

The Rényi entropy and the uncertainty relations in quantum mechanics

Iwo Białynicki-Birula
Center for Theoretical Physics
Warsaw, Poland

www.cft.edu.pl/~birula

November 2008

Heisenberg uncertainty relation

The standard formulation of the uncertainty relations for the measurements of the position and momentum has the following form

$$\Delta x \Delta p \geq \frac{\hbar}{2}$$

Δx and Δp are the mean standard deviations

$$\Delta a = \sqrt{\langle (a - \langle a \rangle)^2 \rangle}$$

$$(\Delta x)^2 = \int dx \rho(x) (x - \langle x \rangle)^2 \quad (\Delta p)^2 = \int dp \tilde{\rho}(p) (p - \langle p \rangle)^2$$

This form of the uncertainty is simple to prove

The proof requires only the Schwartz inequality and is usually included in the first course on quantum mechanics

Critique of the standard uncertainty relations

The mean standard deviation Δa (dispersion) is used a lot in the statistical analysis of experiments that measure a

It is a reasonable measure of the spread when (as in measurements) the distribution in question is of a simple "hump" type

For example, it is a very good characteristic for a Gaussian distribution since it determines the half-width of this distribution

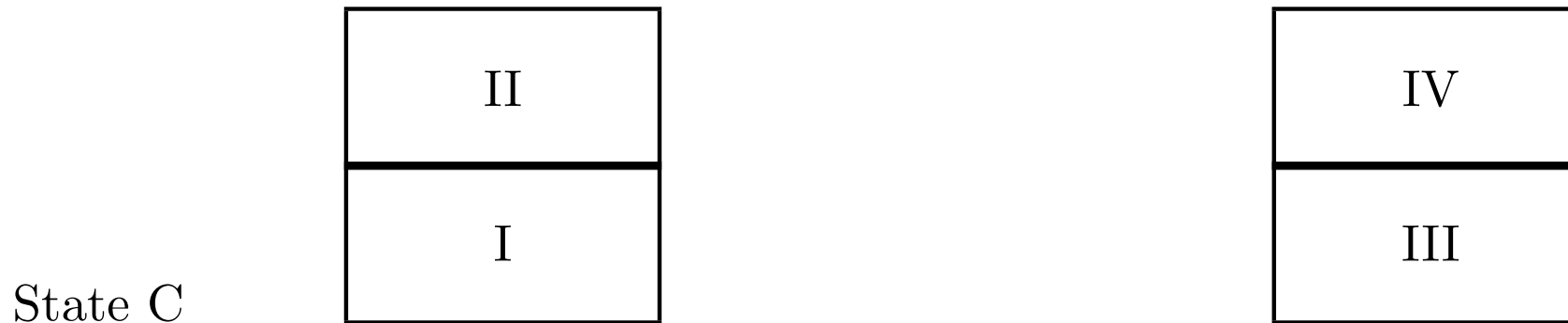
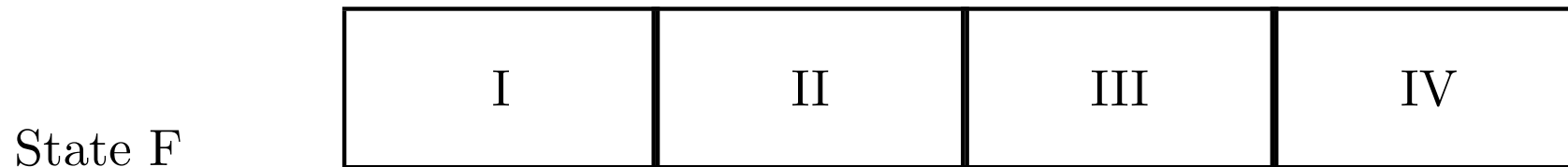
However, when the distribution of the values has more than one hump the standard deviation loses some of its usefulness especially in connection with the notion of uncertainty

Examples of two states of a particle in one dimension

First state describes a particle with a uniformly distributed probability
in a box of total length L

Second state describes a particle localized with equal probabilities
in two smaller boxes each of length $L/4$

The probability distributions corresponding to these two states
F=flat and **C=clustered**, are graphically represented as follows:



State F $\rho = \begin{cases} \frac{1}{L} & \text{inside the box,} \\ 0 & \text{outside the box,} \end{cases}$

State C $\rho = \begin{cases} \frac{2}{L} & \text{inside the box,} \\ 0 & \text{outside the box.} \end{cases}$

In which case, F or C, is the uncertainty in the position greater?

According to our intuition, the uncertainty is greater in the case F
In the case C we know more about the position; we know that the particle **is not** in the regions II and III. However, when we calculate the standard deviation Δx we obtain the opposite result:

Case F

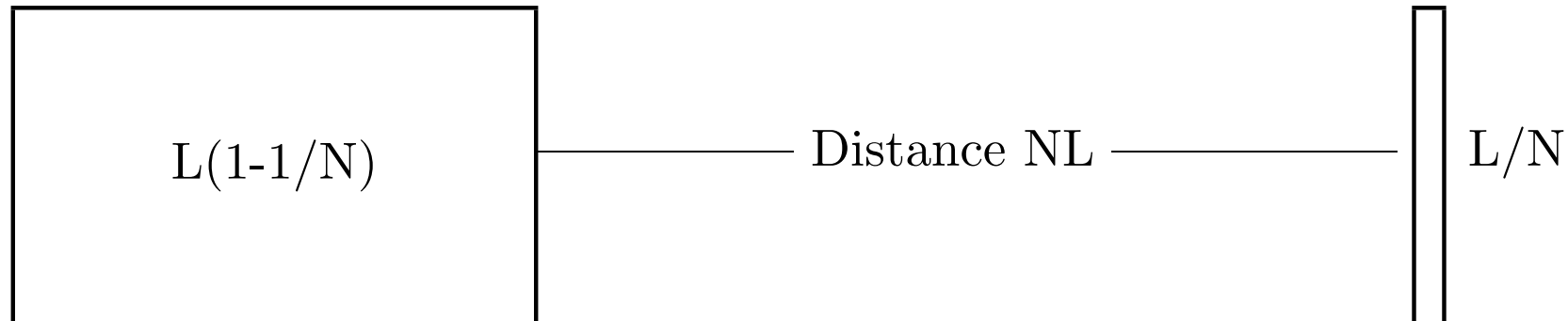
$$\Delta x_F = \frac{L}{\sqrt{12}}$$

Case C

$$\Delta x_C = \sqrt{\frac{7}{4}} \frac{L}{\sqrt{12}}$$

The second, somewhat more dramatic example of the situation where the standard deviation does not give a sensible measure of uncertainty is provided by the following distribution of probability.

State D



The probability density is constant in two regions I and II separated by a large distance NL (N is a large number)
The region I has the size $L(1 - 1/N)$ and the distant region II has the size L/N

$$\text{State D} \quad \rho = \begin{cases} \frac{1}{L} & \text{in region I,} \\ \frac{1}{L} & \text{in region II,} \\ 0 & \text{elsewhere.} \end{cases}$$

For large N the standard deviation Δx is approximately equal to:

Case D

$$\Delta x_D \sim \sqrt{1 + 12N} \frac{L}{\sqrt{12}}$$

This expression tends to infinity with N in spite of the fact that the probability of finding the particle in the region I tends to 1

This example shows clearly what is wrong with the standard deviation

It gets **very high contributions** from distant regions because these contribution enter with a large weight: namely, the distance from the mean value.

Information entropy as a measure of uncertainty

The best measure of information was introduced in 1948 by Shannon in his paper “The mathematical theory of communication”
Shannon entropy closely resembles the physical entropy

$$H = - \sum_i p_i \ln p_i$$

where p_i is the probability of the occurrence of the i -th event
or the a priori probability of the i -th message

In information theory one uses the logarithms to the base of 2

H is then measured in bits

but a change of the base results only in a change of the scale

The information entropy may serve as a precise measure of uncertainty

It has even been described already by Shannon as a measure of
“information, choice or **uncertainty**”

H may indeed serve at the same time as a measure of information and as a measure of uncertainty, because the two notions are directly related

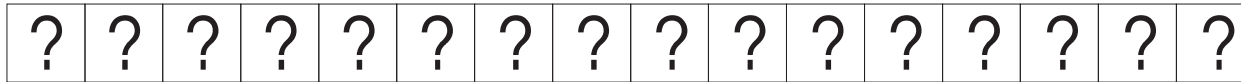
The information gained in the measurement removes the uncertainty about its outcome

The information uncertainty is in many ways a much better **measure of uncertainty** than the standard deviation

The Shannon entropy has a simple intuitive interpretation

The average number of questions needed to discover "the truth" (that is to remove uncertainty) which is hidden in one of the boxes with the probability distribution p_i **is bounded from below by H**

Initially we know nothing:



After the first question we know:



After the second question we know:

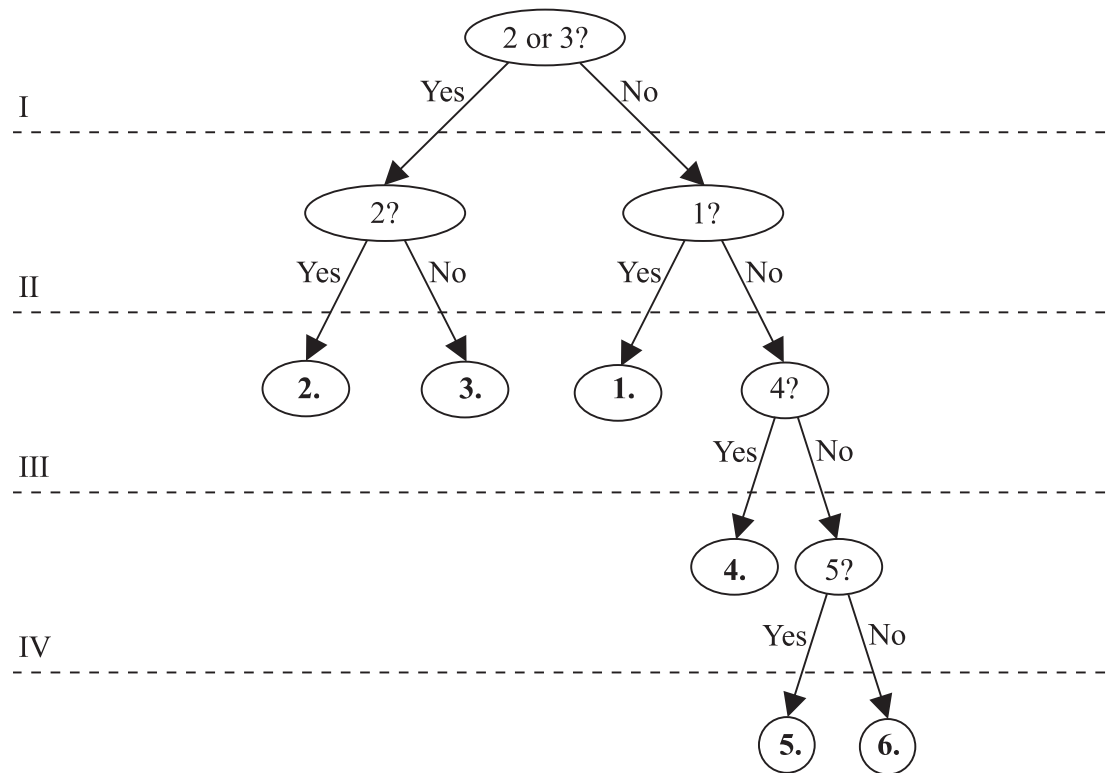


After the third question we know:



After four questions we know where the treasure is:





To apply the Shannon measure of uncertainty to quantum mechanics we must **discretize** the distributions of the position and momentum

Histograms

The probability density to find a quantum particle at x is given by the modulus squared of the wave function $\rho(x) = |\psi(x)|^2$

The probability density to find the momentum p is given by the modulus squared of the Fourier transform $\tilde{\rho}(p) = |\tilde{\psi}(p)|^2$

In order to have a discrete set of probabilities we introduce the **finite resolutions (bins)** for the measurements of x and p

To distinguish them from the statistical dispersions Δx and Δp they will be denoted by δx and δp

The probabilities q_l and p_k to find the values of x and p in the l -th and the k -th bin are

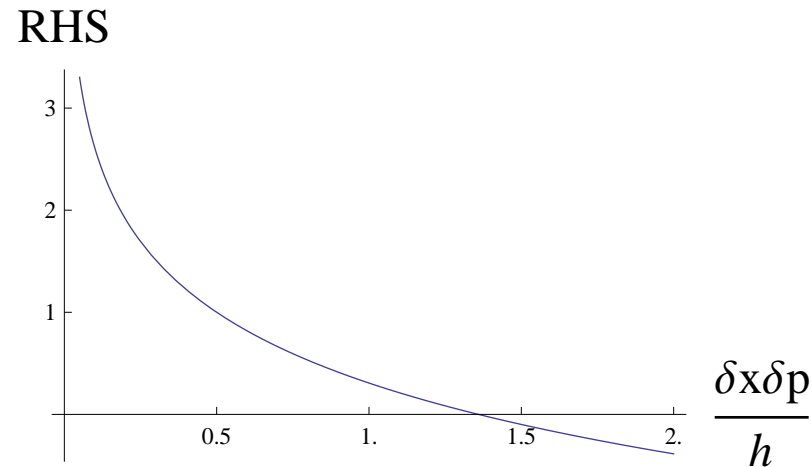
$$q_l = \int_{x_l}^{x_l + \delta x} dx \rho(x) \quad p_k = \int_{p_k}^{p_k + \delta p} dp \tilde{\rho}(p)$$

The values of q_l and p_k represent a histogram of measurements

Following D. Deutsch (1983) and H. M. Partovi (1983) we can formulate the uncertainty relations between the positions and momenta that restrict the available information about x and p measured by the Shannon entropies $H^{(x)} = -\sum q_l \ln q_l$ and $H^{(p)} = -\sum p_k \ln p_k$

The best known (but not sharp) bound expressing the uncertainty relations (I. Bialynicki-Birula 1984) is

$$H^{(x)} + H^{(p)} \geq 1 - \ln(2) - \ln\left(\frac{\delta x \delta p}{h}\right)$$



The Rényi entropy

Is there an uncertainty relation expressed in term of the Rényi entropy?

Rényi entropy is a **one parameter extension** of the Shannon entropy

$$H_\alpha = \frac{1}{1-\alpha} \ln\left(\sum p_k^\alpha\right)$$

The limit $\alpha \rightarrow 1$ produces the Shannon entropy ($\sum p_k = 1$)

$$\frac{d}{d\alpha} \ln\left(\sum p_k^\alpha\right) = \sum p_k \ln p_k$$

Rényi entropy has many applications in information theory (coding and cryptography), physics (particle production, diffusion), signal analysis, nonlinear dynamics, fractals, etc. (150 000 hits in Google)

Rényi entropy shares many properties with Shannon entropy

- H_α is nonnegative
- H_α vanishes when the uncertainty is the smallest (only one $p_k \neq 0$)
- H_α takes on the maximal value $\ln n$ when all n probabilities are equal
- H_α is extensive — satisfies the following condition If p_{ij} is a product of two probability distributions $p_{ij} = p'_i p''_j$ then

$$H_\alpha = \frac{1}{1-\alpha} \ln\left(\sum_{ij} p_{ij}^\alpha\right) = \frac{1}{1-\alpha} \ln\left(\sum_i p_i'^\alpha \sum_j p_j''^\alpha\right) = H'_\alpha + H''_\alpha$$

Rényi entropies for the position and momentum are easily defined using the same probabilities p_k and q_l as for the Shannon entropy

$$H_\alpha^{(x)} = \ln\left(\sum_l q_l^\alpha\right) / (1-\alpha) \quad H_\alpha^{(p)} = \ln\left(\sum_k p_k^\alpha\right) / (1-\alpha)$$

What are the limitations on the values of $H_\alpha^{(x)}$ and $H_\alpha^{(p)}$?

Brief mathematical interlude

The (q, q') -norm of T is the smallest $k(q, q')$ such that

$$\|T\psi\|_q \leq k(q, q') \|\psi\|_{q'} \quad 1/q + 1/q' = 1 \quad q \geq q'$$

$$\|\psi\|_q = \left(\int dx |\psi(x)|^q \right)^{\frac{1}{q}}$$

K. I. Babenko (1961) and W. Beckner (1975)

found $k(q, q')$ for the Fourier transformation

Choosing the physical normalization of $\tilde{\psi}$ we obtain

$$\tilde{\psi}(p) = \frac{1}{\sqrt{2\pi\hbar}} \int dx e^{-ipx/\hbar} \psi(x)$$

$$\|\tilde{\psi}\|_q \leq \hbar^{\frac{1}{q} - \frac{1}{2}} \left(\frac{q}{2\pi} \right)^{\frac{1}{2q}} \left(\frac{q'}{2\pi} \right)^{-\frac{1}{2q'}} \|\psi\|_{q'} \quad (\text{I})$$

This inequality is saturated for all Gaussian wave functions

It is convenient to change $q \rightarrow 2\alpha$ and $q' \rightarrow 2\beta$

$$\frac{1}{\alpha} + \frac{1}{\beta} = 2 \quad \frac{\alpha}{\alpha-1} = \frac{\beta}{1-\beta} \quad \alpha > 1 \implies \beta < 1$$

and rewrite the norms in the inequality (I) as follows

$$\|\tilde{\psi}\|_{2\alpha} = \left(\int dp |\tilde{\rho}(p)|^\alpha \right)^{\frac{1}{2\alpha}} \quad \|\psi\|_{2\beta} = \left(\int dx |\rho(x)|^\beta \right)^{\frac{1}{2\beta}}$$

$$\rho(x) = |\psi(x)|^2 \quad \text{and} \quad \tilde{\rho}(p) = |\tilde{\psi}(p)|^2$$

are quantum mechanical probability densities

After squaring both sides of the inequality (I) we obtain (for $\alpha > \beta$)

$$\|\tilde{\rho}\|_\alpha \leq n(\alpha, \beta) \|\rho\|_\beta \quad n(\alpha, \beta) = \hbar^{\frac{1}{\alpha}-1} \left(\frac{\alpha}{\pi} \right)^{\frac{1}{2\alpha}} \left(\frac{\beta}{\pi} \right)^{\frac{1}{2\beta}} \quad (\text{II})$$

To make use of this mathematical inequality in the proof of the physical uncertainty relations we have to introduce into this analysis the probabilities p_k and q_l . This is done by splitting the integrals into bins

$$\int dp (\tilde{\rho}(p))^\alpha = \sum \int_{x_k}^{x_k + \delta p} dp (\tilde{\rho}(p))^\alpha$$

and using the famous property of convex functions:

The value of the function at the average point cannot exceed the average value of the function

In the present case I need a somewhat more general version of this statement in the form of an integral Jensen inequality

$$\Phi \left(\int_D dx h(x) w(x) \right) \leq \int_D dx \Phi (h(x)) w(x)$$

which is valid for a convex function Φ , an arbitrary function $h(x)$ and a nonnegative weight function $w(x)$ normalized to 1 ($\int_D dx w(x) = 1$)

The power x^α is a convex function for $\alpha > 1$ hence

$$\left(\frac{p_k}{\delta p}\right)^\alpha = \left(\frac{\int_{x_k}^{x_k+\delta p} dp \tilde{\rho}(p)}{\delta p}\right)^\alpha \leq \frac{\int_{x_k}^{x_k+\delta p} dp (\tilde{\rho}(p))^\alpha}{\delta p}$$

$$\delta p^{1-\alpha} \sum p_k^\alpha \leq \int dp (\tilde{\rho}(p))^\alpha = (\|\tilde{\rho}\|_\alpha)^\alpha \quad (\text{A})$$

For concave functions we have: The average value of the function cannot exceed the value of the function at the average point

$$(\|\rho\|_\beta)^\beta = \int dx (\rho(x))^\beta \leq \delta x^{1-\beta} \sum q_l^\beta \quad (\text{B})$$

Raising (A) to the power $1/\alpha$ and (B) to the power of $1/\beta$ and we the use of the basic inequality (II) we obtain

$$\delta p^{\frac{1-\alpha}{\alpha}} \left(\sum p_k^\alpha\right)^{\frac{1}{\alpha}} \leq n(\alpha, \beta) \delta x^{\frac{1-\beta}{\beta}} \left(\sum q_l^\beta\right)^{\frac{1}{\beta}}$$

$$\frac{1-\alpha}{\alpha} \ln(\delta p) + \frac{1}{\alpha} \ln\left(\sum p_k^\alpha\right) \leq \frac{1-\beta}{\beta} \ln(\delta x) + \ln n(\alpha, \beta) + \frac{1}{\beta} \ln\left(\sum q_l^\beta\right)$$

Final form of the uncertainty relation

From the last inequality with the use of $\alpha/(\alpha - 1) = \beta/(1 - \beta)$ after some work we obtain finally

$$H_{\beta}^{(x)} + H_{\alpha}^{(p)} \geq -\frac{1}{2} \left(\frac{\ln \alpha}{1 - \alpha} + \frac{\ln \beta}{1 - \beta} \right) - \ln(2) - \ln \left(\frac{\delta x \delta p}{h} \right)$$

In the limit when $\alpha \rightarrow 1$ also $\beta \rightarrow 1$

With the help of the l'Hôpital rule one obtains

$$\lim_{\alpha \rightarrow 1} \frac{\ln \alpha}{1 - \alpha} = -1$$

The inequality for Rényi entropies reduces to the uncertainty relation for the Shannon entropies proved by me twenty years ago

$$H^{(x)} + H^{(p)} \geq 1 - \ln(2) - \ln \left(\frac{\delta x \delta p}{h} \right)$$

Uncertainty relations for N -level systems

For quantum systems described by vectors in the N -dimensional Hilbert space the analog of the uncertainty relation for the Rényi entropies is

$$\frac{1}{1-\alpha} \ln \left(\sum_{k=1}^N \tilde{\rho}_k^\alpha \right) + \frac{1}{1-\beta} \ln \left(\sum_{l=1}^N \rho_l^\beta \right) \geq \ln N,$$

where $\tilde{\rho}_k = |\tilde{a}_k|^2$, $\rho_l = |a_l|^2$ and the amplitudes \tilde{a}_k and a_l are connected by the discrete Fourier transformation

$$\tilde{a}_k = \frac{1}{\sqrt{N}} \sum_{l=1}^N \exp(2\pi i k l / N) a_l.$$

This uncertainty relation is saturated for the states that are either **localized in “position space”** exactly one of the amplitudes a_l is different from zero or are **localized in “momentum space”** exactly one of the amplitudes \tilde{a}_k is different from zero
The bound does not depend on α and β

Uncertainty relations for mixed states

The uncertainty relations for the Rényi entropies also
hold for all mixed states

This result is not obvious because the Rényi entropy is not a convex function of the probability distributions for all values of α

The terms on the left hand side of the uncertainty relation may decrease as a result of mixing

In order to extend our results to mixed states one has to go back to the original inequalities that were used in the proof and show that they still hold

Uncertainty relations for continuous distributions

There also exist purely mathematical versions of the uncertainty relations that **do not involve** the experimental resolutions δx and δp of the measuring devices

$$\begin{aligned} \frac{1}{1-\alpha} \ln \left(\int_{-\infty}^{\infty} dp (\tilde{\rho}(p))^{\alpha} \right) + \frac{1}{1-\beta} \ln \left(\int_{-\infty}^{\infty} dx (\rho(x))^{\beta} \right) \\ \geq -\frac{1}{2(1-\alpha)} \ln \frac{\alpha}{\pi} - \frac{1}{2(1-\beta)} \ln \frac{\beta}{\pi} \end{aligned}$$

On the left hand side of this inequality we have, what might be called, the continuous or integral versions of the Rényi entropies

In order to derive this purely mathematical inequality that includes **no reference** to the finite resolution of the physical measurements

I have dropped \hbar in the definition of the Fourier transform

This inequality has been independently proven by Zozor and Vignat

Analogous relations for the continuous Tsallis entropies

for x and p were obtained by Rajagopal

Uncertainty relations for continuous Shannon entropies

Assuming that the distributions $\rho(x)$ and $\tilde{\rho}(p)$ are normalized in the limit, when $\alpha \rightarrow 1$, $\beta \rightarrow 1$ from the formula for the entropic uncertainty relation in terms of the Rényi entropies we obtain the classic uncertainty relation for the continuous Shannon entropies

$$-\int_{-\infty}^{\infty} dp \tilde{\rho}(p) \ln \tilde{\rho}(p) - \int_{-\infty}^{\infty} dx \rho(x) \ln \rho(x) \geq \ln(e\pi)$$

This inequality had been conjectured by Hirschman (1957) and later proved by Bialynicki-Birula and Mycielski (1975) and by Beckner (1975)

References

1. I. Bialynicki-Birula, “Renyi entropy and the uncertainty relations”, Foundations of Probability and Physics, Ed. G. Adenier, C. A. Fuchs, and A. Yu. Khrennikov, American Institute of Physics, Melville, 2007.
2. I. Bialynicki-Birula, “Formulation of the uncertainty relations in terms of the Renyi entropies” Phys. Rev. A 74, 052101 (2006).

Both references can be found on my home page: www.cft.edu.pl/~birula

Summary

I have shown that the Rényi entropies for canonically conjugate physical variables obey the inequalities that have the interpretation of the uncertainty relations

In order to use the Rényi entropies in their standard form (i.e. using a discrete set of probabilities), I had to introduce the finite resolutions δx and δp of the measuring devices

The resulting new uncertainty relations for the position and momentum acquire an additional meaning not found in their standard version

The more accurately we wish to measure x and p the stronger the limitations imposed by the uncertainty relation

The lower limit $-\frac{1}{2} \left(\frac{\ln \alpha}{1 - \alpha} + \frac{\ln \beta}{1 - \beta} \right) - \ln \left(\frac{\delta x \delta p}{\pi \hbar} \right)$ increases