A discrete entropic uncertainty relation

H. Maassen*
University of Nijmegen,the Netherlands

Recently [MaU] a new class of 'generalised entropic' uncertainty relations for the probability distributions of non-commuting random variables was proved as a simple consequence of the Riesz-Thorin interpolation theorem. Here we shall give a quite explicit proof of the central inequality of this class, an 'entropic' uncertainty relation, which has been conjectured by Kraus [Kra].

We consider the following situation, not uncommon in quantum mechanics. Two observables of a physical system are represented by symmetric complex $n \times n$ matrices A and B, which we shall assume to have non-degenerate spectra. We can write A and B in the form

$$A = \sum_{i=1}^{n} \alpha_i P_i \quad \text{and} \quad B = \sum_{i=1}^{n} \beta_i Q_i,$$

where $P_1, \cdots, P_n$ and $Q_1, \cdots, Q_n$ are sequences of mutually orthogonal one-dimensional projections, and the sequences $\alpha_1, \cdots, \alpha_n$ and $\beta_1, \cdots, \beta_n$ consist of distinct real numbers, to be interpreted as the values which the observables can take. Each state ω on the algebra M_n of all complex $n \times n$ matrices then induces probability distributions on the spectra of A and B: $\omega(P_i)$ (or $\omega(Q_i)$) is the probability to find the value α_i (or β_i) when measuring the observable A (or B). One now defines the *uncertainty* $H(A, \omega)$ *of* A *in the state* ω as the Shannon entropy of this probability distribution:

$$H(A, \omega) = -\sum_{i=1}^{n} \omega(P_i) \log \omega(P_i).$$

The question was raised ([BBM], [Deu], [Kra]), what can be said about $H(A, \omega) + H(B, \omega)$, more in particular about its lower bound

$$d(A, B) = \inf_{\omega} \left(H(A, \omega) + H(B, \omega) \right).$$

*Supported by the Netherlands Organisation of Scientific Research NWO.

One may regard this infimum as a "degree of incompatibility" of the observables A and B.

As a first reduction, let us note that $H(A, \omega)$ does not depend on the real numbers $\alpha_1, \cdots, \alpha_n$, but only on the projections $P_1, \cdots, P_n$, which are the minimal projections in the maximal abelian von Neumann algebra $\mathcal{A}$ generated by A. Let us therefore write $H(\mathcal{A}, \omega), H(\mathcal{B}, \omega)$ and $d(\mathcal{A}, \mathcal{B})$ in what follows. When viewed in this way, d becomes a natural distance function between maximal abelian von Neumann algebras, comparable to the distance of point *sets* (not of points) in geometry: $d(\mathcal{A}, \mathcal{B}) = 0$ if and only if $\mathcal{A}$ and $\mathcal{B}$ have a minimal projection in common.

The latter observation suggests to consider the following easily computable functional on pairs of abelian von Neumann algebras in M_n:

$$m(\mathcal{A}, \mathcal{B}) = \max\{\mathrm{tr}PQ | P \in \mathcal{A}, Q \in \mathcal{B} \text{ minimal projections}\}.$$

This definition amounts to

$$m(\mathcal{A}, \mathcal{B}) = \max_{1 \leq i,j \leq n} \mathrm{tr}(P_i Q_j) = \max_{1 \leq i,j \leq n} |{<}e_i, f_j{>}|^2,$$

where e_i and f_j are unit vectors in the ranges of P_i and Q_j respectively. Note that

$$m(\mathcal{A}, \mathcal{B}) \leq 1,$$

with equality if and only if $d(\mathcal{A}, \mathcal{B}) = 0$. On the other hand, since for all j

$$\sum_{i=1}^{n} \mathrm{tr}(P_i Q_j) = \mathrm{tr} Q_j = 1,$$

we have

$$m(\mathcal{A}, \mathcal{B}) \geq \frac{1}{n}.$$

It was observed by Kraus [Kra] that this lower bound is reached for 'complementary' observables A and B, which corresponds to e_j and f_j of the form

$$(e_j)_k = 1 \text{ if } j = k, \text{ 0 otherwise};$$

$$(f_j)_k = \frac{1}{\sqrt{n}} e^{\frac{2\pi i j k}{n}}.$$

Note that the algebras $\mathcal{A}$ and $\mathcal{B}$ then take the form:

$$\mathcal{A} = \{X \in M_n | X \text{ diagonal}\},$$

$$\mathcal{B} = \{X \in M_n | X_{i+k,j+k} = X_{i,j} \text{ for all } i,j,k\},$$

where the addition of indices is taken modulo n.

Kraus went on to conjecture that for such complementary observables (or algebras in our terminology)

$$d(\mathcal{A},\mathcal{B}) = \log n.$$

Indeed, the inequality $d(\mathcal{A},\mathcal{B}) \le \log n$ is easily established: choose $\omega(X) = <e_1, Xe_1>$, so that $(\omega(P_1), \cdots, \omega(P_n)) = (1, 0, \cdots, 0)$ and $(\omega(Q_1), \cdots, \omega(Q_n)) = (\frac{1}{n}, \frac{1}{n}, \cdots, \frac{1}{n})$; then $H(\mathcal{A}, \omega) = 0$ and

$$H(\mathcal{B}, \omega) = -\sum_{i=1}^{n} \frac{1}{n} \log \frac{1}{n} = \log n.$$

Kraus' conjecture is therefore a consequence of the following theorem.

Theorem 1 *For all maximal abelian von Neumann subalgebras $\mathcal{A}$ and $\mathcal{B}$ of M_n one has*

$$d(\mathcal{A},\mathcal{B}) \ge -\log m(\mathcal{A},\mathcal{B}).$$

Proof. Let $\{P_1, \cdots, P_n\}$ and $\{Q_1, \cdots, Q_n\}$ be complete sets of minimal projections in $\mathcal{A}$ and $\mathcal{B}$ respectively, and let e_j and f_j be unit vectors in the ranges of P_j and Q_j respectively ($j = 1, \cdots, n$). We may assume that $\{e_j\}$ is the canonical basis of $\mathbf{C}^n$. From the concavity of the function $\eta : [0,1] \to [0,\infty) : x \mapsto -x \log x$ (with $\eta(0) := 0$) it follows that the minimum of $H(\mathcal{A},\omega) + H(\mathcal{B},\omega) = \sum_{j=1}^{n} (\eta(\omega(P_j)) + \eta(\omega(Q_j)))$ is taken in a vector state $\omega(X) = <\psi, X\psi>$ on M_n. It therefore suffices to prove that for all $\psi \in \mathbf{C}^n$:

$$\sum_{j=1}^{n} (\eta(|<e_j,\psi>|^2) + \eta(|<f_j,\psi>|^2)) \ge -\log \max_{i,j} |<e_i, f_j>|^2. \qquad (1)$$

Now let $m = m(\mathcal{A},\mathcal{B}) = \max_{i,j} |<e_i, f_j>|^2$ and let a unitary map $T : \mathbf{C}^n \to \mathbf{C}^n$ be defined by $Tf_j = e_j$. Then $\psi_i = <e_i, \psi>$ and $(T\psi)_i = <e_i, T\psi> = <T^{-1}e_i, \psi> = <f_i, \psi>$. If we now write $h(\psi)$ for $\sum_{j=1}^{n} \eta(|\psi_j|^2)$, then the inequality takes the form

$$h(\psi) + h(T\psi) \ge -\log m. \qquad (2)$$

For $n \in \mathbf{N}$ and $p \in [1, \infty]$ let $l^p(n)$ denote the Banach space $\mathbf{C}^n$with norm

$$\|\psi\|_p = \begin{cases} (\sum_{i=1}^n |\psi_i|^p)^{\frac{1}{p}} & \text{if } 1 \leq p < \infty \\ \max_{1 \leq i \leq n} |\psi_i| & \text{if } p = \infty. \end{cases}$$

We now make the following observation.

Lemma 2 *For all unit vectors ψ in $\mathbf{C}^n$*

$$\frac{d}{dp^{-1}} \|\psi\|_p|_{p=2} = h(\psi).$$

Proof. First we note that

$$\frac{d}{dp} \|\psi\|_p^p|_{p=2} = \frac{d}{dp} \sum_{j=1}^n |\psi_j|^p|_{p=2} =$$

$$\sum_{j=1}^n |\psi_j|^2 \log |\psi_j| = -\tfrac{1}{2} \sum_{j=1}^n \eta(|\psi_j|^2) = -\tfrac{1}{2} h(\psi).$$

Therefore, since $\|\psi\|_2 = 1$,

$$\frac{d}{dp^{-1}} \|\psi\|_p|_{p=2} = \frac{d}{dp^{-1}} (\|\psi\|_p^p)^{p^{-1}}|_{p=2} =$$

$$\log \|\psi\|_2^2 + \frac{1}{p} (\|\psi\|_p^p)^{\frac{1}{p}-1} \cdot \frac{dp}{dp^{-1}} \cdot \frac{d}{dp} \|\psi\|_p^p|_{p=2}$$
$$= \tfrac{1}{2} \cdot (-4)(-\tfrac{1}{2} h(\psi)) = h(\psi).$$

□

Let $\|T\|_p$ denote the norm of the linear map $T : \mathbf{C}^n \to \mathbf{C}^n$, viewed as an operator $l^p(n) \to l^q(n)$, where $\frac{1}{p} + \frac{1}{q} = 1$. (Here we make the usual convention that $\frac{1}{\infty} = 0$.) Then $\|T\|_2 = 1$ and $\|T\|_1 = \max_{j,k} |T_{jk}| = \max_{j,k} |\langle f_j, e_k \rangle| = \sqrt{m}$.

Theorem 3 (Riesz-Thorin interpolation) *For a linear map $T : \mathbf{C}^n \to \mathbf{C}^n$ the function*

$$f_T : [0, 1] \to \mathbf{R} : \frac{1}{p} \mapsto \log \|T\|_p$$

is convex.

Proof: [Rie]; see also [HLP]. □

It follows that f_T has a right derivative $f_T'(\frac{1}{2})$ at $\frac{1}{2}$, and that, since $f_T(\frac{1}{2}) = \log \|T\|_2 = 0$ and $f_T(1) = \log \|T\|_1 = \frac{1}{2}\log m$, we have

$$f_T'(\tfrac{1}{2}) \leq \frac{f_T(1) - f_T(\frac{1}{2})}{1 - \frac{1}{2}} = \log m.$$

On the other hand, by the definition of the operator norm $\|T\|_p$, we have for all $p \in [1, \infty]$ and all unit vectors $\psi \in \mathbf{C}^n$:

$$\log \|T\|_p \geq \log \|T\psi\|_q - \log \|\psi\|_p,$$

where $\frac{1}{p} + \frac{1}{q} = 1$. Equality holds here for $p = 2$, hence we may differentiate with respect to $\frac{1}{p}$ at $\frac{1}{p} = \frac{1}{2}$:

$$f_T'(\tfrac{1}{2}) \geq -h(T\psi) - h(\psi).$$

It follows that $h(T\psi) + h(\psi) \geq -\log m$. □

The equality (2) is optimal if $|T_{ij}| = 1$ for some pair (i, j) and in the case of complementary observables, when

$$T_{jk} = \frac{1}{\sqrt{n}} e^{\frac{2\pi ijk}{n}}.$$

In general however, f_T will be strictly convex, so that $f_T'(\frac{1}{2}) < \log m$ and no ψ exists reaching equality in (2).

References

[BBM] I. Białynicki-Birula, J. Mycielsky: "Uncertainty Relations for Information Entropy in Wave Mechanics"; Commun. math. Phys. *44* (1975) 129-132.

[Deu] D. Deutsch: "Uncertainty in quantum measurements" Phys. Rev. Lett. *50* (1983) 631-633.

[HLP] G. Hardy, J.E. Littlewood, G. Pólya: Inequalities. Cambridge University Press 1934.

[Kra] K. Kraus: "Complementary observables and uncertainty relations"; Phys. Rev. D *35* (1987) 3070-3075.

[MaU] H. Maassen, J. Uffink: "Generalized Entropic Uncertainty Relations"; Phys. Rev. Lett. *60* (1988) 1103-1106.

[Rie] M. Riesz: "Sur les maxima des formes bilinéaires et sur les fonctionnelles linéaires"; Acta Math. *49* (1927) 465-497.