DE GRUYTER
OPEN

DEMONSTRATIO MATHEMATICA
Vol. XLVII No 2 2014

Maciej J. Mączyński

ON SOME CONSEQUENCES OF THE FUNCTIONAL GENERALIZATION OF THE PARALLELOGRAM IDENTITY

Abstract. The aim of this paper is to unify the partial results, which up to now, have been dispersed in various publications in order to show the importance of the functional form of parallelogram identity in mathematics and physics. We study vector spaces admitting a real non-negative functional which satisfies an identity analogous to the parallelogram identity in normed vector spaces. We show that this generalized parallelogram identity also implies an equality analogous to the Cauchy–Schwarz inequality. We study the consequences of this identity in real and complex vector spaces, in generalized Riesz spaces and in abelian groups. We give a physical interpretation to these results. For vector spaces of observables and states, we show that the parallelogram identity implies an inequality analogous to Heisenberg's uncertainty principle (HUP), and we show that we can obtain the standard structure of quantum mechanics from the parallelogram identity, without assuming from the beginning the HUP. The role of complex numbers in quantum mechanics is discussed.

0. Introduction

The parallelogram identity (called also the parallelogram law) expresses an elementary property of a parallelogram telling that the sum of squares of the lengths of sides is equal to the sum of squares of the lengths of its diagonals. It can be expressed also as a property of norm in normed vector spaces. It turns out that if the parallelogram identity holds in a normed vector space then the norm arises from an inner product—i.e. a normed vector space satisfying the parallelogram identity is an inner product space. This was proved by P. Jordan and J. von Neumann in [1] and is called the Jordan–von Neumann theorem. The Jordan–von Neumann theorem has a direct physical interpretation—it implies that a normed space with the parallelogram identity, which is a Banach space, admits a structure of Hilbert space, and consequently can be used as a model for quantum me-

2010 *Mathematics Subject Classification*: Primary 81P10, Secondary 15A63.
Key words and phrases: parallelogram identity, inner product space, Cauchy–Schwarz inequality, Heisenberg inequality, quantum mechanics.

DOI: 10.2478/dema-2014-0040

chanics. In this model, we have immediately the Cauchy–Schwarz inequality, which implies Heisenberg's inequality. In the present paper, we would like to show that the parallelogram identity even without the structure of normed vector space implies an inequality which admits a physical interpretation analogous to Heisenberg's inequality. Our considerations are based on some earlier papers of the author et. al. We would like to unify these results and derive physical consequences from them. We will formulate the parallelogram identity in abstract form for vector space with a functional f and we will study the consequences of this identity in general vector spaces (Sec. 1), also in Sec. 2 when the functional f is real semi-continuous at zero (a norm is an example of such functional), and when the vector space is a generalized Riesz space or K-space with a probability measure (Sec. 3). We also consider in Sec. 4 the consequences of the parallelogram identity in abelian groups (in particular, the addition of vectors in a vector space provides the structure of abelian group). We will show that in all these structures the parallelogram identity implies an inequality analogous to the Cauchy–Schwarz inequality, which can be given a physical interpretation. Finally in Sec. 5, we will apply our results to the vector spaces of observables and states in the axiomatic structure of quantum mechanics. We will show that in these structures the parallelogram identity leads to Heisenberg's inequality, hence to the standard structure of quantum mechanics. We show why, to obtain Heisenberg's uncertainty principle for position and momentum observables, it is necessary to assume the structure of a complex vector space—without complex numbers (i.e. only in real spaces) the uncertainty principle cannot be derived. We hope that our consideration will show that quantum mechanics can be considered as a mathematical theory with equational axioms (the parallelogram identity is expressed in the form of equality). It is known from mathematical logic that the theories with equational axioms have a simpler structure - the theorem of Birkhoff can be applied to them.

1. The parallelogram identity

The parallelogram identity for a parallelogram with the vertices ABCD can be expressed in the following form:

$$|AB|^2 + |BC|^2 + |CD|^2 + |DA|^2 = |AC|^2 + |BD|^2.$$

In the normed spaces, this expression can be written in the form:

$$2\|x\|^2 + 2\|y\|^2 = \|x+y\|^2 + \|x-y\|^2.$$

As we mentioned above, in normed spaces this identity implies that the norm arises from an inner product (the Jordan–von Neumann theorem).

If we put $f(x) := \|x\|^2$, then the above identity can be written in the form:

$$(*) \qquad f(x+y) + f(x-y) = 2f(x) + 2f(y).$$

This is a functional form of the parallelogram identity.

The parallelogram identity is a very strong property. This is seen in the following theorem, which is a generalization of the Jordan–von Neumann theorem. With some additional assumption on f (real semi-continuity at zero, see below) this theorem was proved in [2]. It turns out that the inequality $(**)$ can be derived even without this additional assumption.

Theorem 1.1. *Let V be a real or complex vector space and $f : V \to \mathbb{R}$ a function on V, with positive values for all $x \neq 0$, satisfying the above parallelogram identity $(*)$ for all $x, y \in V$. Then the function f satisfies the following inequality:*

$$(**) \qquad f(x)f(y) \geq \frac{1}{4}|f(x+y) - f(x) - f(y)|^2$$

for all $x, y \in V$.

Substituting $a = f(x+y)$, $b = f(x)$, $c = f(y)$, the inequality $(**)$ can be written in the following more symmetrical form:

$$2(ab + bc + ac) \geq a^2 + b^2 + c^2$$

or in the form

$$\frac{\sqrt{ab + bc + ac}}{a + b + c} \geq \frac{1}{2}.$$

This inequality can be interpreted as some kind of the uncertainty rule implied by the parallelogram identity $(*)$. We can also interpret it geometrically.

Proof. Because the idea of the proof of this theorem will be used in the sequel, we will present this proof. The proof uses some methods from the original proof of Jordan–von Neumann [1]. First, let us observe that if we put in $(*)$ $x = y = 0$, then we obtain $f(0) = 0$.

Now let

$$h(x, y) = \frac{1}{2}((f(x+y) - f(x) - f(y)),$$

for all $x, y \in V$. It is obvious that h is symmetric.

We will show that h is additive with respect to x and consequently with respect to y. From $(*)$ we obtain

$$f(x+y) - f(x) - f(y) = -f(x-y) + f(x) + f(y).$$

Putting in $(*)$ $x = -y$, we obtain $f(-y) = f(y)$ for all $y \in V$. This implies that $h(x,y) = -h(x,-y) = -h(-x,y)$. From $(*)$ we obtain also

$$\begin{aligned} f(x+y) - f(x-y) &= 2(-f(x-y) + f(x) + f(x)) \\ &= -2(f(x+(-y)) - f(x) - f(-y)) \\ &= -4h(x,-y) = 4h(x,y). \end{aligned}$$

Hence we have

(i) $$f(x+y) - f(x-y) = 4h(x,y), \quad \text{for all } x, y \in V.$$

Substituting $x+z$ and $x-z$ for x in $(*)$, we obtain

$$f((x+y)+z) + f((x-y)+z) = 2f(x+z) + 2f(y),$$

and

$$f((x+y)-z) + f((x-y)-z) = 2f(x-z) + 2f(y).$$

Subtracting side by side and applying (i), we obtain

(ii) $$h(x+y,z) + h(x-y,z) = 2h(x,z).$$

Putting $y = x$ in (ii), we obtain $h(2x,z) = 2h(x,z)$. Substituting $x+y$ by x_0 and $x-y$ by y_0 in (ii), we obtain

$$h(x_0,z) + h(y_0,z) = h(2x,z) = h(x_0+y_0,z).$$

Hence h is additive with respect to x, and also to y.

This implies that $h(2x,z) = 2h(x,z)$. By induction, we infer that $h(mx,z) = mh(x,z)$ for all non-negative integers m. We also have that $h(-mx,z) = -mh(x,z)$. Putting mx = y, we obtain $h((1/m)y,z) = (1/m)h(y,z)$ for $m \neq 0$. Hence $h(-x,y) = \lambda h(x,y)$ for all rational λ.

Next putting $x = y$ in $(*)$, we obtain $f(2x) = 4f(x)$. Similarly as above we infer that $f(\lambda x) = \lambda^2 f(x)$ for all rational λ.

Now for fixed $x, y \in V$, $x \neq 0$, we define a function $p : \mathbb{R} \to \mathbb{R}$ as follows:

$$p(\lambda) = f(\lambda x + y) = 2h(\lambda x, y) + f(\lambda x) + f(y), \qquad \lambda \in \mathbb{R}.$$

We have $p(\lambda) \geq 0$ for all $\lambda \in \mathbb{R}$. Hence for all rational λ, we have

$$0 \leq p(\lambda) = \lambda^2 f(x) + 2\lambda h(x,y) + f(y).$$

Because for fixed x and y, $p(\lambda)$ is continuous with respect to λ (quadratic form of λ for every $x \neq 0$), we have

$$0 \leq \lambda^2 f(x) + 2\lambda h(x,y) + f(y),$$

for all real $\lambda, x \neq 0$. Hence the determinant must be non-positive, that is

$$|h(x,y)|^2 \leq f(x)f(y), \text{ for arbitrary } x, y \in V.$$

(For $x = 0$ this inequality also holds.)

Substituting $h(x,y) = \frac{1}{2}(f(x+y) - f(x) - f(y))$, we obtain the thesis of the theorem. ∎

2. The functional characterization of inner product

In this section, we will assume some additional condition on the function f and we will show that then we can obtain the structure of an inner product space.

Definition 2.1. Let V be a real or complex vector space. Let f be a real or complex function on V. We say that f is real semi-continuous at zero, if for every $x \in V$ and for every sequence λ_n of real numbers tending to 0, $\lim_{n\to\infty} \lambda_n = 0$, we have

$$\lim_{n\to\infty} f(\lambda_n x) = 0.$$

Theorem 2.2. *Let V be a real vector space and $f : V \to \mathbb{R}$ a real function on V which is real semi-continuous in zero with the following properties:*

1. $f(x+y) + f(x-y) = 2f(x) + 2f(y)$ *for all* $x, y \in V$,
2. $f(x) > 0$, *for* $x \neq 0$.

Then $\langle x, y\rangle := \frac{1}{2}(f(x+y) - f(x) - f(y))$ *is an inner product on* V.

We have $f(x) = \langle x, x\rangle$ for all $x \in V$, and consequently $f(x)$ is continuous with respect to the topology defined on V by the norm $\|x\| = \sqrt{f(x)}$.

On the basis of Theorem 1.1. this implies the Schwarz inequality:

$$|\langle x, y\rangle|^2 \leq \|x\|\,\|y\|.$$

(Of course the Schwarz inequality can be deduced also from the fact that $\langle x, y\rangle$ is an inner product on V.)

The above theorem can be generalized to the case of complex vector spaces.

Theorem 2.3. *Let V be a complex vector space and $f : V \to \mathbb{R}$ a real function on V, real semi-continuous at zero, with the following properties:*

1. $f(x+y) + f(x-y) = 2f(x) + 2f(y)$ *for all* $x, y \in V$,
2. $f(x) > 0$, *for* $x \neq 0$,
3. $f(\omega x) = f(x)$, *for all* $|\omega| = 1$.

Then

$$\langle x, y\rangle = \frac{1}{4}((f(x+y) - f(x-y)) + \frac{1}{4}i((f(x+iy) - f(x-iy))$$

is an inner product on V. We have $f(x) = \langle x, x\rangle$ for all $x \in V$, and consequently $\|x\| = \sqrt{f(x)}$ is a norm on V. The Schwarz inequality also holds.

For the detailed proofs of these theorems, the reader is referred to [2].

3. The parallelogram identity in generalized K-spaces with a probability measure

In this section, we will assume some additional structure on the vector space V and we will show that also in this case, an analog of Cauchy–Schwarz inequality holds. We also show the role of orthomodularity in vector spaces. This section is based on [3].

Let V be a partially ordered real vector space. Hence V is a vector space with a partial order $\leq$ satisfying the following conditions:

1. if $x \geq y$ then $x + z \geq y + z$,
2. if $x \geq 0$ and $\lambda \in \mathbb{R}$, $\lambda \geq 0$ then $\lambda x \geq 0$, for all $x, y, z \in V$.

A binary relation C in V (we write $(x, y) \in C \equiv xCy$) is called a commutativity relation in V if it is reflexive and symmetric. If A is a subset of V then $[A]$ will denote the set of elements in V which commute with all the elements of A. Now we define $C(A) := [[A]]$. If $A = \{a\}$ for some $a \in V$, then instead of $C(\{a\})$ we write $C(a)$.

Definition 3.1. A partially ordered vector space V with a commutativity relation C is called a generalized Riesz space if the following conditions are satisfied:

1. For every $a \in V$, $C(a)$ is a Riesz subspace of V, i.e. $C(a)$ is a partially ordered subspace of V with the order $\leq$ induced from V (restricted to $C(a)$) and such that for any two elements $x, y \in C(a)$ the l.u.b. (join) $x \vee y$ and g.l.b. (meet) $x \wedge y$ with respect to the order in $C(a)$ exist and belong to $C(a)$.
2. There exists $1 \in V$, which is a unit in every Riesz space $C(a)$ (that is, $1 \geq x$ for every $x \in C(a)$).

Before we formulate the third condition, we define an element $e \in V$ to be basic if $e \in C(a)$ for some $a \in V$ implies $e \wedge (1 - e) = 0$, where the g.l.b. (meet) $\wedge$ is taken in $(C(a), \leq)$. Hence an element basic in V is basic in every Riesz space $C(a)$, to which it belongs. The set of all basic elements in V will be called the basis of V and will be denoted by $B(V)$.

Now we will formulate condition 3:

3. For every triple (e_1, e_2, e_3) of elements of $B(V)$ such that $e_i + e_j \leq 1$ for $i \leq j$, we have $e_1 + e_2 + e_3 \in B(V)$.

Theorem 3.2. *The basis $B(V)$ of a generalized Riesz space V with a commutativity relation C is an orthomodular partially ordered set (an orthoposet) with the order $\leq$ induced from V on $B(V)$ and with the orthocomplementation $e \to e'$ defined by $e' = 1 - e$ for all $e \in B(V)$.*

Let us recall that an orthoposet $(P, \leq, ')$ is a partially ordered set $(P, \leq)$ with a mapping $a \to a'$ called orthocomplementation satisfying the following conditions:

1. $a'' = a$,
2. $a \leq b$ implies $b' \leq a'$,
3. $a \vee a' = b \vee b'$ for all $a, b \in P$,
4. for all $a_1, a_2, \ldots, a_n \in P$ such that $a_i \leq a_j'$ for $i \leq j$, there exists the join in $(P, \leq)$, $a_1 \vee a_2 \vee \cdots \vee a_n \in P$,
5. $a \leq b$ implies $b = a \vee (a \vee b')'$ (orthomodularity).

Definition 3.3. A generalized Riesz space (V, C) is called a generalized K-space, if for every $a \in V, C(a)$ is a K-subspace of the space V and the basis $B(V)$ is σ-orthoposet, that is $(B(V), \leq, ')$ is an orthoposet and if $e_1, e_2, \ldots \in B(V)$, where $e_i + e_j \leq 1$ for $i \neq j$, then there exists in $(B(V), \leq)$ the l.u.b. (join) $e_1 \vee e_2 \vee \cdots \in B(V)$.

Let us recall that $C(a)$ is a K-subspace of V (K after Kantorovich), if $C(a)$ is a complete Dedekind lattice, that is, for every bounded subset of $C(a)$ there exist in $(C(a), \leq)$ g.l.b. and l.u.b. (meets and joins).

Definition 3.4. Let V be a generalized K-space and $B(V)$ a basis of V. A mapping $m : B(V) \to [0, 1]$ is called a probability measure on $B(V)$ if $m(0) = 0$, $m(1) = 1$ and $m(e_1 \vee e_2 \vee \ldots) = m(e_1) + m(e_2) + \ldots$ whenever $e_i \perp e_j$ (i.e. $e_i \leq e_j'$) for $i \neq j$. A set M of probability measures on $B(V)$ is called full if $m(e_1) \leq m(e_2)$ for all $m \in M$ implies $e_1 \leq e_2$ (the converse implication always holds by orthomodularity).

In generalized K-spaces with a basis, we can use a functional calculus: namely, for every Borel funcion $f : \mathbb{R} \to \mathbb{R}$ and for each element $a \in V$ we can construct an element $f(a) \in V$. In particular, for every $a \in V$ the element a^2 is defined and $a^2 \in V$.

Theorem 3.5. *Let (V, C) be a regular generalized K-space which basis $B(V)$ admits a full set M of probability measures. Assume that for all $a, b \in V$, the parallelogram identity holds:*

$$(a + b)^2 + (a - b)^2 = 2a^2 + 2b^2.$$

Let $m_0 \in M$ be a probability measure on $B(V)$ and m a linear functional on V extending m_0. Then for any $a, b \in V$, we have the following inequality:

$$m(a^2) \cdot m(b^2) \geq \frac{1}{4} |m([a, b])|^2, \qquad \text{where } [a, b] = (a + b)^2 - a^2 - b^2.$$

The space V is called regular, if every probability measure m_0 can be extended to a linear functional m by putting $m(a) = \int \lambda d\mu_a$, where μ_a is a measure on $B(R)$ (Borel subsets of the real line $\mathbb{R}$) defined by

$$\mu_a(E) = m_0 \circ h_a(E), \quad \text{for every } E \in B(R)$$

(h_a is a σ-homomorphism from $B(R)$ into $B(V)$ defined by the decomposition of unity induced by $a \in C(a)$).

4. The parallelogram identity in abelian groups

We can now consider only the structure of an abelian group in a vector space V - namely the structure of the additive group of vectors and we will show that even in such a general structure, the parallelogram identity has a consequence which admits a physical interpretation. Algebraic structures on abelian groups were studied in [4].

Let $(G, +)$ be an abelian group. We write $x \equiv y$ if there exists a natural number $n \neq 0$ such that $nx = ny$.

Let G_1 and G_2 be abelian groups and let $H : G_1 \times G_1 \to G_2$. We say that H is symmetric if $H(x,y) \equiv H(y,x)$ for all $x, y \in G_1$; H is homogeneous if $nH(x,y) \equiv H(nx,y) \equiv H(x,ny)$ for all $x, y \in G_1$ and $n \in Z$. H is additive if $H(x+y,z) \equiv H(x,z) + H(y,z)$ and $H(z,x+y) \equiv H(z,x) + H(z,y)$ for all $x, y \in G_1$. H is called bilinear if it is additive and homogeneous.

Theorem 4.1. *Let G_1, G_2 be abelian groups and $F : G_1 \to G_2$. We define $H : G_1 \times G_1 \to G_2$ by $H(x,y) = F(x+y) - F(x) - F(y)$. The following conditions are equivalent:*

1. *$H(-x,y) \equiv -H(x,y)$ for all $x, y \in G_1$,*
2. *H is symmetric and additive,*
3. *H is homogeneous.*

As usual, a mapping $F : G_1 \to G_2$ is called a quadratic form if there exists a symmetric bilinear $H : G_1 \times G_1 \to G_2$ such that $2F(x) \equiv H(x,x)$ for all $x \in G_1$.

Theorem 4.2. *The mapping $F : G_1 \to G_2$ is a quadratic form if and only if for all $x, y \in G_1$, we have*

$$F(x+y) + F(x-y) \equiv 2F(x) + 2F(y).$$

Application to vector spaces. Let V be a real vector space. Then V is an abelian group with respect to the addition of vectors with some additional properties (among others $x \equiv y$ if and only if $x = y$). Hence all results from the above section holds on V with $\equiv$ replaced by $=$. If V_1 and V_2 are real vector spaces, then $H : V_1 \times V_1 \to V_2$ is real homogeneous if $\lambda H(x,y) = H(\lambda x, y) = H(x, \lambda y)$ for all $\lambda \in \mathbb{R}$. H is called real bilinear, if it is additive and real homogeneous.

Applying Theorem 4.1 to the abelian group of addition in a vector space, we obtain the following corollary.

Corollary 4.2. *Let V_1, V_2 be real topological vector spaces and $F : V_1 \to V_2$ be continuous. We define $H : V_1 \times V_1 \to V_2$ by $H(x, y) = F(x + y) - F(x) - F(y)$. Then the following conditions are equivalent:*

1. *$H(-x, y) = -H(x, y)$ for all $x, y \in V_1$,*
2. *H is symmetric and additive,*
3. *H is real homogeneous.*

5. Applications to quantum mechanics

Heisenberg's uncertainty principle concerns mutual relations between dispersions of some non-compatible quantities (observables) A and B as statistical states of variables between the dispersion free state of A and the dispersion free state of B. It is known that this principle can be derived in the formalism of quantum mechanics in Hilbert space and is a consequence of Schwarz's inequality. Namely, if H is a Hilbert space with the inner product $\langle \cdot, \cdot \rangle$ and A and B are self-adjoint (hermitian) operators in H, then for any $\varphi \in H$ in the domains of AB and BA (so that $AB\varphi$ and $BA\varphi$ are well-defined) the following inequality holds:

$$\frac{1}{4}|\langle (AB - BA)\varphi, \varphi \rangle|^2 \leq \langle (A - a)^2\varphi, \varphi \rangle \langle (B - b)^2\varphi, \varphi \rangle, \tag{5.1}$$

where $a = \langle A\varphi, \varphi \rangle$, and $b = \langle B\varphi, \varphi \rangle$.

In agreement with the standard statistical interpretation of quantum mechanics, the quantity $\langle (A - a)^2\varphi, \varphi \rangle$ is the variation of A in the state φ. Moreover, if A and B are two self-adjoint operators satisfying the commutation relation

$$AB - BA = i(h/2\pi)I \tag{5.2}$$

(e.g. A corresponds to the position observable, and B to the momentum observable, I is the identity operator), then the inequality (5.1) takes the form

$$(h/4\pi)\|\varphi\| \leq \sqrt{\mathrm{var}(A, \varphi)}\sqrt{\mathrm{var}(B, \varphi)}. \tag{5.3}$$

If we introduce a new norm $\|\varphi\|_0 = \sqrt{h/2\pi}\|\varphi\|$, then (5.3) takes the form

$$\|\varphi\|_0 \leq \sqrt{\mathrm{var}(A, \varphi)}\sqrt{\mathrm{var}(B, \varphi)}. \tag{5.4}$$

For normalized states $\|\varphi\| = 1$, the inequality (5.3) takes the usual form of Heisenberg's uncertainty principle

$$\sigma(A, \varphi)\sigma(B, \varphi) \geq h/4\pi, \tag{5.5}$$

where $\sigma(A, \varphi)$, the dispersion of A in the state φ.

Now we will write Heisenberg's inequality in another form. Let us substitute in (5.1) $A_1 = A - a$ and $B_2 = B - b$. We have $AB - BA = A_1B_1 - B_1A_1$. Now we can write (5.1) in the form

$$\frac{1}{4}|\langle (A_1B_1 - B_1A_1)\varphi, \varphi\rangle|^2 \leq \langle A_1^2\varphi, \varphi\rangle\langle B_1^2\varphi, \varphi\rangle.$$

Denoting $m_2 = \langle A_1^2\varphi, \varphi\rangle = \langle A_1\varphi, A_1\varphi\rangle = \|A_1\varphi\|^2$, we see that m_2 can be interpreted as the second (statistical) moment of the observable A_1 in the state φ. Clearly m_2 is a function of A_1 and φ, so we can write $m_2 = m(A_1, \varphi)$. Because from now on we will consider only the second moment m_2, we can omit 2 in the subscript m_2. Hence for the observables A_1 and B_1, we obtain the inequality analogous to (5.4)

$$\|\varphi\|_0 \leq \sqrt{m(A_1, \varphi)}\sqrt{m(B_1, \varphi)}, \tag{5.6}$$

where $\|\varphi\|_0 = (h/4\pi)\|\varphi\|$. Clearly (5.6) is equivalent to (5.4).

Now we would like to show that this form of Heisenberg's inequality can be derived from the general structure without assuming the formalism of quantum mechanics in Hilbert space. We will show that this inequality is a consequence of some simple assumptions on the function $m(A, \varphi)$ interpreted as the second moment of A in the state φ, which is a mapping $m(A, \varphi) : V \times S \to R$ from the vector space of observables V and the vector space S of states into non-negative real numbers (by definition, second moments are always non-negative). Our assumption will be purely algebraic, namely, it will be equivalent to the parallelogram identity. Hence, this will show that Heisenberg's inequality is also a consequence of the parallelogram identity.

Before we formulate our next theorem, we will introduce some notation. Let $f(x)$ be a real function defined on a complex vector space X. Then by $f([x, y])$ we will understand a function of two variables defined on $X \times X$ by

$$f([x, y]) = \frac{1}{2}(f(x + y) - f(x) - f(y)), \quad \text{for all } x, y \in X.$$

Now we have the following theorem.

Theorem 5.1. *Let V and S be complex vector spaces and $m(A, \varphi) : V \times S \to \mathbb{R}$ a real non-negative function with the property:*

$$m([A, -B], \varphi) = -m([A, B], \varphi), \tag{5.7}$$

for all $A, B \in V$, $\varphi \in S$. Then the following inequality holds:

$$|m([A, B], \varphi)|^2 \leq m(A, \varphi)m(B, \varphi), \tag{5.8}$$

for all $A, B \in V$, $\varphi \in S$. If, in addition, there exist A, B belonging to V satisfying the following conditions:

1. $m([A,B],[\varphi,-\psi]) = -m([A,B],\ [\varphi,\psi])$, *for all* $\varphi,\psi \in S$,
2. $m([A,B],c\varphi) = m([A,B],\varphi)$, *when* $|c| = 1$,
3. $m([A,B],\varphi)$ *is real semi-continuous at zero with respect to* φ,
4. $m([A,B],\varphi) > 0$ *if* $\varphi \neq 0$,

then $\langle\varphi,\psi\rangle = m([A,B],\ [\varphi,\psi]) + im([A,B],\ [\varphi,i\psi])$ *defines an inner product on* S, $\|\varphi\|_0 = m([A,B],\varphi)$ *is a norm on* S *and the inequality* (5.8) *takes the form*

$$\|\varphi\|_0 \leq \sqrt{m(A,\varphi)}\sqrt{m(B,\varphi)}, \qquad \textit{for all } \varphi \in S. \tag{5.9}$$

This theorem follows from Theorem 2.3. The inequality (5.9) can be interpreted as an abstract form of the Heisenberg's uncertainty principle, for a detailed discussion of this problem see [5].

If we interpret V as the set of all observables of a physical system, S as the set of all states, and $m(A,\varphi)$ as the second moment of A in the state φ, then the inequality (5.9) is formally analogous to (5.4), i.e. to Heisenberg's uncertainty principle. Hence Heisenberg's uncertainty principle is implied by property (5.7). We will show that this property is equivalent to the parallelogram identity.

Namely, let $f_\varphi(A) := m(A,\varphi)$. From the definition of $[A,B]$, the property $m([A,-B],\varphi) = -m([A,B],\varphi)$ is equivalent to identity

$$f_\varphi(A+B) - f_\varphi(A) - f_\varphi(B) = -(f_\varphi(A-B) - f_\varphi(A) - f_\varphi(B)),$$

hence to the identity

$$f_\varphi(A+B) + f_\varphi(A-B) = 2f_\varphi(A) + 2f_\varphi(B),$$

which is an abstract form of the parallelogram identity. Analogously, if for fixed A and B we put $g_{A,B}(\varphi) := m([A,B],\varphi)$, then the property above is equivalent to the parallelogram identity with respect to the states

$$g_{A,B}(\varphi+\psi) + g_{A,B}(\varphi-\psi) = 2g_{A,B}(\varphi) + 2g_{A,B}(\psi).$$

Theorem 5.1 allows us to precise which properties of observables and states are sufficient to assure that the uncertainty principle holds. If $m(A,\varphi)$ is interpreted as the second moment of A in the state φ, then to obtain the inequality (5.8) we have to assume the property of the second moment expressed by the identity $m([A,-B],\varphi) = -m([A,B],\varphi)$. From the definition $m([A,B],\varphi) = \frac{1}{2}((m(A+B),\varphi) - m(A,\varphi) - m(B,\varphi))$, we see that $([A,B],\varphi)$ is equal to 0, when m is additive with respect to the addition of observables. Hence $m([A,B],\varphi)$ can be interpreted as some measure of non-additivity of m with respect to observables (expressed by a property of the second moment). From Theorem 5.1., it follows that if this non-additive part of the second moment is odd with respect to the observables then the inequality

(5.8) holds. For the same reason, if in addition there exists a pair of observables A, B such that the non-additive part of $m([A,B],[\varphi,\psi])$ is odd with respect to the addition of states then the general form (5.9) of Heisenberg's inequality holds.

Hence, Theorem 5.1 shows that the seemingly weak assumption about the oddness of the non-additive part of the second moment (equivalent to the parallelogram identity) implies the abstract form of Heisenberg's uncertainty principle. This shows once more that the parallelogram identity is not weak, in fact it is very strong, its physical consequences are very deep.

From this it follows that if we want to have Heisenberg's uncertainty principle in the formal structure of quantum mechanics, we have to assume the axiom about the oddness of the non-additive part of the second moment. As we have seen above, this is equivalent to assuming that the parallelogram identity holds for the second moment. In this sense, the parallelogram identity implies Heisenberg's uncertainty principle - rather a non-expected conclusion!

Now we will apply Theorem 5.1 to quantum mechanics based on Hilbert space. Let H be an infinite-dimensional complex Hilbert space and $B(H)$ the complex vector space of all bounded operators on H. Let for every $A \in B(H)$ and $\varphi \in H$, $m(A,\varphi) := \langle A\varphi, A\varphi\rangle$. We shall show that the condition (5.7) from Theorem 5.1 holds. We have

$$\begin{aligned} (5.10)\qquad m([A,B],\varphi) &= -(m(A+B),\varphi) - m(A,\varphi) - m(B,\varphi)) \\ &= \frac{1}{2}(\langle (A+B)\varphi,(A+B)\varphi\rangle - \langle A\varphi, A\varphi\rangle - \langle B\varphi,\varphi\rangle) \\ &= \frac{1}{2}(\langle A\varphi, B\varphi\rangle + \langle B\varphi, A\varphi\rangle). \end{aligned}$$

Consequently

$$m([A,-B],\varphi) = -m([A,B],\varphi).$$

Hence by Theorem 5.1, the inequality (5.8) holds and we have

$$|m([A,B],\varphi)|^2 \leq m(A,\varphi)m(B,\varphi).$$

Now let us go to quantum mechanics and let Q be the self-adjoint operator representing position of a particle (position observable), and P the self-adjoint operator representing the momentum of this particle (momentum observable). The operators Q and P are not bounded, but for the vectors φ in the domain of Q and P, the canonical commutation relation holds, i.e. we have

$$QP - PQ = i(h/2\pi)I. \tag{5.11}$$

Now let $A = iQ$ and $B = P$. The derivation (5.10) also holds in this case (although A and B do not have to be bounded) and we obtain

$$m([A,B],\varphi) = \frac{1}{2}(\langle iQ\varphi, P\varphi\rangle + \langle P\varphi, iQ\varphi\rangle) = \frac{1}{2}i\langle (PQ - QP)\varphi,\varphi\rangle$$
$$= \frac{1}{2}i\langle(-i(h/2\pi)I)\varphi,\varphi\rangle = (h/4\pi)\langle\varphi,\varphi\rangle = (h/4\pi)\|\varphi\|^2.$$

It is clear that the so defined $m([A,B],\varphi)$ satisfies all the conditions of Theorem 5.1. Hence $\|\varphi\|_0 = \sqrt{m([A,B]\varphi,\varphi)} = \sqrt{h/4\pi}\|\varphi\|$ is a norm in H and by Theorem 5.1 the following inequality holds

$$\|\varphi\|_0 \le \sqrt{m(A,\varphi)}\sqrt{m(B,\varphi)} \quad \text{where } A = iQ \text{ and } B = P. \tag{5.12}$$

Moreover we have

$$m(A,\varphi) = \langle A\varphi, A\varphi\rangle = \langle iQ\varphi, iQ\varphi\rangle = \langle Q^2\varphi,\varphi\rangle \text{ and}$$
$$m(B,\varphi) = \langle B\varphi, B\varphi\rangle = \langle P\varphi, P\varphi\rangle = \langle P^2\varphi,\varphi\rangle.$$

Hence, the above inequality takes the form

$$(h/4\pi)\|\varphi\|^2 \le \sqrt{\langle Q^2\varphi,\varphi\rangle}\sqrt{\langle P^2\varphi,\varphi\rangle}. \tag{5.13}$$

This inequality implies the analogous inequality for the observables $(P-p)$ and $(Q-q)$, where p and q the expected values for P and Q, respectively. Denoting $P^\wedge = P - p$ and $Q^\wedge = Q - p$ and taking into account that $\sqrt{\langle (P^\wedge)2\varphi,\varphi\rangle} = \sigma(P,\varphi)$ (analogous for Q) are the standard deviations for P i Q in the state φ, we obtain for the normed states $\|\varphi\| = 1$ the inequality

$$\sigma(Q,\varphi)\sigma(P,\varphi) \ge h/4\pi, \tag{5.14}$$

which is the original Heisenberg's uncertainty principle.

Observe that to derive Heisenberg's uncertainty principle form the parallelogram identity, we have to use complex numbers. If we put $A = Q$ and $B = P$ in the definition of $m([A,B],\varphi)$, we would obtain

$$m([A,B],\varphi) = m([Q,P],\varphi) = \frac{1}{2}\langle (QP + PQ)\varphi,\varphi\rangle, \tag{5.15}$$

hence the anti-commutator of Q and P. (Not commutator when we put $A = iQ$ and $B = Q$). The anti-commutator of Q and P is not constant (does not satisfy the cannonical commutation relation) and consequently we do not obtain Heisenberg's uncertainty relation. However, under the assumptions of Theorem 5.1, the inequality (5.8) would still hold and we aso have the following inequality for position and momentum observables

$$\sqrt{m(Q,\varphi)}\sqrt{m(P,\varphi)} \ge \frac{1}{2}\langle (QP + PQ)\varphi,\varphi\rangle, \tag{5.16}$$

which is useless for physical applications, because the right-hand side of (5.16) is not bounded from below for $\varphi \in H$, $\|\varphi\| = 1$. The question of the role of complex numbers in quantum mechanics has been discussed in [6].

References

[1] P. Jordan, J. von Neumann, *On inner products in linear metric spaces*, Annals of Math. 36 (1935), 719–723.
[2] M. J. Mączyński, *A functional characterization of inner product vector spaces*, Demonstratio Math. 16 (1983), 797–803.
[3] M. J. Mączyński, *Orthomodularity in partially ordered vector spaces*, Bull. Polish Acad. Sci. Math. 36 (1988), 299–306.
[4] S. Gudder, *Algebraic conditions for a function on an abelian groups*, Lett. Math. Phys. 3 (1979), 127–133.
[5] M. J. Mączyński, *An abstract derivation of the inequality related to Heisenberg's uncertainty principle*, Rep. Math. Phys. 21 (1985), 281–289.
[6] P. J. Lahti, M. J. Mączyński, *Heisenberg inequality and the complex field in quantum mechanics*, J. Math. Phys. 28 (1987), 1764–1769.

FACULTY OF MATHEMATICS AND INFORMATICS
UNIVERSITY OF BIAŁYSTOK
15-097 BIAŁYSTOK, POLAND
E-mail: mamacz@mini.pw.edu.pl

Received October 22, 2012; revised version November 24, 2012.

Communicated by J. Smith.