Heisenberg's inequality*

Michael Walter

1 Introduction

1 Definition. Let $\mathbb{P}$ be a probability measure on $\mathbb{R}$. Then its *mean* or *expectation* is defined as

$$\mathbb{E}(\mathbb{P}) := \int x \, d\mathbb{P}(x)$$

Its *variance* is the expected square deviation from the mean, that is

$$V(\mathbb{P}) := \int (x - \mathbb{E}(\mathbb{P}))^2 \, d\mathbb{P}(x)$$

Thus it is small if the probability measure is sharply localized and large if it is spread out widely.

If $\mathbb{P}$ is given by a probability density function $\rho \in L^1(\mathbb{R})$, i.e. if

$$d\mathbb{P} = \rho \, d\lambda,$$

then we also write $\mathbb{E}(\rho)$ and $V(\rho)$.

2 Example. Let $f \in L^2(\mathbb{R})$ with $||f||_{L^2} = 1$. Then $||\hat{f}||_{L^2} = 1$ by Plancherel, and both $|f|^2$ and $|\hat{f}|^2$ are probability density functions.

We can now ask whether there is any relationship between the variance of the induced probability measures.

The following inequality which we will prove in section 3 not only gives an affirmative answer to this question but also a concrete lower bound.

3 Theorem (Heisenberg's inequality)**.** *Let $f \in L^2(\mathbb{R})$ with $||f||_{L^2} = 1$. Then*

$$V(|f|^2)V(|\hat{f}|^2) \geq \frac{1}{4},$$

It is an example of the *uncertainty principle* which in the words of Folland and Sitaram [1] says that

> "a nonzero function and its Fourier transform cannot both be sharply localized"

2 The quantum mechanical motivation

Recall that in **classical mechanics**, the state of a physical system at a certain time is given by a point in phase space, that is, by (generalized) positions q and momenta p.

All observable quantities, which are often simply called observables, are functions of q and p. Measuring an observable is assumed to be deterministic and not to modify the state of the system.

In **quantum mechanics**, the *state* of a physical system is given by a unit vector φ in a Hilbert space H.

An *observable* is represented by a self-adjoint operator

$$A : D(A) \subseteq H \to H$$

and its spectrum $\sigma(A)$ gives the set of *observable values.*

If H is finitely-dimensional (which occurs for example with *spin* observables, e.g. Pauli matrices) then the postulates of quantum mechanics tell us that, given a physical system in state φ, the probability that we measure an eigenvalue λ corresponding to a *single* eigenstate ψ of the observable is simply given by

$$|\langle \varphi, \psi \rangle|^2$$

*Based on Folland and Sitaram [1] and Petersen [2].

(In the case of an arbitrary eigenspace V, we take the square norm of the state's projection onto V.) Thus, measuring an observable in general is a nondeterministic process[1]!

In the general case we consider the Borel functional calculus for A by which we have the *projection-valued spectral measure*

$$\mathcal{B}(\sigma(A)) \to B(H),\ E \mapsto 1_E(A)$$

Then for any unit vector φ

$$\mathbb{P}_{A,\varphi} : \mathcal{B}(\sigma(A)) \to [0,1],\ E \mapsto \langle 1_E(A)\varphi, \varphi\rangle$$

defines a probability measure on $\sigma(A)$ which describes the probability distribution of the outcome of a measurement of A if the system is in state φ.

In stochastical terms, the result of such a measurement can be modelled by a $\sigma(A)$-valued random variable $M_{A,\varphi} \sim \mathbb{P}_{A,\varphi}$.

The following proposition expresses integrals over this probability measure in terms of the functional calculus. It is a special case of a more general theorem from spectral theory.

4 Proposition. *Let A be an observable, $p \in \mathbb{R}[X]$ a polynomial and $\varphi \in D(p(A))$ a state. Then $p \in L^1(\mathbb{P}_{A,\varphi})$ and we have*

$$\int p \; d\mathbb{P}_{A,\varphi} = \langle p(A)\varphi, \varphi\rangle$$

($p(A)$ is defined in the obvious way with maximal domain.)

It is now easy to find concrete formulae for the expectation and variance of a measurement.

5 Corollary. *Let A be an observable and let $\varphi \in D(A)$, $\psi \in D(A^2)$ be states. Then*

$$\begin{aligned} \mathbb{E}(\mathbb{P}_{A,\varphi}) &= \langle A\varphi, \varphi\rangle \\ V(\mathbb{P}_{A,\psi}) &= ||(A - \mathbb{E}(\mathbb{P}_{A,\psi}))\psi||^2 \end{aligned}$$

Proof. We apply the preceding proposition to $p := X$ and $q := (X - \mathbb{E}(\mathbb{P}_{A,\varphi}))^2$ and arrive at

$$\begin{aligned} \mathbb{E}(\mathbb{P}_{A,\varphi}) &= \langle A\varphi, \varphi\rangle \\ V(\mathbb{P}_{A,\psi}) &= \big\langle (A - \mathbb{E}(\mathbb{P}_{A,\psi}))^2\psi, \psi\big\rangle = ||(A - \mathbb{E}(\mathbb{P}_{A,\psi}))\psi||^2 \end{aligned}$$

(In the last step we used that A is self-adjoint.) □

We will now prove the following abstract uncertainty inequality which relates the variance of observables A, B to the expectation of their commutator $[A, B] := AB - BA$ (which in a sense indicates their degree of non-commutativity).

6 Corollary (Heisenberg's inequality for observables)**.** *Let A, B be observables and $\varphi \in D([A,B]) \cap D(A^2) \cap D(B^2)$ a state. Then*

$$V(\mathbb{P}_{A,\varphi})V(\mathbb{P}_{B,\varphi}) \geq \frac{1}{4}|E(\mathbb{P}_{[A,B],\varphi})|^2$$

Proof. We apply the preceding corollary. Let $\alpha := \mathbb{E}(\mathbb{P}_{A,\varphi})$, $\beta := \mathbb{E}(\mathbb{P}_{B,\varphi})$. Then $A-\alpha$ and $B-\alpha$ are observables with

$$[A - \alpha, B - \beta] = [A, B]$$

and we see that without loss of generality we may assume that $\alpha = \beta = 0$. Thus the claim follows from

$$\begin{aligned} & |E(\mathbb{P}_{[A,B],\varphi})| = |\langle [A,B]\varphi, \varphi\rangle| \\ = \; & |\langle AB\varphi, \varphi\rangle - \langle BA\varphi, \varphi\rangle| = |\langle B\varphi, A\varphi\rangle - \langle A\varphi, B\varphi\rangle| \\ = \; & 2|\operatorname{Im}\langle A\varphi, B\varphi\rangle| \leq 2|\langle A\varphi, B\varphi\rangle| \leq 2||A\varphi||||B\varphi|| \end{aligned}$$

□

[1] We shall also mention (but without making this statement mathematically precise) that in quantum mechanics a measurement leads to the system's state collapsing to an eigenstate corresponding to the measured eigenvalue.

7 Remark. The physical interpretation of this inequality is that if two observables do not commute they cannot be both measured exactly at the same time.

There is a subtle problem lurking in the back of this seemingly innocent statement, though, which is that in general the domain of the commutator will be rather small (it can even be trivial!). Thus the preceding proposition is a lot less powerful a statement than one would hope for.

8 Example. A quantum mechanical system consisting of a single particle is modelled using the Hilbert space $L^2(\mathbb{R})$.

The particle's position and momentum are given by the self-adjoint operators

$$\begin{aligned} D(Q) &:= \{\varphi \in L^2(\mathbb{R}) : (\cdot)\varphi \in L^2(\mathbb{R})\} \\ Q\varphi &:= (\cdot)\varphi \end{aligned}$$

and

$$\begin{aligned} D(P) &:= \{\varphi \in L^2(\mathbb{R}) : (\cdot)\mathcal{F}\varphi \in L^2(\mathbb{R})\} \\ P\varphi &:= \mathcal{F}^{-1}((\cdot)\mathcal{F}\varphi) = \mathcal{F}^{-1}Q\mathcal{F}\varphi \end{aligned}$$

By theorem 10 we know that P is simply the weak derivative, i.e.

$$\begin{aligned} D(P) &= W^{1,2}(\mathbb{R}) \\ P\varphi &= -iD\varphi \end{aligned}$$

Thus the commutator is given by

$$\begin{aligned} & [Q, P]\varphi = -i((\cdot)D\varphi - D((\cdot)\varphi)) = i\varphi \\ \Rightarrow\ & [Q, P] = i \end{aligned}$$

on its domain (by the product rule which is valid since $(\cdot) = \mathrm{id} \in C^\infty(\mathbb{R})$; compare lemma 12).

We will now explicitely calculate the inequality provided by corollary 6. We have

$$\begin{aligned} |\mathbb{E}(\mathbb{P}_{[Q,P],\varphi})| &= |\langle [Q, P]\varphi, \varphi\rangle| = ||\varphi||_{L^2}^2 = 1 \\ \mathbb{E}(\mathbb{P}_{Q,\varphi}) &= \langle Q\varphi, \varphi\rangle = \int x|\varphi(x)|^2 dx = \mathbb{E}(|\varphi|^2) \\ \mathbb{E}(\mathbb{P}_{P,\varphi}) &= \langle P\varphi, \varphi\rangle = \langle Q\hat{\varphi}, \hat{\varphi}\rangle = \int x|\hat{\varphi}(x)|^2 dx = \mathbb{E}(|\hat{\varphi}|^2) \end{aligned}$$

$$\begin{aligned} V(\mathbb{P}_{Q,\varphi}) &= ||(Q - \mathbb{E}(\mathbb{P}_{Q,\varphi}))\varphi||_{L^2}^2 = \\ &= \int (x - \mathbb{E}(|\varphi|^2))^2 |\varphi(x)|^2 dx = V(|\varphi|^2) \\ V(\mathbb{P}_{P,\varphi}) &= ||(P - \mathbb{E}(\mathbb{P}_{P,\varphi}))\varphi||_{L^2}^2 = ||(Q - \mathbb{E}(\mathbb{P}_{P,\varphi}))\hat{\varphi}||_{L^2} = \\ &= \int (\xi - \mathbb{E}(|\hat{\varphi}|^2))^2 |\hat{\varphi}(\xi)|^2 d\xi = V(|\hat{\varphi}|^2) \end{aligned}$$

for all $\varphi \in D(Q^2) \cap D(P^2) \cap D([Q, P] =: D$.

Hence corollary 6 applied to Q and P is nothing but Heisenberg's inequality for L^2-functions, that is, theorem 3 from the first section — except that the latter holds for arbitrary $\varphi \in L^2(\mathbb{R})$ and gives interesting results fo $\varphi \in D(Q) \cap D(P)$ which is a strictly larger domain than D!

3 Heisenberg's inequality

We will now state theorem 3 again for the convenience of the reader, and give its proof in multiple steps.

3 Theorem (Heisenberg's inequality)**.** *Let $f \in L^2(\mathbb{R})$ with $||f||_{L^2} = 1$. Then*

$$V(|f|^2)V(|\hat{f}|^2) \geq \frac{1}{4},$$

The following lemma shows that it is enough to consider centered functions.

9 Lemma. *Let $f \in L^2(\mathbb{R})$ with $||f||_{L^2} = 1$. Then there exists $g \in L^2(\mathbb{R})$ with $||g||_{L^2} = 1$ and*

$$V(|f|^2)V(|\hat{f}|^2) = (\int x^2|g(x)|^2dx)(\int \xi^2|\hat{g}(\xi)|^2d\xi)$$

Proof. Note that for $\varphi \in L^2(\mathbb{R})$ and $\psi(x) := \varphi(x+a)$ $(a \in \mathbb{R})$ we have

$$||\psi||_{L^2} = ||\varphi||_{L^2}$$
$$|\hat{\psi}(\xi)| = |e^{ia\xi}\hat{\varphi}(\xi)| = |\hat{\varphi}(\xi)|$$

Thus if we replace f by $h : x \mapsto f(x+\mathbb{E}(|f|^2))$ we get rid of the expectation in the first integral while leaving the second integral untouched. Now substituting $\hat{h}$ by $\hat{g} : \xi \mapsto \hat{h}(\xi+\mathbb{E}(|\hat{h}|^2))$ leads to the desired representation. □

Proof of theorem 3. By the preceding lemma it is enough to prove the inequality

$$(\int x^2|f(x)|^2dx)(\int \xi^2|\hat{f}(\xi)|^2d\xi) \geq \frac{1}{4}$$

We can further assume that $(\cdot)f$, $(\cdot)\hat{f} \in L^2(\mathbb{R})$ since otherwise the claim is trivial.

By theorem 10 we have $f \in W^{1,2}(\mathbb{R})$, i.e. $Df \in L^2(\mathbb{R})$. Note that

$$\begin{aligned}
(\cdot)f\bar{f} &\in L^1(\mathbb{R}) \\
D((\cdot)f\bar{f}) &\overset{12}{=} (\cdot)D(f\bar{f}) + f\bar{f} \\
&\overset{13}{=} f\bar{f} + (\cdot)fD(\bar{f}) + (\cdot)\bar{f}Df \in L^1(\mathbb{R}) \\
&= f\bar{f} + 2\,\mathrm{Re}((\cdot)fD(\bar{f}))
\end{aligned}$$

(using the indicated product rules from the appendix). Thus by proposition 11 we get

$$\begin{aligned}
1 = ||f||_{L^2}^2 &= -2\int \mathrm{Re}((\cdot)fD(\bar{f}))d\lambda \leq 2||(\cdot)fD(\bar{f})||_{L^1} \\
&\leq 2||(\cdot)f||_{L^2}||Df||_{L^2} = 2||(\cdot)f||_{L^2}||(\cdot)\hat{f}||_{L^2}
\end{aligned}$$

(employing Cauchy-Schwartz and Plancherel) □

The proof also shows that equality only occurs when

$$Df = \alpha(\cdot)f$$

for some *real* constant $\alpha < 0$ (otherwise the first inequality would be violated). By theorem 10 f has a continuous representative, hence does Df and a glance at the ODE

$$f'(x) = \alpha x f(x)$$

shows that

$$f(x) = \exp(\frac{1}{2}\alpha x^2) \qquad \text{(a.e.)}$$

and indeed for $\alpha < 0$ we have $f \in L^2(\mathbb{R})$.

4 Perspective

There is a lot more to uncertainty principles in harmonic analysis than Heisenberg's inequality; for a modern survey article see e.g. Folland and Sitaram [1].

For example we can make the first section's *uncertainty principle*, i.e. that

> "a nonzero function and its Fourier transform cannot both be sharply localized",

more precise in the following sense: If $f \in \mathcal{L}^2$ is sharply localized, its Fourier transform must not only have large variance (which follows from Heisenberg's inequality) but it cannot be strongly concentrated around of two or more distant points either! Inequalities in this spirit are often called *local uncertainty inequalities*.

A Some facts about weak derivatives

10 Theorem. *Let $f \in L^2(\mathbb{R})$ such that $(\cdot)\hat{f} \in L^2(\mathbb{R})$. Then $f \in W^{1,2}(\mathbb{R})$. and f has a continuous representant.*

Proof. This follows from lemma 22 and theorem 23 of Stefanie Wolz' talk. □

11 Proposition. *Let $f \in W^{1,1}(\mathbb{R})$. Then*

$$\int Df \; d\lambda = 0$$

Proof. Choose a test function $\varphi \in C_c^\infty(\mathbb{R})$ such that $0 \leq \varphi \leq 1$ and $\varphi = 1$ in a neighborhood of the origin. Put $\varphi_n(x) := \varphi(\frac{x}{n})$. Then $\varphi_n \to 1$ pointwise, and

$$\begin{aligned} |\int (Df)\varphi_n d\lambda| &= |\int f(D\varphi_n) d\lambda| \\ \leq \; ||f||_{L^1} ||(D\varphi_n)||_\infty &= \frac{1}{n} ||f||_{L^1} ||D\varphi||_\infty \to 0 \end{aligned}$$

On the other hand, dominated convergence is applicable as $Df \in L^1(\mathbb{R})$ and we see that the integral converges to $\int (Df) d\lambda$. □

12 Lemma. *Let $f \in L^1_{loc}(\mathbb{R})$ have a weak derivative and let $\varphi \in C^\infty(\mathbb{R})$. Then $f\varphi$ has the weak derivative*

$$D(f\varphi) = (Df)\varphi + f(D\varphi)$$

Proof. For any test function $\psi \in C_c^\infty(\mathbb{R})$ it is clear that $\varphi\psi$ is also a test function and thus we have

$$\begin{aligned} \int f\varphi(D\psi) \; d\lambda &= \int \left(fD(\varphi\psi) - f(D\varphi)\psi\right) d\lambda \\ &= \int \left(-(Df)\varphi\psi - f(D\varphi)\psi\right) d\lambda \\ &= -\int \left((Df)\varphi + f(D\varphi)\right)\psi \; d\lambda, \end{aligned}$$

□

13 Theorem (Product rule)**.** *Let $f, g \in W^{1,2}(\mathbb{R})$. Then*

$$D(fg) = (Df)g + f(Dg) \in L^1(\mathbb{R})$$

and

$$\int f(Dg) d\lambda = -\int (Df) g d\lambda$$

Proof. We first show the latter identity by an approximation argument; the product rule then follows easily.

(1) Let $(\varphi_n) \subseteq C_c^\infty(\mathbb{R})$ be an approximate identity and put $f_n := f \star \varphi_n$. A simple calculation shows that $Df_n = (Df) \star \varphi_n$, and we know that

$$\begin{aligned} &(f_n), (Df_n) \subseteq C^\infty(\mathbb{R}) \cap L^2(\mathbb{R}), \\ &f_n \overset{L^2}{\to} f \quad \text{and} \quad Df_n \overset{L^2}{\to} Df \end{aligned}$$

Note that

$$\begin{aligned} &f_n g \in L^1(\mathbb{R}), \\ &D(f_n g) \overset{12}{=} (Df_n)g + f_n(Dg) \in L^1(\mathbb{R}) \end{aligned}$$

and by proposition 11 we have

$$\int (Df_n) g d\lambda = -\int f_n(Dg) d\lambda$$

Thus using Hölder's inequality we conclude that

$$\int (Df)gd\lambda = -\int f(Dg)d\lambda$$

(2) We now show the product rule. Let $\psi \in C_c^\infty(\mathbb{R})$. Then

$$\begin{aligned} &g\psi \in L^2(\mathbb{R}) \\ &D(g\psi) \overset{12}{=} (Dg)\psi + g(D\psi) \in L^2(\mathbb{R}) \end{aligned}$$

hence we can apply (1) to f and $g\psi$

$$\begin{aligned} \int fg(D\psi)d\lambda &= \int fD(g\psi)d\lambda - \int f(Dg)\psi d\lambda \\ &\overset{(1)}{=} -\int (Df)g\psi d\lambda - \int f(Dg)\psi d\lambda \\ &= -\int ((Df)g + f(Dg))\psi d\lambda \end{aligned}$$

□

References

[1] Gerald B. Folland and Alladi Sitaram. The uncertainty principle: A mathematical survey. *The Journal of Fourier Analysis and Applications*, 3(3):207–238, 1997.

[2] B. E. Petersen. Weak derivatives and integration by parts. *The American Mathematical Monthly*, 85(3):190–191, 1978.