RESULTS

We make the following more general conjecture.

**Conjecture 1** (Wehrl-type conjecture for the affine group). Let \( s \geq 1 \) be given. Then among all normalized \( f \in L^2(\mathbb{R}_+) \) the integral

\[
\int_{-\infty}^{\infty} \int_{0}^{\infty} |h_f(a,b)|^{2s} a^{-2} \, da \, db
\]

is maximal if and only if (up to a phase) \( f \) has the form \( f = \alpha^{-1/2} U(a,b) \eta_\alpha \) for some \( a > 0 \) and \( b \in \mathbb{R} \). The maximal value is

\[
\frac{2\alpha^s}{(2\alpha + 1)s - 1}.
\]

The analog of Wehrl’s original entropy conjecture follows from this conjecture by taking a derivative at \( s = 1 \) as in [8] and gives the minimal entropy

\[
1 + (2\alpha)^{-1} - \ln \alpha,
\]

which agrees with the lower bound \( -\ln \alpha \) mentioned above.

The following theorem is the main result of this paper.

**Theorem 2.** The statement of Conjecture 1 holds for the special cases of \( s \) being a positive integer.

**Proof.** We want to prove that

\[
\int_{-\infty}^{\infty} \int_{0}^{\infty} |h_f(a,b)|^{2s} a^{-2} \, da \, db
\]

is maximized if and only if

\[
f(k) = Ak^\alpha \exp(-Bk)
\]

with \( A, B \in \mathbb{C} \) such that \( f \in L^2(\mathbb{R}_+) \) is normalized. The proof will rely only on a Schwarz inequality and existence and uniqueness will follow from the corresponding uniqueness of optimizers for Schwarz inequalities.

If \( s \) is a positive integer we can write

\[
h_f(a,b)^s = C(\alpha)^s a^{s(\alpha + 1/2)} \int_{0}^{\infty} \cdots \int_{0}^{\infty} \exp \left[ (-a + 2\pi ib)(k_1 + \cdots + k_s) \right] \times (k_1 \cdots k_s)^\alpha f(k_1) \cdots f(k_s) \, dk_1 \cdots dk_s.
\]

Hence doing the \( b \) integration gives

\[
\int_{-\infty}^{\infty} \int_{0}^{\infty} |h_f(a,b)|^{2s} a^{-2} \, da \, db = C(\alpha)^{2s} \int_{0}^{\infty} \cdots \int_{0}^{\infty} a^{s(2\alpha + 1) - 2} \exp \left[ -2a(k_1 + \cdots + k_s) \right] \times k_1^\alpha f(k_1) \cdots k_s^\alpha f(k_s) \frac{k_{s+1}^\alpha f(k_{s+1})}{k_{s+1}^\alpha} \cdots \frac{k_{2s}^\alpha f(k_{2s})}{k_{2s}^\alpha} \, da \, dk_1 \cdots dk_{2s-1}
\]
A WEHRL-TYPE ENTROPY INEQUALITY FOR THE AFFINE $AX + B$ GROUP

where $k_{2s} = k_1 + \cdots + k_s - (k_{s+1} + \cdots + k_{2s-1})$. We now do the $a$ integration and arrive at

$$
\int_{-\infty}^{\infty} \int_{0}^{\infty} |h_f(a, b)|^{2s} a^{-2} dadb = C(\alpha)^{2s} \Gamma(s(2\alpha + 1) - 1) \int_{0}^{\infty} \cdots \int_{0}^{\infty} [2(k_1 + \cdots + k_s)]^{-s(2\alpha+1)+1} \times k_1^\alpha f(k_1) \cdots k_s^\alpha f(k_s) \frac{k_{s+1}^\alpha f(k_{s+1})}{k_{s+1}} \cdots \frac{k_{2s}^\alpha f(k_{2s})}{k_{2s}} \, dk_1 \cdots dk_{2s-1}.
$$

(2.1)

We now change variables to

$$
r = k_1 + \cdots + k_s
$$

$$
u_j = k_j/r, \quad j = 1, \ldots, s - 1
$$

$$
v_j = k_{s+j}/r, \quad j = 1, \ldots, s - 1.
$$

(2.2)

The Jacobian determinant for this change of variables is easily found to be

$$
\det \left[ \frac{\partial(k_1, \ldots, k_{2s-1})}{\partial(r, u_1, \ldots, u_{2s-1}, v_1, \ldots, v_{2s-1})} \right] = r^{2s-2}.
$$

We arrive at

$$
\int_{-\infty}^{\infty} \int_{0}^{\infty} |h_f(a, b)|^{2s} a^{-2} dadb = C(\alpha)^{2s} \Gamma(s(2\alpha + 1) - 1) \int_{0}^{\infty} \cdots \int_{0}^{\infty} 2^{-s(2\alpha+1)+1} r^{s-1} \times [u_1 \cdots u_{s-1}(1 - u_1 - \cdots - u_{s-1})]^{\alpha} \times [v_1 \cdots v_{s-1}(1 - v_1 - \cdots - v_{s-1})]^{\alpha} \times f(u_1 r) \cdots f(u_{s-1} r) f(r(1 - u_1 - \cdots - u_{s-1})) \times f(v_1 r) \cdots f(v_{s-1} r) f(r(1 - v_1 - \cdots - v_{s-1})) \times du_1 \cdots du_{s-1} dv_1 \cdots dv_{s-1} dr.
$$

(2.3)
Let us apply the Cauchy-Schwarz inequality for each fixed \( r \) to conclude that

\[
\int_{-\infty}^{\infty} \int_{0}^{\infty} |h_f(a,b)|^{2s} a^{-2} \, da \, db \\
\leq 2^{-s(2\alpha+1)+1} C(\alpha)^{2s} \Gamma(s(2\alpha+1) - 1) \\
\times \left[ \int \cdots \int [u_1 \cdots u_{s-1}(1 - u_1 - \cdots - u_{s-1})]^{2\alpha} \, du_1 \cdots du_{s-1} \right] \\
\times \int_{0}^{\infty} \left[ \int \cdots \int |f(u_1r) \cdots f(u_{s-1}r)f(r(1 - u_1 - \cdots - u_{s-1}))|^2 \\
\times \, du_1 \cdots du_{s-1} \right] r^{s-1} \, dr \\
= 2^{-s(2\alpha+1)+1} C(\alpha)^{2s} \Gamma(s(2\alpha+1) - 1) \left( \int_{0}^{\infty} |f(r)|^2 \, dr \right)^s \\
\times \left[ \int \cdots \int [u_1 \cdots u_{s-1}(1 - u_1 - \cdots - u_{s-1})]^{2\alpha} \, du_1 \cdots du_{s-1} \right].
\]

If \( f \) is normalized this is a number depending on \( \alpha \) and \( s \). The important observation is that the upper bound is achieved if and only if there is a function \( K(r) \) depending on \( r \) such that

\[
f(u_1) \cdots f(u_{s-1})f(r - u_1 - \cdots - u_{s-1}) = K(r)[u_1 \cdots u_{s-1}(r - u_1 - \cdots - u_{s-1})]^{\alpha}
\]

for almost all \( 0 \leq u_1, \ldots, u_{s-1}, r \) satisfying \( u_1 + \cdots + u_{s-1} \leq r \). If we define \( g(u) = u^{-\alpha} f(u) \) and introduce the variable \( u_s = r - (u_1 + \cdots + u_{s-1}) \) we may rewrite this as

\[
g(u_1) \cdots g(u_s) = K(1 + \cdots + u_s)
\]

for almost all \( 0 \leq u_1, \ldots, u_s \). It is not difficult to show that if a locally integrable function \( g \) satisfies this it must be smooth and it follows easily that

\[
g(u) = A \exp(-Bu)
\]

for complex numbers \( A, B \) which is exactly what we wanted to prove.

The maximal value can be found by a straightforward computation. \( \square \)

3. An analytic formulation

Using the Bergman-Paley-Wiener Theorem in [5] we can rephrase our conjecture and theorem in terms of analytic functions on the complex upper half plane \( \mathbb{C}_+ = \{ z \in \mathbb{C} \mid \text{Im}z > 0 \} \). Introducing the Bergman space,

\[
\mathcal{A}^2(\mathbb{C}_+) = \{ F \in L^2(\mathbb{C}_+) \mid f \text{ analytic } \}
\]

The Paley-Wiener Theorem in this context says that there is a unitary map

\[
L^2(\mathbb{R}_+) \ni f \mapsto F \in \mathcal{A}^2(\mathbb{C}_+),
\]
given by
\[ F(z) = \frac{1}{\sqrt{\pi}} \int_0^\infty e^{ikz} k^{1/2} f(k) dk. \]
If we write \( z = 2\pi b + ia \) we recognize the coherent state transform in (1.3) for \( \alpha = 1/2 \) to be
\[ h_f(a, b) = \sqrt{2\pi} \text{Im} F(z). \]
For simplicity we will restrict attention to \( \alpha = 1/2 \) in this section. The formulation of the analytic transform for all values of \( \alpha \) can be found in [3]. The action of the affine group on \( \mathbb{R} \) may be extended to the upper half plane \( \mathbb{C}_+ \) by \( z \mapsto az + b \). The representation of the affine group on the Bergman space is then
\[ (U(a,b)^* F)(z) = \frac{1}{4\pi} \int_0^\infty e^{ikz} k^{1/2} (U(a,b)^* f)(k) dk = F(az + b) \]
In the analytic representation our conjecture states that
\[ \int_{\mathbb{C}_+} |F(z)|^{2s} |\text{Im} z|^{2s-2} d^2z \leq \frac{\pi^{1-s}}{2s-1} \left( \int_{\mathbb{C}_+} |F(z)|^2 d^2z \right)^s \quad (3.1) \]
for analytic functions of the complex upper half-plane and \( s \geq 1 \) with equality if and only if \( F \) is proportional to \((z - z_0)^{-2}\) for some \( z_0 \) in the lower half-plane.

Our theorem from the last section is therefore equivalent to the following result.

**Theorem 3.** Inequality (3.1) holds for all analytic functions \( F : \mathbb{C}_+ \to \mathbb{C} \) and all positive integers \( s \). Equality holds if and only if \( F \) is proportional to \((z - z_0)^{-2}\) for some \( z_0 \) in the lower half-plane.

The open problem is to remove the restriction that \( s \) is an integer.

**References**

[1] Aslaksen, Erik W.; Klauder, John R. Unitary representations of the affine group. J. Mathematical Phys. 9 1968 206–211.
[2] Carlen, E.A., Some integral identities and inequalities for entire functions and their application to the coherent state transform. J. Funct. Anal., 97 (1991), 231–249.
[3] Daubechies, Ingrid, Klauder, John R. and Paul, Thierry, Wiener measures for path integrals with affine kinematic variables, J. Math. Phys., 28, (1987), 85–102,
[4] De Palma, Giacomo, The Wehrl entropy has Gaussian optimizers. Lett. Math. Phys. 108 (2018), no. 1, 97–116.
[5] Duren, Peter, Gallardo-Gutiérrez, Eva A., Montes-Rodríguez, Alfonso, A Paley-Wiener theorem for Bergman spaces with application to invariant subspaces, Bull. Lond. Math. Soc., 39, 2007, 459–466
[6] Gelfand, I.; Neumark, M. Unitary representations of the group of linear transformations of the straight line. C. R. (Doklady) Acad. Sci URSS (N.S.) 55, (1947). 567–570.
[7] Watson, Glenn and Klauder, John R., Generalized affine coherent states: a natural framework for the quantization of metric-like variables, J. Math. Phys., 41, (2000), 8072–8082.
[8] Lieb, Elliott H. Proof of an entropy conjecture of Wehrl. Comm. Math. Phys. 62 (1978), no. 1, 35–41.
[9] Lieb, Elliott H.; Solovej, Jan Philip, Proof of an entropy conjecture for Bloch coherent spin states and its generalizations. Acta Math. 212 (2014), no. 2, 379–398.
[10] Lieb, Elliott H.; Solovej, Jan Philip, Proof of the Wehrl-type entropy conjecture for symmetric SU(N) coherent states. Comm. Math. Phys. 348 (2016), no. 2, 567–578.

[11] Schupp, Peter On Lieb’s Conjecture for the Wehrl Entropy of Bloch Coherent States. Comm. Math. Phys. 207 (1999), no. 2, 481–493.

[12] Wehrl, Alfred General properties of entropy. Rev. Modern Phys. 50 (1978), no. 2, 221–260.

DEPARTMENTS OF MATHEMATICS AND PHYSICS, JADWIN HALL, PRINCETON UNIVERSITY, PRINCETON, NJ 08544, USA
E-mail address: lieb@princeton.edu

QMATH, DEPARTMENT OF MATHEMATICAL SCIENCES, UNIVERSITY OF COPENHAGEN, UNIVERSITETSPARKEN 5, DK-2100 COPENHAGEN Ø, DENMARK
E-mail address: solovej@math.ku.dk