Chapter 1

Gaussian integral operators

[Draft] 30.06.07 /users/neretin/public_html/lectures/lectures.htm

In this chapter we investigate Gaussian integral operators in $L^2(\mathbb{R}^n)$. Sections 1.2– 1.6 contain a preliminary discussion of such operators. Main statements of the chapter (a description of multiplicative structure and boundedness conditions) are obtained in sections 1.8–1.10. Proofs require some linear algebra; a reader who is not sufficiently familiar with linear algebra can find an exposition of all necessary topics in Chapters 2 and 3.

1.1 Preliminaries and notation

This section is an incoherent collection of definitions that are used below in various places. A reader can consult here if he meet a unfamiliar functional-analytical or matrix term.

1.1. Functional spaces. Recall the standard notation for some functional spaces. All the details can be found in the books of Gelfand, Shilov [52], Hörmander, [69], v.1, or Reed, Simon, [145], [146].

a) $C^\infty(\mathbb{R}^n)$ is the space of all infinitely differentiable functions on $\mathbb{R}^n$. A sequence $f_j \in C^\infty(\mathbb{R}^n)$ converges to f, if for each multiindex $\alpha = (\alpha_1, \dots, \alpha_n)$, and each $r > 0$,

$$\lim_{j\to 0} \left(\max_{x:\,|x|\leqslant r} \bigl|\partial^\alpha (f_j(x) - f(x))\bigr| \right) = 0 \tag{1.1}$$

here [1] $|x| := \sqrt{\sum x_j^2}$, and

$$\partial^\alpha f(x) := \Big(\frac{\partial}{\partial x_1}\Big)^{\alpha_1} \dots \Big(\frac{\partial}{\partial x_n}\Big)^{\alpha_n} f(x_1, \dots, x_n).$$

b) $\mathcal{D}(\mathbb{R}^n)$ is the space of all compactly supported functions $f \in C^\infty(\mathbb{R}^n)$. A function f is *compactly supported* if $f = 0$ outside some compact set. A sequence f_j converges to f if

1. one can find $r > 0$ such that $f_j = 0$ for $|x| \leqslant r$
2. the condition (1.1) is satisfied.

c) $\mathcal{S}(\mathbb{R}^n)$ is the*Schwartz space* , it consists of functions f rapidly decreasing with all their derivatives, i.e., for each partial derivative $\partial^\alpha f$ and for each integer $M \geqslant 0$, we have

$$\lim_{|x|\to\infty} |x|^M \cdot |\partial^\alpha f(x)| = 0.$$

A sequence f_j converges to f if for all M and all the multiindices α,

$$\lim_{j\to\infty} \Big(\max_{|x|\to\infty} |x|^M \cdot |\partial^\alpha (f_j(x) - f(x))| \Big) = 0.$$

d) The Hilbert space $L^2(\mathbb{R}^n)$ consists of Lebesgue measurable functions f on $\mathbb{R}^n$ satisfying the condition

$$\int_{\mathbb{R}^n} |f(x)|^2 \, dx < \infty. \tag{1.2}$$

This space is endowed with the inner product

$$\langle f, g \rangle = \int_{\mathbb{R}^n} f(x)\overline{g(x)} \, dx. \tag{1.3}$$

If $f - g = 0$ outside a set of zero measure, we assume that f and g coincide as elements of $L^2(\mathbb{R}^n)$.

We recall a definition of a *pre-Hilbert* space. It is a linear space V endowed with an inner product $\langle , \cdot, \cdot \rangle$ such that

$$\langle v_1 + v_2, w \rangle = \langle v_1, w \rangle + \langle v_2, w \rangle,\ \langle v, w \rangle = \overline{\langle w, v \rangle},\ \langle \lambda v, w \rangle = \lambda \langle v, w \rangle,\ \langle v, v \rangle \geqslant 0$$

We denote by L the subspace of vectors v satisfying $\langle v, v \rangle = 0$. The quotient space V/L is a Euclidean space, we complete it with respect to the norm $\|v\| := \langle v, v \rangle^{1/2}$ and obtain the *Hilbert space associated with the pre-Hilbert space* V.

For instance, formulae (1.2) and (1.3) determine a structure of pre-Hilbert space, and L^2 is the associated Hilbert space.

[1] The ALGOL-58 symbol := means that the right-hand side is a definition of the left-hand side, the meaning of the symbol =: is similar

e) By $\mathcal{D}'(\mathbb{R}^n)$ we denote the space of all the *distributions* on $\mathbb{R}^n$. Recall that this space is the space of all the continuous linear functionals on $\mathcal{D}(\mathbb{R}^n)$. We denote pairing of a distribution χ and a function f in the two ways, namely

$$\ll \chi, f \gg = \int_{x\in\mathbb{R}^n} \chi(x) f(x).$$

A linear functional χ on $\mathcal{D}(\mathbb{R}^n)$ is continuous if

$$\ll \chi, f_j \gg \to \ll \chi, f \gg \quad \text{whenever } f_j \text{ converges to } f$$

A sequence $\chi_j \in \mathcal{D}'(\mathbb{R}^n)$ converges to χ if for each $f \in \mathcal{D}(\mathbb{R}^n)$

$$\lim_{j\to\infty} \ll \chi_j, f \gg = \ll \chi, f \gg .$$

Recall that a measurable function F on $\mathbb{R}^n$ is *locally integrable* if

$$\int_{|x|\leqslant r} |F(x)|\, dx < \infty \quad \text{for all } r$$

A locally integrable function F determines the distribution $F(x)dx$ by the formula

$$\varphi \mapsto \int_{\mathbb{R}^n} F(x)\varphi(x)\, dx, \qquad \varphi \in \mathcal{D}(\mathbb{R}^n) \tag{1.4}$$

As usual $\delta(x)$ denotes the *delta-function*

$$\ll \delta, \varphi \gg = \delta(0)$$

f) By $\mathcal{S}'(\mathbb{R}^n) \subset \mathcal{D}'(\mathbb{R}^n)$ we denote the space of all the *tempered distributions* on $\mathbb{R}^n$. By definition, this space is the space of all continuous linear functionals on the Schwartz space $\mathcal{S}(\mathbb{R}^n)$.

If a function a locally integrable function F has polynomial growth, then it determines an element of $\mathcal{S}'$ by (1.4).

e) There are obvious inclusions

$$L^2 \supset \mathcal{S} \supset \mathcal{D}, \tag{1.5}$$

$$L^2 \subset \mathcal{S}' \subset \mathcal{D}'. \tag{1.6}$$

g) Let Ξ be a countable set. We denote by $\ell_2(\Xi)$ the Hilbert space of functions on Ξ satisfying the condition $\sum_{\xi\in\Xi} |f(\xi)|^2 < \infty$. The inner product on $\ell_2(\Xi)$ is

$$\langle f, g\rangle := \sum\nolimits_{\xi\in\Xi} f(\xi)\overline{g(\xi)}$$

1.2. Integral operators and the kernel theorem. Recall some variants of L.Schwartz's *kernel theorem*, see Hörmander [69], Gelfand, Shilov[52].

A. For $K(x,y) \in \mathcal{D}'(\mathbb{R}^m \times \mathbb{R}^n)$ define a bounded operator $A : \mathcal{D}(\mathbb{R}^n) \to \mathcal{D}'(\mathbb{R}^m)$ as follows. For $f \in \mathcal{D}(\mathbb{R}^n)$ we define a linear functional Af on $\mathcal{D}(\mathbb{R}^m)$, i.e., an element of $\mathcal{D}'(\mathbb{R}^m)$, by

$$Af(g) := \ll K(x,y), g(x)f(y) \gg .$$

By the *kernel theorem*, each bounded (=continuous) linear operator $A : \mathcal{D}(\mathbb{R}^n) \to \mathcal{D}'(\mathbb{R}^m)$ has such a form. We write it symbolically as an *integral operator*

$$Af(x) = \int_{x\in\mathbb{R}^n} K(x,y)f(y), \tag{1.7}$$

where $K(x,y) \in \mathcal{D}'(\mathbb{R}^m \times \mathbb{R}^n)$.

In the same way, a distribution $K(x,y) \in \mathcal{S}'(\mathbb{R}^m \times \mathbb{R}^n)$ determines an integral operator $\mathcal{S}(\mathbb{R}^n) \to \mathcal{S}'\mathbb{R}^m)$, and each bounded operator $\mathcal{S}(\mathbb{R}^n) \to \mathcal{S}'\mathbb{R}^m)$ has such a form.

B. Let $K(x,y) \in \mathcal{S}(\mathbb{R}^n \times \mathbb{R}^m)$. Then the integral operator

$$Af(x) = \int K(x,y)f(y)\,dy$$

is a bounded operator $\mathcal{S}'(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^m)$. Indeed, for a fixed x the expression $K(x,y)$ is a function depending on y, hence we can consider its pairing with a distribution $f(y)$. Each bounded operator $\mathcal{S}'(\mathbb{R}^n) \to \mathcal{S}'(\mathbb{R}^m)$ has this form.

C. Let $A : L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^m)$ be a bounded operator. Due to inclusions (1.5), (1.6), A also is an operator $\mathcal{S}(\mathbb{R}^n) \to \mathcal{S}'(\mathbb{R}^m)$. Therefore the restriction of A to $\mathcal{S}(\mathbb{R}^n)$ is defined by formula (1.7) with a certain kernel $K \in \mathcal{D}(\mathbb{R}^m \times \mathbb{R}^n)$.

But the operator A is bounded in L^2. Hence for any sequence $f_j \in \mathcal{S}(\mathbb{R}^n)$ that converges to $f \in L^2(\mathbb{R}^n)$ in the L^2-sense we have

$$Af(x) = \lim_{j\to\infty} \int K(x,y)\, f_j(y) \qquad \text{in } L^2\text{-sense.}$$

In this sense, *all the bounded operators $L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^m)$ are integral operators.* It can be difficult to establish L^2-boundedness/unboundedness of the operator A determined by a given kernel $K(x,y)$.

D. Let $K(x,y) \in L^2(\mathbb{R}^m \times \mathbb{R}^n)$, i.e.,

$$\int_{\mathbb{R}^m\times\mathbb{R}^n} |K(x,y)|^2\, dx\, dy < \infty.$$

Then the corresponding integral operator

$$Af(x) = \int K(x,y)\, f(y)\, dy$$

is a bounded operator $L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^m)$. Moreover, A is a Hilbert–Schmidt operator (and each Hilbert-Schmidt operator can be represented in this form).

Let H, H' be Hilbert spaces. Let $A : H \to H'$ be a compact operator, let μ_j be its singular values, see below Subsection 1.7. We recall that A is a *Hilbert–Schmidt operator*, (see [52], [145]) iff $\sum \mu_j^2 < \infty$. Let $e_k \in H$, $e'_l \in H'$ be orthogonal bases. We denote by $a_{kl} = \langle Ae_k, e'_l \rangle$ the matrix elements of A. The Hilbert–Schmidt condition is equivalent to the convergence of the series

$$\operatorname{tr} A^* A = \sum_{k,l} |a_{kl}|^2 < \infty.$$

1.3. The Fourier transform. The *Fourier transform* $\mathcal{F}$ (see for instance, [146]) is the operator in space of functions on $\mathbb{R}^n$ defined by the kernel $K(x, y) = \exp\{i \sum x_k y_k\}$, i.e.,

$$\mathcal{F} f(x) = \widehat{f}(x) = (2\pi)^{-n/2} \int_{\mathbb{R}^n} f(y) \exp\Big\{ i \sum x_k y_k \Big\}\, dy \tag{1.8}$$

We recall some its properties:

1) $\mathcal{F}$ is a continuous bijection $\mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n)$

2) Therefore $\mathcal{F}$ is a continuous bijection $\mathcal{S}'(\mathbb{R}^n) \to \mathcal{S}'(\mathbb{R}^n)$

3) $\mathcal{F}$ is a unitary operator $L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$. In other words, for f, $g \in L^2(\mathbb{R}^n)$ the *Plancherel formula* holds

$$\int_{\mathbb{R}^n} f(x)\, \overline{g(x)}\, dx = \int_{\mathbb{R}^n} \mathcal{F} f(y)\, \overline{\mathcal{F} g(y)}\, dy$$

4) The inverse operator is defined by the kernel $K(x, -y) = \exp\{-i \sum x_k y_k\}$

5) Therefore $\mathcal{F}^2 f(x) = f(-x)$; hence $\mathcal{F}^4 = 1$.

6) Consider the operators

$$T(a) f(x) = f(x + a), \qquad S(b) f(x) = \exp\{i \sum b_k x_k\}$$

Then

$$\mathcal{F} T(a) = S(-a) \mathcal{F}, \qquad \mathcal{F} S(b) = T(b) \mathcal{F} \tag{1.9}$$

The main subject of this book is several kinds of Fourier-type operators, see Subsections 1.3.3, 4.2.1, 5.1.3, 7.6.2, 8.1.1, 8.4.5, 9.2.3, 9.3.1, 11.2.3, 11.2.5, 11.5.4.

1.4. Weak convergence in Hilbert spaces. Let H be a Hilbert space. A sequence $h_n \in H$ *weakly converges* to $h \in H$ if for each $\eta \in H$ we have convergence $\langle h_n, \eta \rangle \to \langle h, \eta \rangle$ (see [145]).

Problem 1.1 a) Let e_n be an orthonormal system in a Hilbert space H. The sequence e_n weakly converges to 0.

b) The weak closure of the sphere $\sum x_j^2 = 1$ is the ball $\sum x_j^2 \leqslant 1$.

Now we formulate a constructive criterion for weak convergence of vectors. First, recall that a family of vectors ξ_α in a Hilbert space (or in a topological vector space) is called *total*, if their finite linear combinations are dense in the whole space.

Theorem 1.1 *Let $\xi_\alpha \in H$ be a total family of vectors. The following two assertions are equivalent.*

A) A sequence h_n weakly converges to h.

B) For each ξ_α, the sequence $\langle h_n, \xi_\alpha\rangle$ converges to $\langle h, \xi_\alpha\rangle$ as $n \to \infty$ and $\sup_n \|h_n\| < \infty$.

In particular, we can take an orthonormal basis as a system ξ_α.

1.5. Convergences of operators in Hilbert space. For details, see [145].

a) A sequence of bounded operators A_j in a Hilbert space H *uniformly converges* to a bounded operator A if $\|A_j - A\| \to 0$ as $j \to \infty$.

b) A sequence of bounded operators A_j *strongly converges* to A if for each $h \in H$ one has $\|A_j h - Ah\| \to 0$ as $j \to \infty$.

c) A sequence of bounded operators A_j *weakly converges* to A if for each pair of vectors ξ, $\eta \in H$ the sequence $\langle A_j\xi, \eta\rangle$ converges to $\langle A\xi, \eta\rangle$.

One has the following obvious implications

$$\textit{uniform convergence} \Longrightarrow \textit{strong convergence} \Longrightarrow \textit{weak convergence}$$

Problem 1.2 Consider the shift operators in $L^2(\mathbb{R})$,

$$A(t)f(x) = f(x+t) \qquad t \in \mathbb{R}.$$

1°. The map $t \mapsto A(t)$ is uniformly discontinuous at each point, since

$$\|A(t) - A(s)\| = \sqrt{2} \qquad \text{for } t \neq s.$$

2°. The same map is strongly continuous.

3°. $A(t)$ tends weakly to 0 as $t \to \infty$. However $A(t)$ has no strong limit as $t \to \infty$.

Theorem 1.2 *Let $\xi_\alpha \in H$ be a total family of vectors. The following two conditions are equivalent.*

A) A sequence A_j weakly converges to A.

B) For each ξ_α, ξ_β the sequence $\langle A_j\xi_\alpha, \xi_\beta\rangle$ weakly converges to $\langle \xi_\alpha, \xi_\beta\rangle$ as $j \to \infty$ and $\sup_j \|A_j\| < \infty$.

1.6. Tensor products of Hilbert spaces. Let V, W be Hilbert spaces. Let e_k be an orthonormal basis in V and f_j an orthonormal basis in W. We define the *tensor product* $V \otimes W$ as a Hilbert space whose orthonormal basis is formed by symbols $e_k \otimes f_j$.

For vectors $v = \sum \alpha_k e_k \in V$ and $w = \sum \beta_j f_j \in W$, we define

$$v \otimes w := \sum_{k,j} \alpha_k \beta_j \, v_k \otimes f_j \ \in \ V \otimes W$$

Let $A : V \to \widetilde{V}$, $B : W \to \widetilde{W}$ be linear operators. Their *tensor product* is the linear operator $A \otimes B : V \otimes W \to \widetilde{V} \otimes \widetilde{W}$ determined by the condition $(A \otimes B)(v \otimes w) = Av \otimes Bw$.

If M_1, M_2 are measure spaces, then

$$L^2(M_1) \otimes L^2(M_2) \simeq L^2(M_1 \times M_2)$$

1.7. Matrix notation.

a) We shall denote by 1_n the unit $n \times n$ matrix. Sometimes we omit the index n and write merely 1. If it acts in a space V, we write 1_V.

b) Let $A = \{a_{ij}\}$ be a matrix.

— A^t is the *transposed matrix* (with matrix elements a_{ji}),

— $\overline{A}$ is the element-wise *complex conjugate* matrix; its matrix elements are $\overline{a}_{ij}$,

— $A^* = \overline{A}^t$ is the *adjoint matrix* (its matrix elements are $\overline{a}_{ji}$.)

A matrix A is

— *symmetric* if $A = A^t$;

— *skew-symmetric* if $A = -A^t$;

— *Hermitian* or *self-adjoint* if $A = A^*$;

— *anti-Hermitian* if $A = -A^*$;

We denote by $\mathrm{Mat}(p,q)$ the space of all $p \times q$ matrices, for $n \times n$ matrices we write $\mathrm{Mat}(n)$. We denote by

$$\mathrm{Symm}(n), \quad \mathrm{ASymm}(n), \quad \mathrm{Herm}(n), \quad \mathrm{AHerm}(n)$$

the spaces of symmetric, skew-symmetric, Hermitian, and anti-Hermitian $n \times n$ matrices respectively.

c) Let S be a complex matrix. Represent it in the form $S = A + iB$, where A and B are real matrices. We denote these summands by $A = \mathrm{Re}\, S$ and $B = \mathrm{Im}\, S$.

d) For a Hermitian $n \times n$-matrix R, we define the Hermitian form on $\mathbb{C}^n$ by

$$xRy^* := \begin{pmatrix} x_1 & \dots & x_n \end{pmatrix} R \begin{pmatrix} \overline{y}_1 \\ \vdots \\ \overline{y}_n \end{pmatrix} = \sum_{k,l} r_{kl} x_k \overline{y}_l, \qquad x, y \in \mathbb{C}^n.$$

e) A Hermitian matrix A is *positive definite* (notation: $A > 0$), if the corresponding Hermitian form is positive on all nonzero vectors $x \in \mathbb{C}^n$,

$$xAx^* > 0 \qquad \text{for } x \neq 0. \tag{1.10}$$

A Hermitian matrix is positive if

– all the eigenvalues are positive

– or if all the principal minors are positive (the Sylvester criterion)

We say that A is *non-negative definite* or *positive semi-definite* if we have $\geqslant$ in (1.10). In the same way we define *negative definite* and *non-positive definite* (*negative semi-definite*) matrices.

f) *Polar decomposition*

Theorem 1.3 *Each invertible operator A in an Euclidean space admits a unique decomposition ('polar decomposition') $A = US$, where U is unitary and S is positive self-adjoint.*

PROOF. The operator A^*A is self-adjoint because $(A^*A)^* = A^*A^{**} = A^*A$. It is positive because $\langle A^*Av, v\rangle = \langle Av, Av\rangle > 0$. Hence A^*A can be represented in the form $A^*A = R\Lambda R^{-1}$, where R is unitary and Λ is a diagonal matrix with positive eigenvalues. We define $S := \sqrt{A^*A} := R\sqrt{\Lambda}R^{-1}$. Assume $U := AS^{-1}$. Then $UU^* = AS^{-2}A^* = A(A^*A)^{-1}A = 1$, i.e., U is unitary. $\square$

We use the notation

$$|A| := \sqrt{A^*A}$$

g) Let V, W be Euclidean spaces, $A : V \to W$ be a linear operator. The *norm* of A is

$$\|A\| = \sup_{\|x\|_V=1} \|Ax\|_W = \sup_{\|x\|_V=1,\, \|y\|_W=1} |\langle Ax, y\rangle_W|.$$

For a self-adjoint operator $A : V \to V$ with eigenvalues λ_j,

$$\|A\| = \max |\lambda_j|$$

h) The *singular values* ν_j of a $p \times q$-matrix A are the eigenvalues of $\sqrt{A^*A}$, see [145]. The singular values are invariants of a matrix with respect to the transformations $A \mapsto U_1AU_2$, where U_1, U_2 are unitary matrices.

Each matrix admits a unique representation in the form

$$A = U_1 \begin{pmatrix} \nu_1 & 0 & \dots \\ 0 & \nu_2 & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix} U_2, \text{ where } \nu_1 \geqslant \nu_2 \geqslant \ldots \geqslant 0 \text{ and } U_1,\, U_2 \text{ are unitary}$$

the numbers ν_j coincide with the singular values.

We have

$$\|A\| = \max \nu_j$$

i) The following theorem is a corollary of formula (2.13) given below.

Theorem 1.4 *Let $g = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ be a $(p+q) \times (q+p)$-matrix with an invertible block a. Then*

$$\det g = \det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \det a \cdot \det(d - ca^{-1}b) \tag{1.11}$$

1.8. The Gauss integral.

Theorem 1.5 *Let K be a symmetric matrix, let the matrix $\operatorname{Re} K$ be positive definite. Let $b \in \mathbb{C}^n$ be a vector-row. Then*

$$\int_{\mathbb{R}^n} \exp\Bigl\{-\frac{1}{2}xKx^t + ibx^t\Bigr\}\, dx = \frac{(2\pi)^{n/2}}{\det K^{1/2}} \exp\left\{-\frac{1}{2}bK^{-1}b^t\right\}. \tag{1.12}$$

First, we explain the meaning of the factor $\det K^{1/2}$. We give the three equivalent definitions.

1) If K is real, then $\det K > 0$ and hence $(\det K)^{1/2}$ is a well-defined positive function. It has a unique holomorphic continuation on the domain $K > 0$ in $\mathbb{C}^{n(n+1)/2}$. Indeed, this domain is convex. The function $\det K$ is nonzero in this domain[2], hence the square root has no points of branching.

2) Let λ be an eigenvalue of K, let v be a corresponding eigenvector. Then

$$0 < v(\operatorname{Re} K)v^* = \operatorname{Re}(vKv^*) = \operatorname{Re}(\lambda \cdot vv^*) = (\operatorname{Re}\lambda) \cdot vv^*.$$

Thus, $\operatorname{Re}\lambda > 0$. Define $\det K^{1/2} := \prod \lambda_j^{1/2}$, where λ_j are the eigenvalues of K.

3) Let C be a contour in the domain $\operatorname{Re} z > 0$ surrounding the specter λ_j of K. Then

$$K^{1/2} := \frac{1}{2\pi i}\int_C (z-K)^{-1} z^{1/2} dz.$$

Proof of Theorem. Step 1. Let K be real, and $b = 0$. Represent K in the form $K = U\Lambda U^t$, where U is an orthogonal matrix (i.e., $UU^t = 1$), and Λ is a diagonal matrix with entries $\lambda_k > 0$,

$$\int_{\mathbb{R}^n} \exp\Bigl\{-\frac{1}{2}xKx^t\Bigr\}\,dx = \int_{\mathbb{R}^n} \exp\Bigl\{-\frac{1}{2}(xU)\Lambda(xU)^t\Bigr\}\,d(xU^t) =$$
$$= \int_{\mathbb{R}^n} \exp\Bigl\{-\frac{1}{2}x\Lambda x^t\Bigr\}\,dx = \prod_{k=1}^n \int_{\mathbb{R}} \exp\Bigl\{-\frac{1}{2}\lambda_k x_k^2\Bigr\} dx_k =$$
$$= \prod_{k=1}^n \lambda_k^{-1/2} \cdot \Bigl(\int_{\mathbb{R}} \exp\Bigl\{-\frac{1}{2}x^2\Bigr\}\,dx\Bigr)^n = (\det K)^{-1/2} \cdot \Bigl(\int_{\mathbb{R}} \exp\Bigl\{-\frac{1}{2}x^2\Bigr\}\,dx\Bigr)^n$$

Next,

$$\Bigl(\int_{\mathbb{R}} \exp\Bigl\{-\frac{1}{2}x^2\Bigr\}\,dx\Bigr)^2 = \int_{\mathbb{R}^2} \exp\Bigl\{-\frac{1}{2}(x^2+y^2)\Bigr\}\,dx\,dy =$$
$$= \int_0^{2\pi} d\varphi \int_0^\infty \exp\{-r^2/2\}\, r\,dr = 2\pi$$

This finishes our evaluation of the integral in this special case.

Step 2. Let K, $ib =: c$ be real,

$$\int_{\mathbb{R}^n} \exp\Bigl\{-\frac{1}{2}xKx^t + cx^t\Bigr\}\,dx =$$
$$= \exp\Bigl\{\frac{1}{2}cK^{-1}c^t\Bigr\} \int_{\mathbb{R}^n} \exp\Bigl\{-\frac{1}{2}(x+cK^{-1})K(x+cK^{-1})^t\Bigr\}\,dx$$

and after the substitution $y = x + cK^{-1}$ we obtain the desired statement.

Step 3. The identity (1.12) is proved if a matrix K and a vector ib are real. It remains to observe that both sides of (1.12) holomorphically depend on K, b.
$\square$

[2] Since $\operatorname{Re} xKx^t > 0$ for $x \neq 0$, we have $xK \neq 0$, i.e., the matrix K is nondegenerate

1.2 The Heisenberg group and the Weil representation of $\mathrm{Sp}(2n, \mathbb{R})$.

2.1. Representations. Let G be a group, let V be a complex linear space. A *linear representation* of the group G in the space V is a map that takes each $g \in G$ to an invertible linear operator $\rho(g)$ such that

$$\rho(g_1)\rho(g_2) = \rho(g_1 g_2) \qquad \text{for all } g_1,\ g_2 \in G. \tag{2.1}$$

This condition also implies

$$\rho(e) = 1; \qquad \rho(g^{-1}) = \rho(g)^{-1}, \qquad \text{where } e \text{ is the unit of } G.$$

We also can say that a representation is a homomorphism of G to the group $\mathrm{GL}(V)$ of all the invertible linear operators in V.

2.2. Unitary representations. A representation ρ of a topological group G in a Hilbert space H is *unitary* if

a) The operators $\rho(g)$ are unitary

b) The function $g \mapsto \rho(g)$ is *continuous with respect to the weak operator topology*[3]. In other words, for all v, $w \in H$, the *matrix elements*

$$\gamma(g) = \langle \rho(g)v, w\rangle$$

of the representation ρ are continuous functions[4] on G.

2.3. Projective representations. Let G be a group. A *projective representation* of G in a linear space V is a map that takes each $g \in G$ into an invertible operator $\rho(g)$ in V such that

$$\rho(g_1)\rho(g_2) = \gamma(g_1, g_2)\rho(g_1 g_2)$$

where $\gamma(g_1, g_2) \in \mathbb{C}^*$.

EXAMPLE. Let G be the additive group $\mathbb{R}^2$. For each $(a, b) \in \mathbb{R}^2$ define the operator $\rho(a, b)$ in $L^2(\mathbb{R})$ by

$$\rho(a, b) = f(x + a)\exp\{ibx\}$$

The operators ρ determine a unitary projective representation of $\mathbb{R}^2$. The same example is discussed in the next subsection (the group $\mathbb{R}^2$ is the quotient of the Heisenberg group by the center, see below). □

Some over definitions of representation theory are contained in Addendums B and C.

2.4. The Heisenberg group. The *Heisenberg group* is the group of operators in $L^2(\mathbb{R}^n)$ generated by the following transformations

$$T(a)f(x) = f(x + a), \qquad S(b)f(x) = f(x)\exp\{ixb^t\}, \tag{2.2}$$

[3] See 1.5 and Problem 1.2

[4] Officially, this requirement is necessary. Nevertheless, for locally compact groups measurability of matrix elements implies continuity. For this reason, usually b) holds by itself

where, a, $b \in \mathbb{R}^n$.

We wish to give a more formal definition. Let vectors v_+, v_- range in $\mathbb{R}^n$. The Heisenberg group $\mathrm{Heis}_n = \mathrm{Heis}_n(\mathbb{R})$ is the group of block $(1+n+1)\times(1+n+1)$-matrices having the structure

$$H\big[v_+ \oplus v_-; \xi\big] = \begin{pmatrix} 1 & v_+ & \xi \\ 0 & 1 & v_-^t \\ 0 & 0 & 1 \end{pmatrix}, \tag{2.3}$$

Their product is given by

$$\begin{pmatrix} 1 & v_+ & \xi \\ 0 & 1 & v_-^t \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & w_+ & \zeta \\ 0 & 1 & w_-^t \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & v_+ + w_+ & \xi+\zeta+v_+w_-^t \\ 0 & 1 & v_-^t + w_-^t \\ 0 & 0 & 1 \end{pmatrix},$$

or

$$H\big[v_+ \oplus v_-; \xi\big]H\big[w_+ \oplus w_-; \zeta\big] = H\big[(v_- + w_+) \oplus (v_- + w_+);\, \xi+\zeta+v_+w_-^t\big] \tag{2.4}$$

Therefore we really obtain a group.

Matrices $H(0\oplus 0, \zeta)$ commute with all the matrices $H(v_-, v_+, \xi)$, i.e., such matrices form the center $\mathcal{Z} \simeq \mathbb{R}$ of the Heisenberg group. The quotient group $\mathrm{Heis}_n/\mathcal{Z}$ is isomorphic to the additive group $\mathbb{R}^{2n}$ (see (2.4)).

For a matrix (2.3), we define the operator

$$U\big[v_+ \oplus v_-, \xi\big] = f(x+v_-)\exp(ixv_+^t + i\xi)$$

in the space $L^2(\mathbb{R})$. The following statement is trivial.

Theorem 2.1 *U is a unitary representation of* Heis_n

2.5. Irreducibility.

Theorem 2.2 *a) Let A be a bounded operator $L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^n)$ satisfying*

$$AU(g) = U(g)A \qquad \text{for all } g \in \mathrm{Heis}_n.$$

Then A is a scalar operator.

b) Let $A : \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n)$ be a bounded operator satisfying the same equation. Then A is a scalar operator.

Problem 2.1 Let B be a diagonal matrix with pairwise distinct entries. If $AB = BA$, then A is a diagonal matrix.

THE FIRST PROOF OF THE THEOREM. In both cases, our operator A is determined by some kernel $K \in \mathcal{S}'(\mathbb{R}^n \times \mathbb{R}^n)$, see Subsection 1.2. The kernel of $AS(b)$ is $K(x,y)\exp\{iyb^t\}$, the kernel of $S(b)A$ is $\exp\{ixb^t\}K(x,y)$. Hence

$$\begin{aligned} \exp\{ixb^t\}K(x,y) = K(x,y)\exp\{iyb^t\} \iff \\ \implies K(x,y)\big(\exp\{ixb^t\} - \exp\{iyb^t\}\big) = 0 \quad \forall b \in \mathbb{R}^n. \end{aligned} \tag{2.5}$$

Therefore $K(x,y)$ is supported by the set of (x,y) satisfying the condition

$$\exp\{ixb^t\} - \exp\{iyb^t\} = 0 \text{ for all } b$$

In other words, $\frac{1}{2\pi}(x-y)b^t \in \mathbb{Z}$. Thus $K(x,y)$ is supported by the diagonal $x=y$, and moreover

$$K(x,y) = \psi(x)\delta(x-y)$$

Therefore the operator A has the form

$$Af(x) = \psi(x)f(x).$$

Next, we write the condition

$$AT(c)f(x) = T(c)Af(x)$$

or

$$\psi(x)f(x+c) = \psi(x+c)f(x+c),$$

i.e., $\psi = \text{const}$. □

THE SECOND PROOF OF THEOREM 2.2.A. In fact we wish to prove again the following lemma that (as wee seen above) immediately implies the desired statement.

Lemma 2.3 *Let A be a bounded operator in $L^2(\mathbb{R}^n)$ commuting with all the operators $S(b)f(x) = \exp(ixb^t)$. Then A is an operator of multiplication by a function.*

PROOF OF THE LEMMA. The operator A must commute with all the linear combinations of the operators $S(b)$. Therefore it must commute with all the operators $S(\varphi)$ of the form

$$S(\varphi)f(x) := \left[\int_{\mathbb{R}^n} S(b)\varphi(b)\,db\right] f(x) = \left[\int_{\mathbb{R}^n} e^{ibx}\varphi(b)\,db\right] f(x)$$

where φ is an integrable function on $\mathbb{R}^n$.

In particular, A commutes with any operator $f(x) \mapsto r(x)f(x)$, where $r \in \mathcal{S}(\mathbb{R}^n)$ (see above 1.3). Hence,

$$A(rf) = r(Af).$$

Thus, if we know Af, then we know all the $A(rf)$ for all r.

For definiteness, assume $f(x) = \exp(-|x|^2)$. Let $g \in \mathcal{D}(\mathbb{R}^n)$ be a compactly supported smooth function. Then

$$Ag = A\left(\frac{g}{f}\cdot f\right) = \frac{g}{f}\cdot Af = \frac{Af}{f}\cdot g$$

Thus, the operator A coincides with the operator of multiplication by Af/f on the dense subspace $\mathcal{D} \subset L^2$. Since A is continuous in $L^2(\mathbb{R}^n)$, it follows that A is a multiplication by a function. □

Corollary 2.4 *The representation U is irreducible, i.e., there is no proper closed subspaces in $L^2(\mathbb{R})$ invariant with respect to all the operators $U(g)$.*

PROOF. Let $V \subset L^2(\mathbb{R})$ be a proper subspace. Since the operators $U(g)$ are unitary, it follows that the orthogonal complement $V^\perp$ to V also is invariant. The operator of orthogonal projection on V (along $V^\perp$) commutes with $U(g)$. It remains to apply Theorem 2.2. $\square$

2.6. Automorphisms of the Heisenberg group. We slightly modify the coordinates on Heis_n. Denote

$$R\big[v_+ \oplus v_-; \alpha\big] := \begin{pmatrix} 1 & v_+ & \alpha + \frac{1}{2} v_+ v_-^t \\ 0 & 1 & v_- \\ 0 & 0 & 1 \end{pmatrix} \in \mathrm{Heis}_n. \tag{2.6}$$

Then

$$R\big[v_+ \oplus v_-; \alpha\big]\, R\big[w_+ \oplus w_-; \beta\big] = R\big[(v_+ + w_+) \oplus (v_- + w_-);\ \alpha + \beta + \frac{1}{2}(v_- w_+^t - v_+ w_-^t)\big].$$

Let us change our language. Consider the real linear space

$$V_{2n} := \mathbb{R}^n \oplus \mathbb{R}^n,$$

Let us write its elements as $v = v_+ \oplus v_-$. We define a skew-symmetric bilinear form on V_{2n} by

$$\{v, w\} = v_+ w_-^t - v_- w_+^t.$$

In this notation, the space Heis_n is $V_{2n} \oplus \mathbb{R}$; the multiplication is given by

$$\big[v; \alpha\big] * \big[w; \beta\big] = \big[v + w; \alpha + \beta + \frac{1}{2}\{v, w\}\big]. \tag{2.7}$$

The operators of the representation U in new coordinates are

$$U\big[v_+ \oplus v_-, \xi\big] = f(x + v_-) \exp\big\{ixv_+^t + i\xi + \frac{i}{2} v_+ v_-^t\big\} \tag{2.8}$$

Denote by $\mathrm{Sp}(2n, \mathbb{R})$ (*the real symplectic group*) the group of all operators in V_{2n} preserving the form $\{\cdot, \cdot\}$ (for a detailed discussion of this group, see Chapter 3). We write out elements of $\mathrm{Sp}(2n, \mathbb{R})$ as block $(n + n) \times (n + n)$-matrices $g = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$.

By (2.7), it follows that the transformations

$$\sigma(g) : \ \big[v; \alpha\big] \ \mapsto \ \big[vg; \alpha\big], \qquad g \in \mathrm{Sp}(2n, \mathbb{R}),$$

are automorphisms of the Heisenberg group, i.e., they take products in Heis_n to products.

Therefore for a fixed $g \in \mathrm{Sp}(2n, \mathbb{R})$, the operators

$$U_g\big[v; \alpha\big] := U\big[vg; \alpha\big]$$

form a unitary representation of the Heisenberg group (since the composition of an automorphism and a representation is a representation).

The key fact is the equivalence of representations U_g for all g.

Theorem 2.5 *a) For each $g \in \mathrm{Sp}(2n, \mathbb{R})$, there is a unique to within a scalar factor bounded invertible operator $\mathrm{W}(g) : L^2(\mathbb{R}) \to L^2(\mathbb{R})$ such that*

$$U\big[vg; \alpha\big] = \mathrm{W}(g)^{-1} U\big[v; \alpha\big] \mathrm{W}(g) \tag{2.9}$$

for each $\big[v; \alpha\big] \in \mathrm{Heis}_n$.

b) The operators $\mathrm{W}(g)$ are unitary to within scalar factors and satisfy

$$\mathrm{W}(g_1)\mathrm{W}(g_2) = c(g_1, g_2)\mathrm{W}(g_1 g_2), \tag{2.10}$$

where $c(g_1, g_2) \in \mathbb{C}^*$. In other words, $\mathrm{W}(g)$ is a unitary projective representation of $\mathrm{Sp}(2n, \mathbb{R})$.

c) The operators $\mathrm{W}(g)$ are bounded in $\mathcal{S}(\mathbb{R}^n)$.

Remark By definition, the operators $\mathrm{W}(g)$ are defined to within scalars, therefore the appearance of a scalar factor in (2.10) is not surprising. Changing a normalization of $\mathrm{W}(g)$, we change $c(g_1, g_2)$; there is a (noncanonical) way to make $c(g_1, g_2) = \pm 1$ for all the g_1, g_2. But it is impossible to obtain $c(g_1, g_2) = 1$. For a detailed discussion, see below Sections 3.6–3.10 □

There are several names of the representation $g \mapsto \mathrm{W}(g)$, the most usual terms are *Weil representation*, *harmonic representation*, *oscillator representation*. The construction was discovered by K.O.Friedrichs near 1950, see [43].

2.7. Refinement of Theorem 2.5. [5]

Lemma 2.6 *a) Matrices of the form*

$$\begin{pmatrix} A & 0 \\ 0 & A^{t-1} \end{pmatrix}, \quad \begin{pmatrix} 1 & B \\ 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 \\ C & 1 \end{pmatrix}, \quad \text{where } B = B^t,\ C = C^t, \tag{2.11}$$

generate the symplectic group $\mathrm{Sp}(2n, \mathbb{R})$.

b) Matrices of the form

$$\begin{pmatrix} A & 0 \\ 0 & A^{t-1} \end{pmatrix}, \quad \begin{pmatrix} 1 & B \\ 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad \text{where } B = B^t,, \tag{2.12}$$

also generate the symplectic group.

[5] This theorem is completely trivial as a fact of representation theory. By the Stone–von Neumann theorem (see Addendum B), the representations U and U_g of the Heisenberg group are equivalent. By the Schur Lemma (see Addendum B), there exists a unique intertwining operator $\mathrm{W}(g)$ connecting these two irreducible representations. Also a product of intertwining operators is an intertwining operator. This implies the remaining part of the theorem.
The reader can consult with Addendum B and convince yourself, that this footnote is a complete proof. Our Subsection 2.8 is equivalent to the reference to the Schur Lemma, and Theorem 2.7 gives a constructive proof of equivalence of U and U_g avoiding a reference to the Stone–von Neumann Theorem.

PROOF. Let $g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{Sp}(2n, \mathbb{R})$. Suppose the block a is invertible. We wish to decompose g as

$$g = \begin{pmatrix} 1 & 0 \\ x & 1 \end{pmatrix} \begin{pmatrix} y & 0 \\ 0 & z \end{pmatrix} \begin{pmatrix} 1 & u \\ 0 & 1 \end{pmatrix} \quad \Longrightarrow \quad \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} y & yu \\ xy & z + xyu \end{pmatrix}.$$

We get the system of equations

$$y = a, \qquad yu = b, \qquad xy = c, \qquad z + xyu = d$$

Solving it, we obtain

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ ca^{-1} & 1 \end{pmatrix} \begin{pmatrix} a & 0 \\ 0 & d - ca^{-1}b \end{pmatrix} \begin{pmatrix} 1 & a^{-1}b \\ 0 & 1 \end{pmatrix}. \tag{2.13}$$

Problem 2.2 a) The matrices (2.11) are contained in $\mathrm{Sp}(2n, \mathbb{R})$.

b) If $g \in \mathrm{Sp}(2n, \mathbb{R})$, then the three factors in (2.13) are contained in $\mathrm{Sp}(2n, \mathbb{R})$ i.e., they have the form (2.11). See also (2.8.2).

Thus, an element g being in an open dense subset in $\mathrm{Sp}(2n, \mathbb{R})$ can be decomposed as a product of matrices (2.11). Products of such elements cover the whole group.

This proves a). The statement b) follows from a) and

$$\begin{pmatrix} 1 & 0 \\ C & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} 1 & -C \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}^{-1}.$$

Theorem 2.7 *For generators (2.12) of* $\mathrm{Sp}(2n, \mathbb{R})$*, the operators* $\mathrm{W}(\cdot)$ *are given by*

$$\mathrm{W} \begin{pmatrix} A & 0 \\ 0 & A^{t-1} \end{pmatrix} f(x) = f(xA^{-1}) |\det A|^{-1/2}, \tag{2.14}$$

$$\mathrm{W} \begin{pmatrix} 1 & B \\ 0 & 1 \end{pmatrix} f(x) = f(x) \exp\Big\{ \frac{i}{2} x B x^t \Big\}, \tag{2.15}$$

$$\mathrm{W} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} f(x) = \int_{\mathbb{R}^n} f(y) \exp\{ixy^t\}\, dy. \tag{2.16}$$

A verification of the relations (2.9) is trivial and we leave it as an exercise, for the Fourier transform this is reduced to (1.9). Due to scalar factor $|\det A|^{1/2}$ the operator (2.14) is unitary.

2.8. Proof of Theorem 2.5. Assume, that one can find two operators $\mathrm{W}(g)$, $\mathrm{W}'(g)$ satisfying (2.9). Then

$$\mathrm{W}(g)^{-1} U[v; \alpha] \mathrm{W}(g) = \mathrm{W}'(g)^{-1} U[v; \alpha] \mathrm{W}'(g)$$

Hence $\mathrm{W}(g)^{-1}\mathrm{W}'(g)$ commutes with $U[v; \alpha]$. By Theorem 2.2, it follows that $\mathrm{W}(g)^{-1}\mathrm{W}'(g)$ is a scalar. This proves uniqueness of $\mathrm{W}(g)$.

Now, let $\mathrm{W}(g_1)$, $\mathrm{W}(g_2)$ be solutions of the defining equation (2.9). Then

$$\begin{aligned} U\left[vg_1g_2;\alpha\right] &= \mathrm{W}(g_2)^{-1}U\left[vg_1;\alpha\right]\mathrm{W}(g_2) = \\ &= \mathrm{W}(g_2)^{-1}\mathrm{W}(g_1)^{-1}U\left[v;\alpha\right]\mathrm{W}(g_1)\mathrm{W}(g_2) = \\ &= \left(\mathrm{W}(g_1)\mathrm{W}(g_2)\right)^{-1}U\left[v;\alpha\right]\mathrm{W}(g_1)\mathrm{W}(g_2) \end{aligned}$$

But the last row is the defining relation for $\mathrm{W}(g_1)\mathrm{W}(g_2)$. Thus, we obtain the following Lemma.

Lemma 2.8 *If $\mathrm{W}(g_1)$, $\mathrm{W}(g_2)$ are the solutions of the defining equation (2.9) for g_1, g_2, then $\mathrm{W}(g_1)\mathrm{W}(g_2)$ is the solution of the defining equation for g_1g_2.*

By Theorem 2.7, it follows that the function $\mathrm{W}(\cdot)$ is defined on the generators of $\mathrm{Sp}(2n,\mathbb{R})$ and hence $\mathrm{W}(\cdot)$ is defined everywhere.

The identity (2.10) follows from the last lemma and the uniqueness of $\mathrm{W}(\cdot)$.

c) The operators (2.14)–(2.16) are bounded and invertible in $\mathcal{S}(\mathbb{R}^n)$. □

2.9. Reformulation. Consider the operators P_k, Q_k in $\mathcal{S}(\mathbb{R}^n)$ given by

$$P_k f(x) = i\frac{\partial}{\partial x_k}f(x), \qquad Q_k f(x) = x_k f(x).$$

They satisfy the commutation relations[6]

$$[P_k, Q_k] = i, \qquad \text{all other commutators are zero.}$$

Consider their linear combinations

$$\widehat{a}(v) = \widehat{a}(v_+ \oplus v_-) := \sum v_-^{(k)}Q_k + \sum v_+^{(k)}P_k = \sum v_-^{(k)}x_k + i\sum v_+^{(k)}\frac{\partial}{\partial x_k}.$$

Then

$$[\widehat{a}(v), \widehat{a}(w)] = i(v_+w_-^t - v_-w_+^t)\cdot 1,$$

Therefore for each $g \in \mathrm{Sp}(2n,\mathbb{R})$, we have

$$[\widehat{a}(vg), \widehat{a}(wg)] = [\widehat{a}(v), \widehat{a}(w)]$$

Theorem 2.9 *a) For each $g \in \mathrm{Sp}(2n,\mathbb{R})$ there is a unique to within a scalar bounded operator $Y(g) : \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n)$ such that*

$$\widehat{a}(vg) = Y(g)\widehat{a}(v)Y(g)^{-1} \tag{2.17}$$

b) $Y(g) = \lambda \cdot \mathrm{W}(g)$, where $\lambda \in \mathbb{C}$.

PROOF. UNIQUENESS. We watch considerations of Subsection 2.7, 2.8, for this we need the lemma given below.

Lemma 2.10 *An operator in $\mathcal{S}(\mathbb{R}^n)$ commuting with operators P_k and Q_k is a scalar operator.*

[6] Here $[A,B] := AB - BA$ is the *commutator* of operators

PROOF OF THE LEMMA. We repeat the first proof of Theorem 2.2. Our operators commute with Q_k. Instead of (2.5), we obtain the equation $K(x,y)(x-y)=0$. Hence the kernel $K(x,y)$ is a delta-function supported by the diagonal $x=y$ and we repeat the end of the proof of Theorem 2.2 literally. □

EXISTENCE. First, the operators (2.14)–(2.16) satisfy the desired commutation relations. Lemma 2.8 and all the considerations of Subsection 2.8 survive literally. □

Problem 2.3 a) Show that $\exp \widehat{a}(v) = U(v,0)$.
b)* Show a priory equivalence of (2.17) and (2.9).

1.3 Gaussian operators

3.1. Gaussian vectors. Let T be an $n \times n$ complex symmetric matrix, $\mathrm{Re}\, T > 0$. We define a *Gaussian vector* $b[T] \in L^2(\mathbb{R}^n)$ by

$$b[T](x) = \exp\Big\{-\frac{1}{2} x T x^t\Big\}.$$

By Theorem 1.5,

$$\langle b[P], b[Q]\rangle_{L^2(\mathbb{R}^n)} = (2\pi)^{n/2} \det(P + \overline{Q})^{-1/2}.$$

Vectors of the form $\lambda \cdot b[T]$, where $\lambda \in \mathbb{C}$, also are said to be Gaussian.

3.2. Gaussian distributions. By definition the set of Gaussian vectors form a cone. This cone is closed in the space $L^2(\mathbb{R}^n)$, but is not closed in the space $\mathcal{S}'(\mathbb{R}^n)$ of tempered distributions.

Proposition 3.1 *The family φ_α of Gaussian functions*

$$\varphi_\alpha(x) = \frac{1}{\sqrt{2\pi\alpha}} \exp\Big\{-\frac{x^2}{2\alpha}\Big\}$$

tends to the delta-function $\delta(x)$ as $\alpha \to 0$.

By definition of a convergence of distributions (see Subsection 1.1), this means

$$\lim_{\alpha \to +0} \frac{1}{\sqrt{2\pi\alpha}} \int_{-\infty}^{\infty} \exp\Big\{-\frac{x^2}{2\alpha}\Big\} f(x)\, dx = f(0) \qquad (3.1)$$

for each $f \in \mathcal{S}(\mathbb{R})$.

PROOF. The reason of this phenomenon is following. First,

$$\int_{\mathbb{R}} \varphi_\alpha(x)\, dx = 1 \qquad (3.2)$$

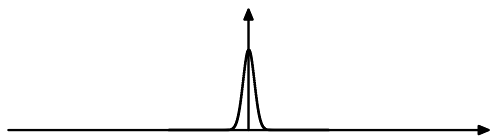

Fig. 1.1: A graph of a function $\exp(-x^2/2\alpha)/\sqrt{2\pi\alpha}$ for small α.

for all α. Secondly φ_α is concentrated near zero for small α, see Fig.1.1. This implies

$$\int_{-\infty}^{\infty} \varphi_\alpha(x) f(x)\,dx \approx \int_{-\varepsilon}^{\varepsilon} \varphi_\alpha(x) f(x)\,dx \approx$$
$$\approx f(0) \int_{-\varepsilon}^{\varepsilon} \varphi_\alpha(x)\,dx \approx f(0) \int_{-\infty}^{\infty} \varphi_\alpha(x)\,dx = f(0)$$

We wish to repeat the same arguments more formally. Fix small $\varepsilon > 0$,

$$\int_{-\infty}^{\infty} \varphi_\alpha(x)\, f(x)\,dx = \int_{|x|<\varepsilon} + \int_{|x|>\varepsilon} =$$
$$= f(0) \int_{|x|<\varepsilon} \varphi_\alpha(x)\,dx + \int_{|x|<\varepsilon} \varphi_\alpha(x)\,(f(x) - f(0))\,dx + \int_{|x|>\varepsilon} \varphi_\alpha(x)\, f(x)\,dx.$$

Now let α tends to 0 and ε remains fixed.

— The last summand tends to 0, since $\varphi_\alpha \leqslant \varphi_\alpha(\varepsilon)$ in the domain $|x| > \varepsilon$ and $\varphi_\alpha(\varepsilon) \to 0$ as $\alpha \to 0$.

— The middle summand remains small, because $|f(x) - f(0)|$ is small in the domain $|x| \leqslant \varepsilon$ and (3.2) is valid.

— the integral in the first summand tends to 1 as $\alpha \to 0$ (due to (3.2). $\square$

Consider a linear subspace $L \subset \mathbb{R}^n$. Denote by $d_L(x)$ a Lebesgue measure on L (its normalization is not important for us). Let H be a quadratic form on L and let $\operatorname{Re} H$ be non-negative definite.

Definition 3.2 *A Gaussian distribution is a tempered distribution χ of the form*

$$\ll \chi, f \gg = \lambda \int_L \exp\Big\{-\frac{1}{2}H(x)\Big\}\, f(x)\, d_L(x), \qquad f \in \mathcal{S}(\mathbb{R}^n), \qquad (3.3)$$

where $\lambda \in \mathbb{C}$.

Problem 3.1 Each Gaussian distribution is a limit of a sequence of Gaussian vectors (this is slightly modified Proposition 3.1).

Proposition 3.3 *The cone of Gaussian distributions is closed in $\mathcal{S}'(\mathbb{R}^n)$.*

Below the proposition becomes obvious (see Section 1.6 or 6.2), since we will parameterize the set of Gaussian distributions by a certain compact set in $\mathbb{C}^{n(n+1)/2}$. So now we can propose the statement as a non-obvious exercise in theory of distributions.

3.3. Gaussian operators. Consider a complex symmetric $(m+n)\times(m+n)$ matrix

$$S = \begin{pmatrix} A & B \\ B^t & C \end{pmatrix}; \qquad S = S^t. \tag{3.4}$$

In particular, the matrices A and B also are symmetric. Let $\operatorname{Re} S > 0$. A Gaussian operator is an integral operator of the form

$$B[S]f(x) = \int_{\mathbb{R}^n} \exp\Big\{-\frac{1}{2}\begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} A & B \\ B^t & C \end{pmatrix} \begin{pmatrix} x^t \\ y^t \end{pmatrix}\Big\} f(y)\, dy \tag{3.5}$$

Here f is a function on $\mathbb{R}^n$ and $B[S]f$ is a function on $\mathbb{R}^m$ (a case $n = m$ is sufficiently interesting).

More generally, let $\chi(x,y)$ be a Gaussian distribution on $\mathbb{R}^m \times \mathbb{R}^n$. We define a *Gaussian operator* as an integral operator of the form

$$Bf(x) = \int_{y\in\mathbb{R}^n} \chi(x,y) f(y).$$

By Subsection 1.2, it follows that B is a well-defined operator

$$B : \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}'(\mathbb{R}^m).$$

A detailed discussion of L^2-boundedness of Gaussian operators is given below.

3.4. Example: the identity operator. The identity operator $E : f \mapsto f$ can be represented in the form

$$Ef(x) = \int_{y\in\mathbb{R}^m} \delta(x-y) f(y). \tag{3.6}$$

By definition $\delta(x-y)$ is a Gaussian distribution on $\mathbb{R}^n \times \mathbb{R}^n$.

3.5. Example: the Fourier transform. The Fourier transform

$$\mathcal{F}f(x) = \int_{\mathbb{R}} f(y) e^{ixy}\, dy$$

is a Gaussian operator. Denote

$$Qf(x) = xf(x), \qquad Pf(x) = i\frac{d}{dx}f(x). \tag{3.7}$$

Recall that

$$\mathcal{F}^{-1}P\mathcal{F} = -Q, \qquad \mathcal{F}^{-1}Q\mathcal{F} = P. \tag{3.8}$$

Below we extend these formulae to arbitrary Gaussian operators.

3.6. Example. The heat equation. The Poisson formula. Let $t \in \mathbb{C}$ satisfies $\operatorname{Re} t \geqslant 0$. Let us consider the operators

$$V(t)f(x) := \frac{1}{\sqrt{2\pi t}} \int_{\mathbb{R}} \exp\Big\{-\frac{(x-y)^2}{2t}\Big\}\, f(y)\, dy. \tag{3.9}$$

For $f \in \mathcal{S}'(\mathbb{R})$, we denote

$$\Psi(x,t) := V(t)f(x).$$

Theorem 3.4 *a)* $V(t)V(s) = V(s+t)$.

b) For $\operatorname{Re} t > 0$,

$$\frac{\partial}{\partial t}\Psi(x,t) = \frac{1}{2}\frac{\partial^2}{\partial x^2}\Psi(x,t). \tag{3.10}$$

c) For each $f \in \mathcal{S}'(\mathbb{R})$,

$$\Psi(x,0) := \lim_{t\to+0} V(t)f(x) = f(x),$$

i.e., $V(0)$ *is the identity operator.*

d) The operators $V(is)$, *where* $s \in \mathbb{R}$, *are unitary in* $L^2(\mathbb{R})$.

e) The map $t \mapsto V(t)$ *is a weakly continuous map from the half-plane* $\operatorname{Re}(t) \geqslant 0$ *to the set of operators in* $L^2(\mathbb{R})$.

This theorem is nothing but the *Poisson formula* for solution $\Psi(x,t)$ of the *heat equation* (3.10) with a given initial data $f(x)$. We also can write

$$V(t) = \exp\Big\{\frac{t}{2}\frac{\partial^2}{\partial x^2}\Big\}.$$

Proof is trivial. The statement a) can be verified by a straightforward calculation. The statement b) is equivalent to the identity

$$\Big[\frac{\partial}{\partial t} - \frac{1}{2}\frac{\partial^2}{\partial x^2}\Big]\frac{1}{\sqrt{2\pi t}}\exp\Big\{-\frac{(x-y)^2}{2t}\Big\} = 0.$$

For a continuous f, the statements c) is (3.1). We leave a general case $f \in \mathcal{S}'(\mathbb{R})$ as an exercise.

To prove d), we write out

$$V(is)f(x) = \exp\Big\{-\frac{x^2}{2is}\Big\} \cdot (2\pi is)^{-1/2} \int_{-\infty}^{\infty} \exp\Big\{\frac{1}{is}xy\Big\}\Big[f(y)\exp\Big\{-\frac{y^2}{2is}\Big\}\Big]\, dy.$$

Thus, the operator $V(is)$ is a product of the 3 unitary operators:

— the operator of multiplication by the function $\varphi(y) := \exp\{-y^2/2is\}$; this function satisfies $|\varphi(y)| = 1$.

— the Fourier transform

— the operator of the multiplication by the same function again.

Therefore $V(s)$ is unitary.

We leave e) as an exercise in functional analysis. (one of possible proofs: in notation of Theorem 1.2, we assume $\xi_\alpha(x) = \exp(-x^2/2 + i\alpha x)$.) □

Problem 3.2 $V(is) = \mathrm{W}\begin{pmatrix} 1 & 0 \\ s & 1 \end{pmatrix}$, where $\mathrm{W}(\cdot)$ is the Weil representation defined in the previous section.

3.7. Example. The Harmonic oscillator. The Mehler formula. Let $\operatorname{Re} t \leqslant 0$. We define the operators

$$U(t) = \frac{1}{\sqrt{2\pi \sinh t}} \cdot B \begin{bmatrix} -\cosh t/\sinh t & 1/\sinh t \\ 1/\sinh t & -\cosh t/\sinh t \end{bmatrix}. \tag{3.11}$$

For $f \in \mathcal{S}'(\mathbb{R})$ we denote

$$\Psi(x,t) := U(t)f(x).$$

Theorem 3.5 *a)* $U(t)U(s) = U(s+t)$.

b) For $\operatorname{Re} t > 0$,

$$\frac{\partial}{\partial t}\Psi(x,t) = \frac{1}{2}\Big[-\frac{\partial^2}{\partial x^2} + x^2\Big]\Psi(x,t) \tag{3.12}$$

c) For each $f \in \mathcal{S}'(\mathbb{R})$,

$$\Psi(x,0) := \lim_{t \to +0} U(t)f(x) = f(x),$$

i.e., $U(0)$ *is the identity operator. Also,* $U(2\pi i k) = (-1)^k$.

d) $U(\pi i)f(x) := \lim_{\varepsilon \to +0} U(\pi i + \varepsilon)f(x) = e^{\pi i/2} f(-x)$

e) $U(\pi i/2) := \lim_{\varepsilon \to +0} U(\pi i/2 + \varepsilon) = e^{\pi i/4}\mathcal{F}$
where $\mathcal{F}$ *is the Fourier transform.*

f) The operators $U(is)$, *where* $s \in \mathbb{R}$, *are unitary in* $L^2(\mathbb{R})$.

g) The map $t \mapsto U(t)$ *is a weakly continuous map from the half-plane* $\operatorname{Re}(t) \geqslant 0$ *to the set of operators in* $L^2(\mathbb{R})$.

PROOF. We repeat the arguments from the proof of the previous theorem.□

REMARK. The partial differential equation (3.12) is the well-known *harmonic oscillator* equation from quantum mechanics. Theorem 3.5 is the famous Mehler formula (Mehler, [105], 1866, Myller-Lebedeff, [107], 1907), for a solution of the oscillator equation.

We can also say that we have evaluated the exponent

$$\exp\Big\{\frac{t}{2}\big(-\frac{d^2}{dx^2} + x^2\big)\Big\} = U(t).$$

I hope that the considerations given below make Theorem 3.5 obvious.

Proposition 3.6 *The following identities hold*

$$xU(t)f = U(t)(\cosh t \cdot y - \sinh t\frac{i\partial}{\partial y})f, \tag{3.13}$$

$$\frac{i\partial}{\partial x}U(t)f = U(t)(-\sinh t \cdot y + \cosh t\frac{i\partial}{\partial y})f. \tag{3.14}$$

Proof is a direct verification. □

These formulae extend the similar identities (3.8) for the Fourier transform.

Corollary 3.7 *For* $\varphi \in \mathbb{R}$, $\mathrm{W}\begin{pmatrix} \cos\varphi & \sin\varphi \\ -\sin\varphi & \cos\varphi \end{pmatrix} = U(i\varphi)$

3.8. Example: multiplication operator. Let $\operatorname{Re} t \geqslant 0$. We consider the operator

$$A(t)f(x) = \exp\{-tx^2/2\}f(x). \tag{3.15}$$

It is a Gaussian operator, whose kernel is the Gaussian distribution

$$\exp\{-tx^2/2\}\delta(x-y)$$

We met operators $A(it)$ in the previous section.

3.9. Example: linear change of variables. Let $t \in R$. Consider the operator

$$B(t)f(x) = f(e^t x).$$

This operator is a Gaussian operator defined by the kernel $\delta(x - e^{-t}y)$. Also

$$B(t) = \exp\{tx\frac{d}{dx}\}.$$

More generally, let $g : \mathbb{R}^n \to \mathbb{R}^n$ be an invertible linear transformation. Then the operator

$$C(g)f(x) = f(xg)$$

is Gaussian; its kernel is $\delta(xg - y)$.

Such operators also appeared in the previous section.

3.10. Example: a projector to Gaussian vector. The operator

$$\frac{1}{\sqrt{2\pi}} B\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \int_{\mathbb{R}} f(y) e^{-y^2/2} dy$$

is the orthogonal projection on the function $e^{-x^2/2}$.

More generally,

$$B\begin{bmatrix} \lambda & 0 \\ 0 & \mu \end{bmatrix} f(x) = e^{-\lambda x^2/2} \int_{\mathbb{R}} f(y) e^{-\mu y^2/2} dy$$

is a rank 1 operator.

Remark. In notation of Subsection 3.7,

$$B\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \lim_{t\to+\infty} \sqrt{2\pi \sinh t}\, U(t).$$

3.11. Example: Gaussian vectors and covectors. Consider the space $L^2(\mathbb{R}^0) = \mathbb{C}$. By definition, a Gaussian operator $L^2(\mathbb{R}^0) \to L^2(\mathbb{R}^n)$ is the form

$$s \mapsto s \cdot \exp\{-\tfrac{1}{2} x K x^t\}, \qquad s \in \mathbb{C}. \tag{3.16}$$

Gaussian operators $L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^0)$ are linear functionals of the form

$$f \mapsto \int_{\mathbb{R}^n} f(x) \exp\Bigl\{-\frac{1}{2} x K x^t\Bigr\}\, dx. \tag{3.17}$$

3.12. Example. Tensor products of Gaussian operators. Consider two Gaussian operators

$$A_1 = B\begin{bmatrix} K_1 & L_1 \\ L_1^t & M_1 \end{bmatrix} : L^2(\mathbb{R}^{n_1}) \to L^2(\mathbb{R}^{m_1}),$$
$$A_2 = B\begin{bmatrix} K_2 & L_2 \\ L_2^t & M_2 \end{bmatrix} : L^2(\mathbb{R}^{n_2}) \to L^2(\mathbb{R}^{m_2}).$$

Their tensor product (see 1.6)

$$A_1 \otimes A_2 : L^2(\mathbb{R}^{n_1+n_2}) \to L^2(\mathbb{R}^{m_1+m_2})$$

is

$$A_1 \otimes A_2 = B\begin{bmatrix} K_1 & 0 & L_1 & 0 \\ 0 & K_2 & 0 & L_2 \\ L_1^t & 0 & M_1 & 0 \\ 0 & L_2^t & 0 & M_2 \end{bmatrix}.$$

1.4 Product of Gaussian operators

Here we derive a formula (4.1)-(4.2) for a product of Gaussian operators. It looks as rather complicated, but it is clarified below in Section 1.8.

4.1. The product formula. Let a symmetric $(m+n)\times(m+n)$ matrix $S = \begin{pmatrix} A & B \\ B^t & C \end{pmatrix}$ be the same as above (3.4).

Proposition 4.1 *Let* $\operatorname{Re} S > 0$. *Then*

a) $B[S]$ *is a bounded operator* $\mathcal{S}'(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^m)$.

b) $B[S]$ *is a Hilbert-Schmidt operator* $L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^m)$.

PROOF. The first statements is a special case of the Kernel Theorem, see Subsection 1.2, the second statement is obvious, see Subsection 1.2. □

Theorem 4.2 *Consider Gaussian operators*

$$B[S_1] = B\begin{bmatrix} K & L \\ L^t & M \end{bmatrix} : \mathcal{S}'(\mathbb{R}^m) \to \mathcal{S}(\mathbb{R}^k), \qquad B[S_2] = B\begin{bmatrix} P & Q \\ Q^t & R \end{bmatrix} : \mathcal{S}'(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^m),$$

such that

$$\operatorname{Re} S_1 > 0, \qquad \operatorname{Re} S_2 > 0.$$

Then

$$B[S_1]\, B[S_2] = (2\pi)^{m/2} \det(M+P)^{-1/2}\, B[S_1 * S_2], \tag{4.1}$$

where
$$S_1 * S_2 = \begin{bmatrix} K + L(M+P)^{-1}L^t & L(M+P)^{-1}Q \\ Q^t(M+P)^{-1}L^t & R + Q^t(M+P)^{-1}Q \end{bmatrix}. \tag{4.2}$$

PROOF is a trivial calculation. We must evaluate the convolution of kernels of integral operators (see [145]), i.e., the integral

$$\int_{\mathbb{R}^m} \exp\Bigl\{-\frac{1}{2}\begin{pmatrix} x & y\end{pmatrix}\begin{pmatrix} K & L \\ L^t & M\end{pmatrix}\begin{pmatrix} x^t \\ y^t\end{pmatrix}\Bigr\} \exp\Bigl\{-\frac{1}{2}\begin{pmatrix} y & z\end{pmatrix}\begin{pmatrix} P & Q \\ Q^t & R\end{pmatrix}\begin{pmatrix} y^t \\ z^t\end{pmatrix}\Bigr\}\, dy =$$
$$=\exp\bigl\{-\frac{1}{2}xKx^t - \frac{1}{2}zRz^t\bigr\} \int \exp\Bigl\{-\frac{1}{2}y(M+P)y^t - y(L^tx^t + Qz^t)\Bigr\}\, dy.$$

Applying the Gauss integral (1.12), we obtain

$$\exp\bigl\{-\frac{1}{2}xKx^t - \frac{1}{2}zRz^t\bigr\} \cdot (2\pi)^{m/2} \det(M+P)^{-1/2} \times$$
$$\times \exp\Bigl\{-\frac{1}{2}(xL + zQ^t)(M+P)^{-1}(L^tx^t + Qz^t)\Bigr\}$$

and this is the required statement. □

Observation 4.3 *The $*$-product of matrices given by formula (4.2) is associative.*

Indeed, a product of linear operators is an associative operation.

4.2. Gaussian operators and Gaussian vectors.

Corollary 4.4 *Let $S = \begin{pmatrix} K & L \\ L^t & M\end{pmatrix}$ be an $(m+n)\times(m+n)$ matrix and $\|S\| < 1$. Let $b[P] \in L^2(\mathbb{R}^n)$ be a Gaussian vector. Then*

$$B\begin{bmatrix} K & L \\ L^t & M\end{bmatrix} b[P] = b\bigl[K + L(M+P)^{-1}L^t\bigr]. \tag{4.3}$$

PROOF. We can evaluate a Gaussian integral again, but (see Subsetion 3.11) a Gaussian vector is an operator $L^2(\mathbb{R}^0) \to L^2(\mathbb{R}^n)$, and we apply the previous theorem.

4.3. Categories. In fact, we obtain a category, whose objects are the spaces $L^2(\mathbb{R}^n)$ and morphisms are Gaussian operators.

Recall the definition of abstract *categories*. To define a category $\mathcal{K}$, we require the following data

— a set[7] $\mathrm{Ob}(\mathcal{K})$ of *objects* of $\mathcal{K}$

— for any two objects V, $W \in \mathrm{Ob}(\mathcal{K})$ a set $\mathrm{Mor}_{\mathcal{K}}(V, W)$ is defined; its elements are called *morphisms* from V to W

[7] For amateurs of mathematical logic: not a 'set', but a 'class'

— for any $P \in \mathrm{Mor}_{\mathcal{K}}(V, W)$, $Q \in \mathrm{Mor}_{\mathcal{K}}(W, Y)$ their product $QP \in \mathrm{Mor}_{\mathcal{K}}(V, Y)$ is defined; the product is associative, i.e., for any V, W, Y, $Z \in \mathrm{Ob}(\mathcal{K})$ and any $P \in \mathrm{Mor}_{\mathcal{K}}(V, W)$, $Q \in \mathrm{Mor}_{\mathcal{K}}(W, Y)$, $R \in \mathrm{Mor}_{\mathcal{K}}(Y, Z)$,

$$R(QP) = (RQ)P.$$

EXAMPLE 1. The objects are finite-dimensional linear spaces $\mathbb{C}^n$, morphisms are linear operators.

In a similar fashion, we have a category of groups and homomorphisms; a category of smooth manifolds and smooth maps etc. etc. □

EXAMPLE 2. The objects are the spaces $L^2(\mathbb{R}^n)$, and the morphisms $L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^m)$ are Gaussian operators $B[S]$ with $\mathrm{Re}\, S > 0$. □

EXAMPLE 3. The objects are nonnegative integers $n = 0, 1, 2, \ldots$, morphisms $n \to m$ are $(m+n) \times (m+n)$-matrix $S = \begin{pmatrix} K & L \\ L^t & M \end{pmatrix}$ such that $S = S^t$ and $\mathrm{Re}\, S > 0$. The product of two morphisms is given by formula (4.2). This $*$-multiplication is associative. This is obvious, since a product of linear operators is associative.

Problem* **4.1** Show the associativity of the $*$-product directly from (4.2). □

We recall some other definitions.

An *endomorphism* of an object V is an element of the semigroup $\mathrm{Mor}_{\mathcal{K}}(V, V)$. We denote this semigroup by $\mathrm{End}_{\mathcal{K}}(V)$.

A *unit* $1_V \in \mathrm{End}_{\mathcal{K}}(V)$ is an element satisfying the conditions

$$\begin{aligned} \forall W \in \mathrm{Ob}(\mathcal{K})\ \forall Q \in \mathrm{Mor}_{\mathcal{K}}(V, W) :&\quad Q \cdot 1_V = Q, \\ \forall Y \in \mathrm{Ob}(\mathcal{K})\ \forall P \in \mathrm{Mor}_{\mathcal{K}}(Y, V) :&\quad 1_V \cdot P = P \end{aligned}$$

An element $R \in \mathrm{End}_{\mathcal{K}}(V)$ is an *invertible* element or an *automorphism*, if one can find an element $S = R^{-1}$ such that $RS = SR = 1_V$. Automorphisms of V form a group $\mathrm{Aut}_{\mathcal{K}}(V)$.

Below (Section 1.8) we show that the group of invertible Gaussian operators in $L^2(\mathbb{R}^n)$ is the extended real symplectic group $\mathrm{Sp}(2n, \mathbb{R})$.

1.5 Few linear symplectic geometry

In the rest of this chapter, we explain the $*$-multiplication (4.2). We need some simple notions and constructions of linear algebra. A necessary background is contained in Chapters 2 and 3 (written for a reader that has no sufficient experience in linear algebra). In this Section we discuss structures that appear when we complexify a space endowed with a skew-symmetric bilinear form.

5.1. Standard symplectic spaces. Fix $n = 0, 1, 2, \ldots$. Denote by $\mathcal{V}_{2n}(\mathbb{R})$ the space

$$\mathcal{V}_{2n}(\mathbb{R}) := \mathcal{V}_n^+(\mathbb{R}^n) \oplus \mathcal{V}_n^-(\mathbb{R}^n) = \mathbb{R}^n \oplus \mathbb{R}^n$$

equipped with the standard skew symmetric bilinear form

$$\{v^+ \oplus v^-, w^+ \oplus w^-\} := \sum_{j=1}^n v_j^+ w_j^- - \sum_{j=1}^n v_j^- w_j^+ =: \begin{pmatrix} v^+ & v^- \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} (w^+)^t \\ (w^-)^t \end{pmatrix}. \tag{5.1}$$

Denote by $\mathcal{V}_{2n}$ the complexification[8] of the space $\mathcal{V}_{2n}(\mathbb{R})$. We identify it with $\mathbb{C}^n \oplus \mathbb{C}^n$. Denote

$$\mathcal{V}_n^+ := \mathbb{C}^n \oplus 0, \qquad \mathcal{V}_n^- = 0 \oplus \mathbb{C}^n.$$

Next, we are going to extend the form $\{\cdot,\cdot\}$ to the space $\mathcal{V}_{2n}$. There are two ways to do this.

1°. We can extend $\{\cdot,\cdot\}$ as a bilinear form. Let $v = y + iz$, $v' = y' + iz'$, where y, y', z, $z' \in \mathcal{V}_{2n}(\mathbb{R})$. We assume

$$\{v, v'\} := \{y + iz, y' + iz'\} = \{y, y'\} - \{z, z'\} + i\{y, z'\} + i\{z, y'\}. \tag{5.2}$$

This form is given by the same formula (5.1). It is more convenient to change the notation and introduce the form

$$\Lambda(v, v') := -i\{v, v'\} = \begin{pmatrix} v^+ & v^- \end{pmatrix} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \begin{pmatrix} (w^+)^t \\ (w^-)^t \end{pmatrix}. \tag{5.3}$$

2°. We can extend $\{\cdot,\cdot\}$ as a sesquilinear form (see Section 2.1)

$$\langle v, v'\rangle = \langle y + iz, y' + iz'\rangle = \{y, y'\} + \{z, z'\} - i\{y, z'\} + i\{z, y'\}. \tag{5.4}$$

This sesquilinear form satisfies the condition

$$\langle v, w\rangle = -\langle w, v\rangle$$

i.e., it is anti-Hermitian. It is more pleasant to change the notation and to define the Hermitian form

$$M(v, w) := -i\langle v, w\rangle = \begin{pmatrix} v^+ & v^- \end{pmatrix} \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \begin{pmatrix} (w^+)^* \\ (w^-)^* \end{pmatrix}.$$

Remark. The inertia indices of the form M are (n, n). Indeed $\mathcal{V}_n^\pm$ are n-dimensional isotropic subspaces (see Subsection 2.1.9).

3°. In addition, our space $\mathcal{V}_{2n}$ is equipped with the operator

$$J(x + iy) = x - iy, \qquad x, y \in \mathcal{V}_{2n}(\mathbb{R}) \tag{5.5}$$

of the complex conjugation. Obviously, this operator is an *involution*, i.e.,

$$J^2 = 1 \tag{5.6}$$

Also J is antilinear *antilinear*, i.e.,

$$J(\alpha v) = \overline{\alpha} J v, \qquad J(v + w) = Jv + Jw.$$

[8]for a definition of complexification, see below Subsection 3.2.1

The real subspace $\mathcal{V}_{2n}(\mathbb{R})$ consists of fixed points of the involution J, i.e., vectors satisfying $Jv = v$.

The operator J is compatible with our forms in the following sense

$$\Lambda(Jv, Jw) = -\overline{\Lambda(v, w)}, \qquad M(Jv, Jw) = -\overline{M(v, w)}. \tag{5.7}$$

Also

$$M(v, w) = \Lambda(v, Jw), \qquad \Lambda(v, w) = M(v, Jw). \tag{5.8}$$

5.2. Some formalization. Let $\mathcal{Y}$ be a $2n$-dimensional linear space endowed with a nondegenerate skew-symmetric bilinear form Λ, a nondegenerate Hermitian form M, and an antilinear involution J. Assume that J is compatible with these forms (i.e., 5.7) holds), and also $M(v, w) = \Lambda(v, Jw)$.

Proposition 5.1 $\mathcal{Y} \simeq \mathcal{V}_{2n}$.

Problem 5.1 Let I be a linear operator in a real finite-dimensional linear space V such that $I^2 = 1$. Let $H_\pm$ be the eigenspaces of I corresponding the eigenvalues ± 1. Then $V = H_+ \oplus H_-$.

PROOF OF PROPOSITION. Consider the real subspace $\mathcal{Y}(\mathbb{R})$ consisting of fixed point of the involution J. For $y \in \mathcal{Y}(\mathbb{R})$, we have

$$J(iy) = -i(Jy) = -iy$$

By the previous problem, $\mathcal{Y} = \mathcal{Y}(\mathbb{R}) \oplus i\mathcal{Y}(\mathbb{R})$ and we can regard $\mathcal{Y}$ as a complexification of $\mathcal{Y}(\mathbb{R})$. By (5.7), it follows that the form Λ is pure imaginary on $\mathcal{Y}(\mathbb{R})$. By (5.8), $\Lambda = M$ on $\mathcal{Y}(R)$. We define the $\mathbb{R}$-valued form $\{\cdot, \cdot\}$ on $\mathcal{Y}(\mathbb{R})$ being $\frac{1}{i}\Lambda$. We omit a verification of the remaining details. □

Our complex linear space is equipped with 3 additional structures, i.e., a skew-symmetric form, a Hermitian form, and an antilinear involution. In fact, each pair of these structures determines the third structure canonically. We wish to discuss this phenomenon.

a) Let there be given a space with a nondegenerate skew-symmetric bilinear form Λ and an antilinear involution J preserving Λ (as in 5.7). Then we introduce an Hermitian form M by (5.8) and obtain a standard symplectic space.

b) Having a nondegenerate Hermitian form M and an antilinear involution J compatible with M, we again introduce a form Λ by (5.8).

c) Finally, let the be given a nondegenerate Hermitian form M and a nondegenerate skew-symmetric form Λ. We can restore an antilinear operator K from the condition $\Lambda(v, w) = M(v, Kw)$. Incidentally, $K^2 \neq 1$. Bur pair of forms Λ, M is not a pair of forms in a general position; this imply some strange properties of these forms, see next subsection.

5.3. Geometry of the pair of forms. Remarks. Identities (5.7) and (5.8) imply several simple consequences. We formulate them here; they are used as lemmas below, in Section 1.8.

Having two forms, we have two different orthocomplements, say

$$P^{\perp\Lambda}, \qquad P^{\perp M}$$

Lemma 5.2 *For an arbitrary subspace R, we have*

$$J(R^{\perp\Lambda}) = R^{\perp M}, \qquad J(R^{\perp M}) = R^{\perp\Lambda}.$$

PROOF. By (5.7), it follows that a vector w is M-orthogonal to P iff Jw is Λ-orthogonal to P. □

Lemma 5.3 *Let P be a Λ-Lagrangian subspace. Then*
a) $P^{\perp M}$ also is Λ-Lagrangian;
b) $P^{\perp M} = JP$

PROOF. By definition, P is the Λ-orthogonal complement to itself. Hence JP is M-orthogonal complement to P. Also, for v, $v' \in P$,

$$0 = \Lambda(v, v') = -\overline{\Lambda(Jv, Jv')}$$

and hence JP is Λ-Lagrangian. □

Corollary 5.4 *Let Q be a Λ-isotropic subspace. Then JQ is Λ-isotropic and M-orthogonal to Q.*

PROOF. We embed Q to a Λ-Lagrangian subspace P, and apply the previous lemma. □

Lemma 5.5 *Let R be an M-positive subspace. Then JR is M-negative.*

Proof. Indeed, $M(Jv, Jv) = -M(v, v)$. □

Lemma 5.6 *Let R be a J-invariant subspace. Then $R^{\perp\Lambda} = R^{\perp M}$.*

PROOF. Let $w \in R$, $v \in R^{\perp\Lambda}$. Then $0 = \Lambda(v, Jw) = M(v, w)$. □

5.4. An illustration. Apparently, it can be difficult to understand this list of trivial lemmas. To convince the reader, we propose the following problem

Problem 5.2 a) Consider the plane $\mathbb{R}^2$, the operator $J := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$, and two symmetric bilinear forms

$$\Lambda(x, y) = x_1 y_1 - x_2 y_2, \qquad M(x, y) = x_1 y_1 + x_2 y_2$$

Formulate and prove analogs of the previous lemmas.

b) Do the same for $J := \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ and the bilinear forms

$$\Lambda(x, y) = x_1 y_2 - x_2 y_1, \qquad M(x, y) = x_1 y_1 + x_2 y_2$$

REMARK (It is important but it is not used in what follows). Recall that a *semi-involution* in a complex (real, quaternionic) linear space is a linear or anti-linear map J such that J^2 is a scalar operator. The effects from the previous and this subsections are produced by a (symmetric, skew-symmetric, Hermitian, anti-Hermitian) form $\Lambda(\cdot, \cdot)$ and a semi-involution compatible with Λ. Precisely, we define the second form $M(v, w) := \Lambda(v, Jw)$, and obtain a similar picture (tables of such triplets are in [120]).

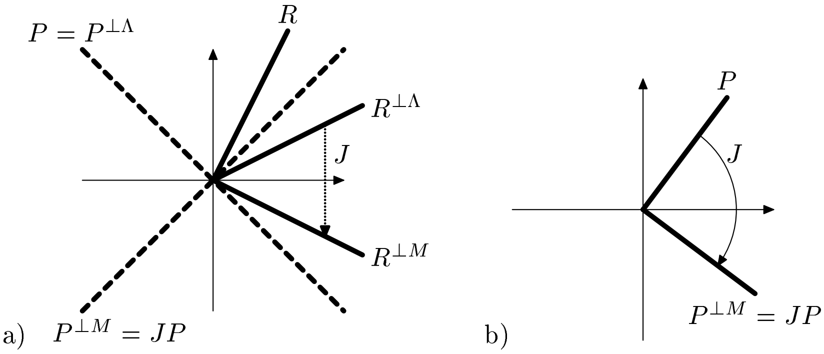

Fig. 1.2: a) To Problem 5.2.a. The involution J is the symmetry with respect to the axis Ox, the form M is the usual inner product in $\mathbb{R}^2$. The line P is Λ-isotropic. the line R is in general position.
b) To Problem 5.2.b. The semi-involution J is the rotation by 90°; M is the usual inner product. All the lines are Λ-isotropic.

1.6 Gaussian operators and Lagrangian subspaces

6.1. Symplectic spaces and differential operators. For $v = v^+ \oplus v^- \in \mathcal{V}_{2n}$ we define the differential operator on $\mathbb{R}^n$ by

$$\widehat{a}(v)f(x) := \Big(\sum v_j^- x_j + i \sum v_j^+ \frac{\partial}{\partial x_j}\Big) f(x) \tag{6.1}$$

Commutators of the operators $\widehat{a}(v)$ satisfy the following identities

$$\widehat{a}(v)^* = \widehat{a}(Jv) \tag{6.2}$$
$$[\widehat{a}(v), \widehat{a}(w)] = \Lambda(v, w) \tag{6.3}$$
$$[\widehat{a}(v), \widehat{a}(w)^*] = M(v, w) \tag{6.4}$$

6.2. Annihilators of Gaussian vectors. Let $b[T] \in \mathcal{S}(\mathbb{R}^n)$ be a Gaussian vector. Consider the set $P = P[T]$ of all $v \in \mathcal{V}_{2n}$ such that

$$\widehat{a}(v)b[T] = 0 \tag{6.5}$$

Theorem 6.1 *a) The subspace $P[T] \subset \mathcal{V}_{2n}$ satisfies the following properties*

1. $\dim P[T] = n$

2. $P[T]$ is a Lagrangian subspace with respect to the symplectic form Λ

3. the Hermitian form M is negative definite on $P[T]$

b) If $T \neq T'$, then $P[T] \neq P[T']$.

c) If $\xi \in \mathcal{D}'(\mathbb{R}^n)$ is a nonzero distribution satisfying

$$\widehat{a}(v)\xi = 0 \qquad \forall v \in P[T],$$

then $\xi = \lambda b[T]$ for some $\lambda \in \mathbb{C}^*$.

d) Each subspace $P \subset \mathcal{V}_{2n}$ satisfying properties 1-3 has the form $P[T]$ for some Gaussian vector $b[T]$.

Thus, we obtain one-to-one correspondence between the set of Gaussian vectors (defined to within a scalar) and a certain domain in the complex Lagrangian Grassmannian.

PROOF. a)

$$\widehat{a}(v)b[T](x) = \Big[\sum v_j^- x_j + i\sum v_j^+ \frac{\partial}{\partial x_j}\Big] \exp\Big\{-\frac{1}{2}\sum_{k,l} t_{kl}x_kx_l\Big\} =$$
$$= \Big[\sum_j (v_j^- - i\sum_\alpha v_\alpha^+ t_{\alpha j})x_j\Big] \exp\Big\{-\frac{1}{2}xTx^t\Big\}$$

Therefore a vector $v = v^+ \oplus v^- \in \mathcal{V}_{2n}$ is contained in $P[T]$ iff

$$v_- = iv_+T \tag{6.6}$$

In particular, $\dim P[T] = n$. Furthermore, the condition $T = T^t$ means that the subspace $P[T]$ is Lagrangian (see below Theorem 3.1.4) and $\operatorname{Re} T > 0$ means the $P[T]$ is M-negative (see below Proposition 2.3.10).

b) This is evident. For different vectors, the conditions (6.6) are different.

c) Represent the distribution ξ in the form

$$\xi = \eta \exp\{-\tfrac{1}{2}xTx^t\},$$

where $\eta \in \mathcal{D}'(\mathbb{R}^n)$. For $v \in P[T]$,

$$0 = \widehat{a}(v)\xi = \Big[\sum v_j^- x_j + i\sum v_j^+ \frac{\partial}{\partial x_j}\Big]\Big(\eta \exp\Big\{-\frac{1}{2}\sum_{k,l} t_{kl}x_kx_l\Big\}\Big) =$$
$$= \exp\Big\{-\frac{1}{2}xTx^t\Big\} i\Big(\sum v_j^+ \frac{\partial}{\partial x_j}\Big)\eta + \eta\cdot\Big[\sum_j x_j(v_j^- - i\sum_\alpha v_\alpha^+ t_{\alpha j})\Big] \exp\Big\{-\frac{1}{2}xTx^t\Big\}.$$

The second summand in the right-hand side is zero because the factor in the square brackets is zero (as we have seen in proof of a)). Hence the first summand also is 0. Therefore

$$\frac{\partial}{\partial x_j}\eta = 0 \qquad \text{for all } j.$$

Thus $\eta = \text{const}$.

d) Since P is M-negative and $\mathcal{V}^-$ is M-isotropic, we have $P \cap \mathcal{V}^- = 0$. Since $\dim P = n$, the subspace P is a graph of an operator $iT : \mathcal{V}^+ \to \mathcal{V}^-$. Since P is Lagrangian, the matrix T is symmetric (see Theorem 3.1.4). Since P is negative, we have $\operatorname{Re} T > 0$, see Proposition 2.3.10. □

There is also an another variant of the same theorem

Theorem 6.2 *a) Let χ be a nonzero Gaussian distribution on $\mathbb{R}^n$. Denote by $P(\chi)$ the set of all $v \in \mathcal{V}_{2n}$ such that*

$$\widehat{a}(v)\chi = 0.$$

The subspace $P(\chi)$ satisfies the conditions 1-2 of the previous theorem (i.e., $P(\chi)$ is a Lagrangian subspace) and also the following condition

$3'$. The Hermitian form M is negative semi-definite on $P(\chi)$.

b) If $P(\chi) = P(\chi')$, then $\chi' = s \cdot \chi$, where $s \in \mathbb{C}^*$.

c) If a distribution ξ satisfies

$$\widehat{a}(v)\xi = 0 \qquad \forall v \in P(\chi)$$

then $\xi = \mathrm{const} \cdot \chi$.

d) Each subspace in $\mathcal{V}_{2n}$ satisfying the properties 1, 2, $3'$ has the from $P(\chi)$.

Again, we obtain the one-to-one correspondence between the set of Gaussian distributions (defined to within a scalar) and the set of Lagrangian semi-negative subspaces.

REMARK. This theorem implies Proposition 3.3. about the closure of the cone of Gaussian distributions. □ □

PROOF. a), b) Without loss of generality, we can consider a Gaussian distribution on $\mathbb{R}^n$ of the form

$$\ll \chi, f \gg = \int_{\mathbb{R}^m} \exp\Bigl\{-\frac{1}{2}\sum_{k,l=1}^{m} t_{kl} x_k x_l\Bigr\} f(x_1, \dots, x_m, \dots, x_n)\, dx_1 \dots dx_m.$$

In this case,

$$x_{m+1}\chi = 0, \qquad \dots \qquad x_n \chi = 0 \tag{6.7}$$

and $P(\chi)$ consists of vectors

$$v^+ \oplus v^- = (v_1^+, \dots, v_m^+, v_{m+1}^+, \dots, v_n^+) \oplus (v_1^-, \dots, v_m^-, 0, 0, \dots, 0) \tag{6.8}$$

satisfying the condition

$$v_j^+ = i \sum_{k=1}^{m} v_k^- t_{kj}, \qquad \text{for } j \leqslant m$$

Now the statement becomes more-or-less obvious.

c) Equations (6.7) imply that χ has the form

$$\psi(x_1, \dots, x_n)\, \delta(x_{m+1}) \dots \delta(x_n),$$

where ψ is a distribution supported by $\mathbb{R}^m$. We represent ψ as

$$\eta \cdot \exp\Bigl\{-\frac{1}{2}\sum_{k,l=1}^{m} t_{k,l} x_k x_l\Bigr\}$$

and repeat the calculation given in the previous proof.

We leave d) as an exercise in linear algebra for amateurs. □

6.3. Annihilators of Gaussian operators. Now, let χ be a Gaussian distribution on $\mathbb{R}^{m+n}$. Consider the corresponding integral operator $B(\chi) : \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}'(\mathbb{R}^m)$. Consider the subspace

$$\mathcal{L}(\chi) \subset \mathcal{V}_{2n} \oplus \mathcal{V}_{2m}$$

consisting of vectors $v \oplus w$ such that

$$\widehat{a}(w)B(\chi) = B(\chi)\widehat{a}(v). \tag{6.9}$$

EXAMPLE. To simplify notation, let us assume $m = n = 1$ and consider a nonsingular Gaussian distribution

$$K(x,y) = \exp\Big\{-\frac{1}{2}\begin{pmatrix} x & y \end{pmatrix}\begin{pmatrix} p & q \\ q & r \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}\Big\}. \tag{6.10}$$

Our equations (6.9) have the form

$$\Big[w^+x + iw^-\frac{\partial}{\partial x}\Big]\int_{\mathbb{R}} K(x,y)f(y)\,dy = \int_{\mathbb{R}} K(x,y)\Big[v^+y + iv^-\frac{\partial}{\partial y}\Big]f(y)\,dy.$$

Integrating by parts in the right-hand side and equating the kernels, we obtain

$$\Big[w^+x + iw^-\frac{\partial}{\partial x}\Big]K(x,y) = \Big[v^+y - iv^-\frac{\partial}{\partial y}\Big]K(x,y).$$

Substituting the explicit expression (6.10) for $K(x,y)$, we obtain

$$\Big[w^+x - iw^-(px+qy)\Big]K(x,y) = \Big[v^+y + iv^-(qx+ry)\Big]K(x,y).$$

Collecting terms with x and y, we obtain

$$\begin{pmatrix} w^+ & v^+ \end{pmatrix} = i\cdot\begin{pmatrix} w^- & v^- \end{pmatrix}\begin{pmatrix} p & -q \\ q & -r \end{pmatrix}. \tag{6.11}$$

□

Thus we observe that the problems (6.5) and (6.9) lead to similar systems of partial differential equations and similar systems (6.6), (6.11) of linear equations. There is only a difference in signs. □

We wish to formulate our observation more precisely

Let $\chi \in \mathcal{S}'(\mathbb{R}^m \oplus \mathbb{R}^n)$ be a Gaussian distribution. Consider the subspace $P(\chi) \subset \mathcal{V}_{2m+2n}$ defined in the previous subsection. We denote its vectors by

$$(v^+ \oplus w^+) \oplus (v^- \oplus w^-) \in (\mathcal{V}_m^+ \oplus \mathcal{V}_n^+) \oplus (\mathcal{V}_m^- \oplus \mathcal{V}_n^-) = \mathcal{V}_{m+n}^+ \oplus \mathcal{V}_{n+m}^-$$

Consider the subspace $\mathcal{L}(\chi) \subset \mathcal{V}_{2m} \oplus \mathcal{V}_{2n}$ defined in this subsection. We denote its elements by

$$(w^+ \oplus w^-) \oplus (v^+ \oplus v^-) \in (\mathcal{V}_m^+ \oplus \mathcal{V}_m^-) \oplus (\mathcal{V}_n^+ \oplus \mathcal{V}_n^-) = \mathcal{V}_{2m} \oplus \mathcal{V}_{2n}$$

Theorem 6.3

$$(v^+ \oplus w^+) \oplus (v^- \oplus w^-) \in P(\chi) \Longleftrightarrow$$
$$\Longleftrightarrow (w^+ \oplus w^-) \oplus ((-v^+) \oplus v^-) \in \mathcal{L}(\chi). \quad (6.12)$$

Below we will formulate this result in more closed form.

6.4. Preliminary remark on products of two Gaussian operators. Now let us return to the product formula (4.1) for Gaussian operators.

Proposition 6.4 *Let $v \oplus w \in \mathcal{L}[S_2]$ and $w \oplus y \in \mathcal{L}[S_1]$. Then $v \oplus y \in \mathcal{L}[S_1 * S_2]$.*

PROOF.

$$B[S_1]B[S_2]\widehat{a}(v) = B[S_1]\widehat{a}(w)B[S_2] = \widehat{a}(y)B[S_1]B[S_2] \qquad \square$$

This statement claims that a product of Gaussian operators corresponds to a product of linear relations, see next two sections.

1.7 Linear relations. Imitation of basic definitions of matrix theory

7.1. Linear relations. Let V, W be linear spaces. A *linear relation* $P : V \rightrightarrows W$ is a linear subspace $P \subset V \oplus W$.

REMARK. Let $A : V \to W$ be a linear operator. Its *graph* $\mathrm{graph}(A) \subset V \oplus W$ consists of all the vectors $v \oplus Av$. By definition, $\mathrm{graph}(A)$ is a linear relation,

$$\dim \mathrm{graph}(A) = \dim V.$$

REMARK. Not all linear relations $P : V \rightrightarrows W$ are graphs of operators. There are two obstacles.

1) $\mathrm{graph}(A) \cap (0 \oplus W) = 0$. This is not a case for an arbitrary linear relation $P : V \rightrightarrows W$.

2) The projection of $\mathrm{graph}(A)$ on $V \oplus 0$ is the whole space V. Now again, this is not valid for an arbitrary linear relation. $\square$

REMARK. Denote by $\mathcal{Z}$ the set of all linear relations $P : V \rightrightarrows W$ of dimension $\dim V$. The set of all graphs of operators forms an open dense subset in $\mathcal{Z}$. $\square$

Our next purpose is to extend standard definitions concerning linear operators to linear relations.

7.2. Product of linear relations. Let $P : V \rightrightarrows W$, $Q : W \rightrightarrows Y$ be linear relations.

Informally, a product of linear relations is a product of many-valued maps. If P takes a vector v to a vector w and Q takes the vector w to a vector y, then QP takes v to y.

Precisely, the product QP is a linear relation $QP : V \to W$ consisting of all $v \oplus y \in V \oplus Y$ such that there exists $w \in W$ satisfying the condition $v \oplus w \in P$, $w \oplus y \in Q$.

Example. Assume that P is a graph of an operator $A : V \to W$ and Q is a graph of an operator $B : W \to Y$. Then QP is the graph of the operator BA. $\square$

Evidently, the product of linear relations is an associative operation.

Remark. Thus the product of linear relations is an operation extending the product of linear operators. Below we reduce the product of linear relations to matrix operations but a reduction is not obvious, see 9.4. $\square$

7.3. Some definitions. Let $P : V \rightrightarrows W$ be a linear relation.

1°. The *kernel* $\ker P$ consists of all $v \in V$ such that $v \oplus 0 \in P$. In other words,

$$\ker P = P \cap (V \oplus 0).$$

Remark. If $P = \mathrm{graph}(A)$ is a graph, then $\ker P = \ker A$. $\square$

2°. The *image* $\mathrm{im}\, P \subset W$ is the projection of P on $0 \oplus W$.

Remark. If $P = \mathrm{graph}\, A$, then $\mathrm{im}\, P = \mathrm{im}\, A$ $\square$

3°. The *domain* of definiteness $\mathrm{dom}\, P \subset V$ of P is the projection of P on $V \oplus 0$.

4°. The *indefinity* $\mathrm{indef}\, P \subset W$ of P is $P \cap (0 \oplus W)$.
Remark. Informally, $\mathrm{indef}\, P$ is the image of 0 under P. $\square$

Remark. Obviously, P is a graph of an operator $V \to W$ iff $\mathrm{indef}\, P = 0$, $\mathrm{dom}\, P = V$. If P is a graph of an operator $W \to V$, then $\ker P = 0$, $\mathrm{im}\, P = W$. $\square$

For arbitrary linear relations $P : V \rightrightarrows W$, $Q : W \rightrightarrows Y$,

$$\ker QP \supset \ker P; \qquad \mathrm{dom}\, QP \subset \mathrm{dom}\, P; , \tag{7.1}$$
$$\mathrm{indef}\, QP \supset \mathrm{indef}\, Q; \qquad \mathrm{im}\, QP \subset \mathrm{im}\, Q. \tag{7.2}$$

5°. We also define the *rank* of P,

$$\mathrm{rk}\, P = \dim P - \dim\ker P - \dim\mathrm{indef}\, P = \\ = \dim\mathrm{dom}\, P - \dim\ker P = \dim\mathrm{im}\, P - \dim\mathrm{indef}\, P.$$

This definition extends the definition of the rank of a linear operator.
Obviously (see (7.1), (7.2),

$$\mathrm{rk}\, QP \leqslant \min(\mathrm{rk}\, Q, \mathrm{rk}\, P).$$

6°. The last definition of this series is the *pseudoinverse* linear relation $P^{\square}$. It is the same subspace $P \subset V \oplus W$ considered as a linear relation $W \rightrightarrows V$.

Remark. If $A : V \to W$ is an *invertible* linear operator, then

$$(\mathrm{graph}\, A)^{\square} = \mathrm{graph}\, A^{-1}$$

Otherwise, incidentally, $PP^\square \ne 1$. $\square$

Evidently,

$$P^\square Q^\square = (QP)^\square \tag{7.3}$$

REMARK. The usual definition of a linear operator in functional analysis can be formulated as: a linear operator is a linear relation P such that $\operatorname{indef} P = 0$. $\square$

7.4. Dimension of product.

Theorem 7.1 *Consider linear relations $P : V \rightrightarrows W$, $Q : W \rightrightarrows Y$.*

a) Assume

$$\operatorname{indef} P \cap \ker Q = 0;, \tag{7.4}$$

$$\operatorname{im} P + \operatorname{dom} Q = W. \tag{7.5}$$

Then

$$\dim QP = \dim Q + \dim P - \dim W. \tag{7.6}$$

b) The product of linear relations is a continuous operation on the domain determined by conditions (7.4), (7.5).

Problem 7.1 a) The group $\mathrm{GL}(V) \times \mathrm{GL}(W) \times \mathrm{GL}(Y)$ acts in a natural way on the space of pairs (P, Q) of linear relations. Find canonical forms of such pairs under this action.

b) Prove the theorem.

We prefer another (more complicated) way in the next subsection.

7.5. Reformulation of the definition of product of linear relations. Consider the space $V \oplus W \oplus W \oplus Y$ and the following subspaces

– Ω consisting of vectors $v \oplus w \oplus w \oplus y$

– $\Xi \subset \Omega$ consisting of vectors $0 \oplus w \oplus w \oplus 0$.

Consider the natural projection $\pi : \Omega \to V \oplus Y$ along Ξ.

Lemma 7.2 *Let $P : V \to W$, $Q : W \to Y$ be linear relations. Then $QP \subset V \oplus Y$ is the image of $(P \oplus Q) \cap \Omega$ under the projection π.*

We leave a proof as an exercise.

PROOF OF THEOREM 7.1. The condition (7.5) implies

$$\Omega + (P \oplus Q) = V \oplus W \oplus W \oplus Y$$

For arbitrary two subspaces, L, M,

$$\dim(L + M) + \dim(L \cap M) = \dim L + \dim M$$

Applying this formula to $L = \Omega$, $M = P \oplus Q$, we obtain that $\dim \Omega \cap (P \oplus Q)$ is (7.6).

The condition (7.4) implies $(P \oplus Q) \cap \Xi = 0$ and hence the projection π is injective on $\Omega \cap (P+Q)$. This finishes proof of a) and also proves b). □

REMARK. Our proof gives the following formula which is valid for all P, Q:

$$\dim QP = \dim Q + \dim P - \dim\bigl[\operatorname{indef} P \cap \ker Q\bigr] - \dim\bigl[\operatorname{im} P + \operatorname{dom} Q\bigr]. \tag{7.7}$$

7.6. Action of linear relations on Grassmannians. Let $P : V \rightrightarrows W$ be a linear relation, $T \subset V$ a subspace. We define the subspace $PT \subset V$ as the set of all $w \in W$ such that there exists $v \in T$ satisfying $v \oplus w \in P$.

This operation is a special case of the product of linear relations. Indeed, we can regard T as a linear relation $T : 0 \rightrightarrows V$, then PT coincides with the product of P and T.

In particular, formula (7.7) implies

$$\dim PT = \dim P + \dim T - \dim\bigl[T \cap \ker P\bigr] - \dim\bigl[T + \operatorname{dom} P\bigr] \tag{7.8}$$

7.7. Isotropic linear relations and unitary operators. Let V be a Euclidean space endowed with an inner product $H(\cdot,\cdot)$. Consider the following indefinite Hermitian form on the space $V \oplus V$:

$$H^{\ominus}_{V\oplus V}(v \oplus w, v' \oplus w') = H_V(v, v') - H_V(w, w')$$

Theorem 7.3 *The following conditions are equivalent*

– $g : V \to V$ *is a unitary operator*

– $\operatorname{graph} g$ *is a maximal isotropic*[9] *subspace in* $V \oplus V$

PROOF. Let g be a unitary operator. For two points $v \oplus gv$, $v' \oplus gv'$ in $\operatorname{graph} g$,

$$H^{\ominus}(v \oplus gv, v' \oplus gv') = H_V(v, v') - H_V(gv, gv') \tag{7.9}$$

Since g is unitary, the right-hand side is zero.

Conversely, let $Z \subset V \oplus V$ be a maximal isotropic subspace. The form $H^{\ominus}$ is zero on Z and is negative definite on $0 \oplus V$. Hence $Z \cap [0 \oplus V] = 0$. Since $\dim Z = \dim V$, the subspace Z is a graph of an operator $V \to V$. Since Z is isotropic, the left hand-side of (7.9) is zero. Hence g is unitary. □

Problem 7.2 Now, consider a linear space V equipped with a skew-symmetric bilinear form $H^{\ominus}(\cdot,\cdot)$. Repeating the same considerations, show that *an operator* $g : V \to V$ *is symplectic iff its graph is a Lagrangian subspace.*

7.8. Isotropic linear relations and self-adjoint operators. Let V be a finite-dimensional complex Euclidean space, let $H(\cdot,\cdot)$ be the inner product. Define the indefinite Hermitian form $H^{\diagdown}$ on $V \oplus V$ by

$$H^{\diagdown}(v \oplus w, v' \oplus w') = i\Bigl(H(v, w') - H(w, v')\Bigr)$$

(we write i to obtain a Hermitian form, otherwise this expression is anti-Hermitian).

[9] see Subsection 2.1.9

Theorem 7.4 *a) For an operator $A : V \to V$,*

$$\mathrm{graph}(A)^{\perp} = \mathrm{graph}(A^*)$$

b) A is self-adjoint iff $\mathrm{graph}(A)$ *is a maximal isotropic subspace*

PROOF. Let A, B be operators, let $v \oplus Av$, $w \oplus Bw$ be points of their graphs,

$$H^{\diagdown}(v \oplus Av, w \oplus Bw) = i\big(H(v, Bw) - H(Av, w)\big)$$

If $A = B^*$, then $\mathrm{graph}(A)$ is orthogonal to $\mathrm{graph}(B)$. □

REMARK. However, there is no one-to-one correspondence between self-adjoint operators and maximal isotropic subspaces P. It can happened that $\mathrm{indef}\, P \neq 0$ and in this case P is not a graph of an operator. □

7.9. Digression. The Cayley transform. Recall that the *Cayley transform*[10] is given by

$$iA = \frac{1-g}{1+g}, \qquad \Longleftrightarrow \qquad g = \frac{1-iA}{1+iA}$$

Theorem 7.5 *A matrix g is unitary iff A is self-adjoint.*

PROOF. Let A be self-adjoint. Then

$$gg^* = \frac{1-iA}{1+iA} \cdot \frac{1+iA}{1-iA} = 1$$

If g is unitary, then

$$A - A^* = \frac{1}{i}\frac{1-g}{1+g} + \frac{1}{i}\frac{1-g^{-1}}{1+g^{-1}} = \frac{1}{i}\Big(\frac{1-g}{1+g} + \frac{g-1}{g+1}\Big) = 0 \qquad \square$$

Previous considerations offer another explanation of this theorem. In Subsection 7.7 we constructed a bijection between the unitary group $\mathrm{U}(n)$ and the Grassmannian of maximal isotropic subspaces. In Subsection 7.8 we obtained a correspondence between the same Grassmannian and the space $\mathrm{Herm}(n)$ of self-adjoint operators. The Cayley transform $\mathrm{U}(n) \leftrightarrow \mathrm{Herm}(n)$ is the composed correspondence.

7.10. Contractive operators and linear relations.

Theorem 7.6 *The following conditions for a linear operator g are equivalent*

— $\|g\| \leqslant 1$

— *the form $H^{\ominus}$ is non-negative on* $\mathrm{graph}(g)$

[10] Below we use its another normalization

PROOF. First, let us verify '$\Longleftarrow$',

$$H^{\ominus}(v \oplus gv, v \oplus gv) = \|v\|^2 - \|gv\|^2 \tag{7.10}$$

If $\|g\| \leqslant 1$, then this expression is positive.

Now, let the form $H^{\ominus}$ be positive on graph(g). Since the form is negative on $0 \oplus V$, the intersection graph$(g) \cap (0 \oplus H)$ is zero. Hence g is a graph of an operator. Since (7.10) is non-negative for all v, we have $\|v\| \leqslant 1$. □

7.11. Digression. Normal forms of linear relations. Let $P : V \rightrightarrows V$ be a linear relation in a complex linear space V. We are going to formulate an analog of Jordan normal form for a linear relation. In other words, we wish to classify linear relations $P : V \rightrightarrows V$ with respect to the equivalence

$$P \sim g^{-1}Pg, \qquad g \in \mathrm{GL}(V)$$

We say that a linear relation P is *decomposable* if there is a decomposition $V = V_1 \oplus V_2$ and linear relations $P_1 : V_1 \rightrightarrows V_1$, $P_2 : V_2 \rightrightarrows V_2$ such that $P = P_1 \oplus P_2$. Classification of linear relations is reduced to classification of indecomposable relations.

Theorem 7.7 *There are the following five types of indecomposable linear relations*

Let $e_1, \dots, e_k$ be a basis in a linear space V. Below we write bases in linear relations $P : V \rightrightarrows V$.

1^*. $P : V \rightrightarrows V$ is a graph of a Jordan box

$$J_n(\lambda) = \begin{pmatrix} \lambda & 1 & \dots & 0 \\ 0 & \lambda & \ddots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda \end{pmatrix}$$

with $\lambda \neq 0$. In other words, we get a linear relation with a basis

$$e_1 \oplus (\lambda e_1 + e_2),\ e_2 \oplus (\lambda e_2 + e_3), \dots, e_n \oplus \lambda e_n$$

2^*. The graph of the Jordan box $J(0)$. Of course, this is nothing but a variant of a). A basis of the linear relation is

$$e_1 \oplus e_2,\ e_2 \oplus e_3,\ e_3 \oplus e_4, \dots, e_n \oplus 0$$

3^*. The linear relation $J(0)^{\square}$ pseudoinverse for the Jordan box $J(0)$. Its basis is

$$e_2 \oplus e_1,\ e_3 \oplus e_2,\ \dots, e_n \oplus e_{n-1},\ 0 \oplus e_n$$

4^*. The $(n-1)$-dimensional linear relation spanned by

$$e_1 \oplus e_2,\ e_2 \oplus e_3, \dots e_{n-1} \oplus e_n$$

5^*. The $(n+1)$-dimensional linear relation spanned by

$$e_1 \oplus 0,\ e_2 \oplus e_1,\ e_3 \oplus e_2, \dots e_n \oplus e_{n-1}, 0 \oplus e_n$$

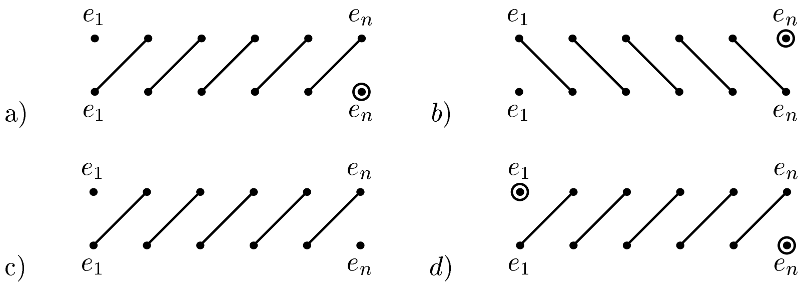

Fig. 1.3: To Subsection 7.11. Indecomposable linear relations $P : V \rightrightarrows V$. Lower fat points are basis elements $e_k \oplus 0$ of the first summand, upper fat points are basis elements $0 \oplus e_k$ of the second summand. If $e_k \oplus e_l \in P$, then we connect e_k, e_l with a segment. If $e_k \oplus 0 \in P$ or $0 \oplus e_l \in P$, then we draw a circle.
a) the Jordan box $J_n(0)$
b) the pseudo-inverse $J_n(0)^{\square}$ for Jordan box.
c) the linear relation of dimension $(n-1)$.
d) the linear relation of dimension $(n+1)$.

Problem* 7.3 Prove the theorem in the following way. For a linear relation $P \subset V \oplus V$ consider two projection operators

$$A: \ P \to V \oplus 0, \qquad B: \ P \to 0 \oplus V$$

We get a pair of operators $P \to V$. Next, apply the Kronecker theorem (see [47]) on pencils of linear operators.

Problem* 7.4 Fix $k \neq \dim V$. Show that the group $\mathrm{GL}(V)$ acts with an open dense orbit on the set of all linear relations $V \rightrightarrows V$ having dimension $= k$.

1.8 The symplectic category

Here we discuss the algebraic structure formed by Gaussian operators.

Recall that standard symplectic spaces $\mathcal{V}_{2n}$ are defined above in 5.1. Recall that they are equipped with two forms, precisely the symplectic form Λ and Hermitian form M. Also they are equipped with the operator J of complex conjugation.

Recall that the spaces $\mathcal{V}_{2n}$ parameterize certain families of partial differential operators in $\mathcal{S}(\mathbb{R}^n)$, see 6.1.

8.1. Contractive Lagrangian relations. Consider two objects $\mathcal{V}_{2n}$ and $\mathcal{V}_{2m}$ of our category. Endow their direct sum $\mathcal{V}_{2n} \oplus \mathcal{V}_{2m}$ with two forms, namely the skew-symmetric bilinear form $\Lambda^{\ominus}$ given by

$$\Lambda^{\ominus}(v \oplus w, v' \oplus w') = \Lambda(v, v') - \Lambda(w, w') \tag{8.1}$$

and the indefinite Hermitian form

$$M^{\ominus}(v \oplus w, v' \oplus w') = M(v, v') - M(w, w') \tag{8.2}$$

In addition, we define the canonical operator J of complex conjugation in $\mathcal{V}_{2n} \oplus \mathcal{V}_{2m}$,

$$J_{\mathcal{V}_{2n}\oplus\mathcal{V}_{2m}} = J_{\mathcal{V}_{2n}} \oplus J_{\mathcal{V}_{2m}} \qquad \square$$

We say that a linear relation $P : \mathcal{V}_{2n} \rightrightarrows \mathcal{V}_{2m}$ is *contractive Lagrangian* if

1^*. P is $\Lambda^{\ominus}$-Lagrangian; in particular, $\dim P = m + n$.

2^*. the form $M^{\ominus}$ is non-negative definite on P

Theorem 8.1 *a) For each contractive Lagrangian linear relation $P : \mathcal{V}_{2n} \rightrightarrows \mathcal{V}_{2m}$ there is a unique to within a scalar factor nonzero linear operator*

$$\mathcal{B} = \mathrm{Be}(P) : \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}'(\mathbb{R}^m)$$

such that

$$\widehat{a}(w)\mathcal{B} = \mathcal{B}\widehat{a}(v) \qquad \text{for all } v \oplus w \in P. \tag{8.3}$$

b) The operator $\mathrm{Be}(P)$ *is a Gaussian operator.*

c) If $P \neq P'$, then $\mathrm{Be}(P) \neq \mathrm{Be}(P')$.

d) Each Gaussian operator has the form $\mathrm{Be}(P)$ *for some contractive linear relation*

Thus, we have one-to-one correspondence between set of Lagrangian linear relations and Gaussian operators defined to within a scalar.

Proof. This is a paraphrasing of Theorems 6.2, 6.3. $\square$

Remark. We use double notation for the Gaussian operators, namely $B[S]$ and $\mathrm{Be}(P)$. In the first case, S is a matrix; in the second case, P is a linear relation.

8.2. Boundedness conditions. It is easy to produce examples of Gaussian operators that are not bounded as operators $L^2 \to L^2$ and $\mathcal{S} \to \mathcal{S}$.

Problem 8.1 Examine for boundedness in $L^2(\mathbb{R})$ and $\mathcal{S}(\mathbb{R})$ the operators with the following kernels

$$\delta(x)\cdot\delta(y), \qquad \delta(x), \qquad \delta(y), \qquad \exp(ix^2+iy^2), \qquad \exp(-x^2-iy^2)$$

Nevertheless, the set of all unbounded Gaussian operators is small. They form a submanifold lying on the boundary of the 'ball' of all the Gaussian operators.

Theorem 8.2 *Let $P : \mathcal{V}_{2n} \rightrightarrows \mathcal{V}_{2m}$ be a contractive Lagrangian linear relation.*

a) The operator $\mathrm{Be}(P)$ *is a bounded operator $L^2(\mathbb{R}^{2n}) \to L^2(\mathbb{R}^{2m})$ iff the subspace* $\mathrm{indef}\, P \subset \mathcal{V}_{2m}$ *is purely negative with respect to the Hermitian form M and the subspace* $\ker P \subset \mathcal{V}_{2n}$ *is purely positive.*

b) Under these conditions, the operator $\mathrm{Be}(P)$ *also is bounded as an operator $\mathcal{S}(\mathbb{R}^{2n}) \to \mathcal{S}(\mathbb{R}^{2m})$ and $\mathcal{S}'(\mathbb{R}^{2n}) \to \mathcal{S}'(\mathbb{R}^{2m})$.*

We present two proofs of this theorem. The first (see next section) is based on certain classification problems in linear algebra. More conceptual proof based on fixed point principle is contained in Section 5.3.

In the next subsection, we produce a definition from the last theorem.

8.3. The symplectic category Sp. Objects of the category **Sp** are the standard symplectic spaces $\mathcal{V}_{2n}$, where $n = 0, 1, 2, \dots$.

Morphisms $\mathcal{V}_{2n} \to \mathcal{V}_{2m}$ are contractive Lagrangian linear relations $P : \mathcal{V}_{2n} \rightrightarrows \mathcal{V}_{2m}$ satisfying the additional conditions

3*a).The subspace $\ker P \subset \mathcal{V}_{2n}$ is M-positive,

3*b) the subspace $\operatorname{indef} P \subset \mathcal{V}_{2m}$ is M-negative.

Remark. Note that the positivity condition 2^* implies the semi-positivity of $\ker P$ and semi-negativity of $\operatorname{indef} P$. Thus 3^* is a minor refinement of 2^*. □

Theorem 8.3 *The category* Sp *is well defined, i.e., a product of two morphisms is a morphism.*

Proof is given below in Subsection 9.1.

Theorem 8.4 *The group* $\mathrm{Aut}_{\mathbf{Sp}}(\mathcal{V}_{2n})$ *of automorphisms of* $\mathcal{V}_{2n}$ *is* $\mathrm{Sp}(2n, \mathbb{R})$.

Proof is given below in 9.2.

8.4. The product formula.

Theorem 8.5 *Let* $P : \mathcal{V}_{2n} \rightrightarrows \mathcal{V}_{2m}$, $Q : \mathcal{V}_{2m} \to \mathcal{V}_{2k}$ *be morphisms of the category* **Sp**. *Then*

$$\mathrm{Be}(Q)\,\mathrm{Be}(P) = c(Q, P)\,\mathrm{Be}(QP), \tag{8.4}$$

where $c(Q, P)$ *is a nonzero complex constant.*

Proof is given below.

Remark. Theorem 8.1 defines the operators $\mathrm{Be}(P)$ to within a scalar factor. Hence a presence of a constant factor in formula (8.4) is not surprising. We discuss this factor in Section 3.7.

1.9 The symplectic category. Details

9.1. The symplectic category Sp is well-defined. For morphisms $P : \mathcal{V}_{2n} \rightrightarrows \mathcal{V}_{2m}$, $Q : \mathcal{V}_{2m} \to \mathcal{V}_{2k}$, we must verify that QP satisfies the conditions $1^* - 3^*$

A. *The subspace QP is $\Lambda^{\ominus}$-isotropic.* Let $v \oplus w$, $v' \oplus w' \in P$ and $w \oplus y$, $w' \oplus y' \in Q$,

$$\begin{aligned} P \text{ is } \Lambda^{\ominus}\text{-isotropic} \qquad &\Longrightarrow \qquad \Lambda(v, v') = \Lambda(w, w'), \\ Q \text{ is } \Lambda^{\ominus}\text{-isotropic}, \qquad &\Longrightarrow \qquad \Lambda(w, w') = \Lambda(y, y') \end{aligned}$$

Therefore for $v \oplus y$, $v' \oplus y' \in QP$ we get $\Lambda(v, v') = \Lambda(y, y')$. Hence QP is $\Lambda^{\ominus}$-isotropic.

The subspace QP is $M^{\ominus}$-nonnegative. Let v, w, y be as above,

$$\begin{aligned} P \text{ is } M^{\ominus}\text{-nonnegative} \qquad &\Longrightarrow \qquad M(v,v) \geqslant M(w,w) \\ Q \text{ is } M^{\ominus}\text{-nonnegative} \qquad &\Longrightarrow \qquad M(w,w) \geqslant M(y,y). \end{aligned}$$

Therefore $M(v,v) \geqslant M(y,y)$ and hence QP is nonnegative.

To continue a proof, we need the following two lemmas.

Lemma 9.1 *Let W be a linear space endowed with a nondegenerate skew-symmetric form $\{\cdot,\cdot\}$. Let P be a Lagrangian subspace, Y a subspace where the form is nondegenerate. Then the orthocomplement of $P \cap Y$ in Y coincides with the projection of P on Y.*

Proof of Lemma. Denote this orthocomplement by R,

$$R = Y \cap (Y \cap P)^{\perp}$$

Let us write vectors in W as $y \oplus z$, where $y \in Y$, $z \in Y^{\perp}$.

For $h \in P \cap Y$, we have $h \oplus 0 \in P$. Let $y \oplus z \in P$. Then

$$0 = \{h \oplus 0, y \oplus z\} = \{h, y\}$$

and hence the vector y (which is the projection of $y \oplus z$ on Y) is contained in $(P \cap Y)^{\perp} \cap Y$.

Let us verify the inverse inclusion. Consider a vector ζ in the orthocomplement of our projection in Y. For $y \oplus \zeta \in P$,

$$\{y \oplus z, \zeta \oplus 0\} = \{y, \zeta\} = 0$$

Thus $\zeta \oplus 0$ is orthogonal to the Lagrangian subspace P, and hence it is contained in P. Therefore $\zeta \in (P \cap Y)^{\perp} \cap Y$. □

Lemma 9.2 *For a morphism $P : \mathcal{V}_{2n} \rightrightarrows \mathcal{V}_{2m}$ of symplectic category, we have*

$$(\ker P)^{\perp\Lambda} = \operatorname{dom} P; \qquad (\operatorname{indef} P)^{\perp\Lambda} = \operatorname{im} P.$$

This follows from the previous lemma, $W = \mathcal{V}_{2n} \oplus \mathcal{V}_{2m}$, $Y = \mathcal{V}_{2n}$, $P = P$, $P \cap Y = \ker P$, and the projection of P on $\mathcal{V}_{2n}$ is $\operatorname{dom} P$.

Now we continue a proof of the theorem.

The subspace QP is $\Lambda^{\ominus}$-Lagrangian. Since QP is isotropic, it is sufficient ырсюЁрэЄto verify that $\dim QP = n + m$. To apply Theorem 7.1 about dimension of product we must verify conditions (7.4) and (7.5). First, $\ker Q \cap \operatorname{indef} P = 0$, because one subspace is M-positive, and another subspace is M-negative. By the lemma,

$$(\ker Q)^{\perp\Lambda} = \operatorname{dom} Q, \qquad (\operatorname{indef} P)^{\perp\Lambda} = \operatorname{im} P.$$

Hence

$$(\mathrm{dom}\, Q + \mathrm{im}\, P)^{\perp\Lambda} = (\mathrm{dom}\, Q)^{\perp} \cap (\mathrm{im}\, P)^{\perp\Lambda} = \ker Q \cap \mathrm{indef}\, P = 0$$

and (7.5) also is valid.

Problem 9.1 We used redundant arguments. Product of any two Lagrangian relations is Lagrangian. Apply Lemma 9.2, and formula (7.7).

The subspace $\mathrm{indef}\, QP$ *is strictly negative.* Let $y \in \mathrm{indef}\, QP$, consider $w \in W$ such that $0 \oplus w \in P$, $w \oplus y \in Q$.

If $w \neq 0$, then, by 3A*, $M(w,w) < 0$. By M-contractivity, it follows that $M(y,y) \leqslant M(w,w) < 0$.

If $w = 0$, then, by 3A*, we have $M(y,y) < 0$.

In both cases, $M(y,y) < 0$. □

9.2. Groups of automorphisms are $\mathrm{Sp}(2n,\mathbb{R})$**.** The symplectic group $\mathrm{Sp}(2n,\mathbb{R})$ acts on the real symplectic space $\mathcal{V}_{2n}(\mathbb{R})$ preserving the form $\{\cdot,\cdot\}$. Hence it acts (via the same matrices) on the complex space $\mathcal{V}_{2n}$ preserving the bilinear from Λ and the Hermitian form M. For $g \in \mathrm{Sp}(2n,\mathbb{R})$, its graph $\mathrm{graph}(g) \subset \mathcal{V}_{2n} \oplus \mathcal{V}_{2n}$ is $\Lambda^{\ominus}$-isotropic and $M^{\ominus}$-isotropic (see Subsection 7.7).

In particular, $\mathrm{graph}(g)$ is a morphism of the category **Sp**[11]. Now we wish to show that all the automorphisms of category **Sp** have this form.

Proof of Theorem 8.4. Let R be an automorphism. Then $RR^{-1} = R^{-1}R = 1$.. By (7.1), (7.2), it follows that $\ker R = 0$, $\mathrm{indef}\, R = 0$. Since $\dim R = 2n = \dim \mathcal{V}_{2n}$, R is a graph of an invertible operator $g : \mathcal{V}_{2n} \to \mathcal{V}_{2n}$.

Since R is $\Lambda^{\ominus}$-Lagrangian, the operator g preserves the bilinear form $\Lambda^{\ominus}$ (see Subsection 7.7). Since R is M-positive, the operator g satisfies the condition

$$M(gv, gv) \leqslant M(v,v).$$

Since R is invertible in the category **Sp**, we have also

$$M(g^{-1}w, g^{-1}w) \leqslant M(w,w).$$

Assuming $w = gv$, we obtain $M(gv,gv) = M(v,v)$.

Thus g preserves both form Λ and M. By (5.8), g also commutes with the complex conjugation J. Indeed, for all v, $v \in \mathcal{V}_{2n}$

$$M(v, Jv') = \Lambda(v,v') = \Lambda(gv,gv') = M(gv, Jgv') = M(v, g^{-1}Jgv)$$

and hence $J = g^{-1}Jg$. Hence, the matrix of g is real. Therefore $g \in \mathrm{Sp}(2n,\mathbb{R})$. □

Theorem 9.3 *a) For* $g \in \mathrm{Sp}(2n,\mathbb{R})$*, the corresponding operator* $\mathrm{Be}(P)$ *in* $L^2(\mathbb{R}^n)$ *is unitary to within a scalar factor*

b) This operator is bounded as an operator $\mathcal{S}(\mathbb{R}^n) \to \mathcal{S}(\mathbb{R}^n)$.

[11] Below we say that g itself is a morphism of the category **Sp**. It is not an inaccuracy of language, since the official set-theoretical definition of a map in the Zermelo axiom system (see, for instance K.Kuratowski, A.Mostowski "Set Theory") is given in terms of its graph.

PROOF. The statement a) is a rephrasing of Theorem 2.5. To prove b) we must note that the operators (2.14)–(2.16) are bounded in $\mathcal{S}$. □

9.3. Remark. Operators that are contained in $\mathrm{End}(\mathcal{V}_n)$.

Proposition 9.4 *An operator $g : \mathcal{V}_{2n} \to \mathcal{V}_{2n}$ is a morphism of the category* Sp *iff*

$$M(gv, gv) \leqslant M(v, v); \qquad \Lambda(gv, gw) = \Lambda(v, w). \tag{9.1}$$

The statement is obvious. □

The second condition (9.1) means that g is an element of complex symplectic group $\mathrm{Sp}(2n, \mathbb{C})$. For a discussion of the first condition, see Subsection 2.2.7.

9.4. Preliminary discussion of the product formula.

QUASI-PROOF OF THEOREM 8.5. This statement is obvious modulo the boundedness Theorem 8.2. Indeed, for $v \oplus w \in P$, $w \oplus y \in Q$,

$$\mathrm{Be}(Q)\,\mathrm{Be}(P)\widehat{a}(v) = \mathrm{Be}(Q)\widehat{a}(w)\,\mathrm{Be}(P) = \widehat{a}(y)\,\mathrm{Be}(Q)\,\mathrm{Be}(P)$$

(this row makes sense since all the present operators are bounded in the Schwartz space $\mathcal{S}$). But

$$\mathrm{Be}(QP)\widehat{a}(v) = \widehat{a}(y)\,\mathrm{Be}(QP).$$

Bearing in mind Theorem 8.1.b, we obtain $\mathrm{Be}(QP) = c \cdot \mathrm{Be}(Q)\,\mathrm{Be}(P)$. □

Thus Theorem 8.5 is proved modulo boundedness Theorem 8.2.

Nevertheless the following special cases of Theorem 8.5 really are proved.

Lemma 9.5 *The product formula (8.4) holds if at least one of the following conditions is satisfied*

– *the operator* $\mathrm{Be}(P) : \mathcal{S}(\mathbb{R}^{2n}) \to \mathcal{S}(\mathbb{R}^{2m})$ *is bounded.*

– *the operator* $\mathrm{Be}(Q) : \mathcal{S}'(\mathbb{R}^{2m}) \to \mathcal{S}'(\mathbb{R}^{2k})$ *is bounded.*

9.5. Images of Gaussian vectors. Let $R \subset \mathcal{V}_{2n}$ be a non-positive Lagrangian subspace. Denote by $\mathrm{be}(R) \in L^2(\mathbb{R}^n)$ the corresponding Gaussian distribution, i.e., $\mathrm{be}(R)$ a nonzero distribution satisfying the partial differential equations

$$\widehat{a}(v)\,\mathrm{be}(R) = 0 \qquad \text{for all } v \in \mathbb{R}$$

If the subspace R is strictly negative, then $\mathrm{be}[P]$ is a Gaussian vector.

Theorem 9.6 *The image of a Gaussian vector with respect to a Gaussian operator is a non-zero Gaussian distribution. Precisely,*

$$\mathrm{Be}(P)\,\mathrm{be}(R) = \lambda \cdot \mathrm{be}(PR), \qquad \lambda \in \mathbb{C} \setminus 0. \tag{9.2}$$

PROOF. Since $R : 0 \rightrightarrows \mathcal{V}_{2n}$ is a contractive linear relation $0 \rightrightarrows \mathcal{V}_{2n}$, we apply Lemma 9.5. □

It remains to show that $\lambda \neq 0$.

Lemma 9.7 *Let χ be a Gaussian distribution on $\mathbb{R}^k$, $\mathrm{be}(R) \in L^2(\mathbb{R}^k)$ a Gaussian vector. Then*

$$\ll \chi, \mathrm{be}(R) \gg \neq 0.$$

PROOF OF LEMMA. By definition (3.3) of Gaussian distribution, the expression $\ll \chi, \mathrm{be}(R) \gg$ is a Gaussian integral of the form

$$\int_L \exp\Big\{-\frac{1}{2} x A x^t\Big\}\, dx = (2\pi)^{\dim L/2} \det A^{-1/2} \neq 0. \qquad \square$$

Now we return back to the theorem. Let $K(x,y)$ be the kernel of $\mathrm{Be}(P)$. Let $\mathrm{be}(R) = \exp\{-\frac{1}{2} y B y^t\}$, let $\exp\{-\frac{1}{2} x C x^t\} \in \mathcal{S}(\mathbb{R}^m)$ be a Gaussian vector in the target space. Then

$$\ll \mathrm{Be}(P) \exp\{-\frac{1}{2} y B y^t\}, \exp\{-\tfrac{1}{2} x C x^t\} \gg =$$
$$= \ll K(x,y), \exp\{-\frac{1}{2} y B y^t\} \cdot \exp\{-\tfrac{1}{2} x C x^t\} \gg \neq 0.$$

Corollary 9.8 *Product of two Gaussian operators is nonzero.*

PROOF. Indeed, a Gaussian operator takes a nonzero Gaussian vector to a non-zero vector. $\square$

9.6. Explicit formula for $\mathrm{Be}(P)$**.** In fact, an explicit formula was obtain above, see Section 1.6 and Theorem 6.3 Here we present he same result more precisely.

Let $P : \mathcal{V}_{2n} \rightrightarrows \mathcal{V}_{2m}$ be a morphism of the category **Sp**. *Assume that P is a graph of an operator*

$$S := \begin{pmatrix} K & L \\ L^t & M \end{pmatrix} : \mathcal{V}_m^- \oplus \mathcal{V}_n^- \to \mathcal{V}_m^+ \oplus \mathcal{V}_n^+. \tag{9.3}$$

REMARK. In fact, our matrix S, see (9.3), defines coordinate system on the set $\mathrm{Mor}_{\mathbf{Sp}}(\mathcal{V}_{2n}, \mathcal{V}_{2m})$. Below, we call S the *Potapov transform* of P. $\square$

Theorem 9.9 *Under this condition,*

$$\mathrm{Be}(P) f(x) = \int_{\mathbb{R}^n} \exp\Big\{-\frac{1}{2} \begin{pmatrix} x & y \end{pmatrix} \begin{pmatrix} K & -L \\ L^t & -M \end{pmatrix} \begin{pmatrix} x^t \\ y^t \end{pmatrix}\Big\}\, f(y)\, dy. \tag{9.4}$$

PROOF. This is a rephrasing of Theorems 6.2 and 6.3. $\square$

REMARK. Theorems 6.2 and 6.3 allow to write explicit formulae for Gaussian kernels in the remaining cases. In Chapter 5, we achieve the same purpose using more convenient language.

9.7. Addendum to the construction. Involution in the category Sp and adjoint operators. For a morphism $P : \mathcal{V}_{2n} \rightrightarrows \mathcal{V}_{2m}$, consider the complex conjugate subspace

$$\overline{P} := JP = P^{\perp M}$$

(we apply Lemma 5.3). Denote

$$P^\sigma := \overline{P}^\square,$$

where $P^\square$ is the pseudo-inverse (see Subsection 7.3).

Proposition 9.10 *a) P^σ is a morphism $\mathcal{V}_{2m} \rightrightarrows \mathcal{V}_{2n}$.*
b) $(PQ)^\sigma = Q^\sigma P^\sigma$.
c) $P^{\sigma\sigma} = P$.

PROOF. b), c) are obvious.

a) To prove a), we refer to 5.3. By Lemma 5.3, it follows that P^σ is Lagrangian. By Lemma 5.5, $\overline{P} = P^{\perp M}$ is non-positive in $\mathcal{V}_{2n} \oplus \mathcal{V}_{2m}$. The transposition of summands implies a changing of a sign of $M^\ominus$. Hence P^σ is non-negative

The condition 3^* is a corollary of the following trivial lemma.

Lemma 9.11 $\ker P^\sigma = (\operatorname{im} P)^{\perp M}$, $\operatorname{indef} P^\sigma = (\operatorname{dom} P)^{\perp M}$.

Proposition 9.12 *For* $\operatorname{Be}(P) : L^2(\mathbb{R}^n) \to L^2(\mathbb{R}^m)$,

$$\operatorname{Be}(P)^* = \operatorname{Be}(P^\sigma).$$

PROOF. Look to the defining equation (8.3) for $\operatorname{Be}(P)$ and take the adjoints of both sides. We obtain

$$\widehat{a}(\overline{v})\mathcal{B}^* = \mathcal{B}^*\widehat{a}(\overline{w}).$$

This is the defining equation for $\operatorname{Be}(P^\sigma)$. $\square$

9.8. Addendum to the construction. Tensor products. Now let us use coordinate-free language as in Subsection 5.2. Let $\mathcal{Y}$, $\mathcal{Z} \in \operatorname{Ob}(\mathbf{Sp})$. We define their direct sum $\mathcal{Y} \oplus \mathcal{Z}$ as the linear space $\mathcal{Y} \oplus \mathcal{Z}$ endowed with two forms

$$\Lambda^\oplus(v \oplus w, v' \oplus w') = \Lambda(v, v') + \Lambda(w, w'),$$
$$M^\oplus(v \oplus w, v' \oplus w') = M(v, v') + M(w, w')$$

and by the antilinear operator $J^\oplus := J \oplus J$.

We have the obvious identification

$$\mathcal{V}_{2\alpha} \oplus \mathcal{V}_{2\beta} \simeq \mathcal{V}_{2(\alpha+\beta)}, \qquad \mathcal{V}^\pm_\alpha \oplus \mathcal{V}^\pm_\beta \simeq \mathcal{V}^\pm_{\alpha+\beta}$$

In particular, for morphisms

$$P : \mathcal{V}_{2\alpha} \rightrightarrows \mathcal{V}_{2\beta}, \qquad R : \mathcal{V}_{2\gamma} \rightrightarrows \mathcal{V}_{2\delta}$$

we have a well-defined morphism

$$P \oplus R : \mathcal{V}_{2(\alpha+\gamma)} \rightrightarrows \mathcal{V}_{2(\beta+\delta)}$$

Proposition 9.13

$$\operatorname{Be}(P \oplus R) = \operatorname{const} \cdot \operatorname{Be}(P) \otimes \operatorname{Be}(R)$$

This is obvious, see also Subsection 3.12.

1.10 Proof of boundedness. Canonical forms of Gaussian operators

In what follows, we derive the conditions of L^2-boundedness of Gaussian operators (see Theorem 8.2). For this purpose, we reduce the boundedness Gaussian operators morphisms of **Sp** to canonical forms of morphisms of the category **Sp**. This way is simple but its exposition is long.

More conceptual proof is given in Chapter 5.

10.1. Problem of description of canonical forms. Let $\mathcal{Y}$, $\mathcal{Z}$ be objects of the category **Sp** (now we prefer coordinate-free language) and $P : \mathcal{Y} \rightrightarrows \mathcal{Z}$ a morphism. Let $g \in \mathrm{Aut}(\mathcal{Y})$, $h \in \mathrm{Aut}(\mathcal{Z})$ be elements of automorphism groups (i.e., elements of real symplectic groups).

Theorem 9.3 and Lemma 9.5 (imperfect product formula) yields that the operators $\mathrm{Be}(gPh^{-1})$ and $\mathrm{Be}(P)$ are bounded or unbounded simultaneously.

Therefore it is reasonable to describe canonical forms of morphisms $P : \mathcal{Y} \rightrightarrows \mathcal{Z}$ under transformations

$$P \mapsto gPh^{-1}, \qquad \text{where } g \in \mathrm{Sp}(2m, \mathbb{R}),\ h \in \mathrm{Sp}(2n, \mathbb{R})$$

We say that a morphism $P : \mathcal{Y} \rightrightarrows \mathcal{Z}$ is *decomposable* if there are decompositions $\mathcal{Y} = \mathcal{Y}_1 \oplus \mathcal{Y}_2$, $\mathcal{Z} = \mathcal{Z}_1 \oplus \mathcal{Z}_2$ and morphisms $P_1 : \mathcal{Y}_1 \rightrightarrows \mathcal{Z}_1$, $P_2 : \mathcal{Y}_2 \rightrightarrows \mathcal{Z}_2$ such that $P = P_1 \oplus P_2$.

Now we can formulate the problem of canonical forms as

— *to describe (to within the equivalence) all the indecomposable morphisms of the category* **Sp**

10.2. Classification of indecomposable morphisms.

Theorem 10.1 *The 5 types of indecomposable morphisms enumerated below exhaust all the indecomposable morphisms of the category* **Sp**. *In particular, each morphism of category* **Sp** *is a direct sum of morphisms enumerated below.*

In fact, all these operators were discussed above in Section 1.3.

1^*. A morphism $\mathcal{V}_0 \to \mathcal{V}_2$. All such morphisms are equivalent.

The corresponding operator $L^2(\mathbb{R}^0) = \mathbb{R} \to L^2(\mathbb{R}^1)$ is a 'Gaussian vector', see (3.16).

2^*. A morphism $\mathcal{V}_2 \to \mathcal{V}_0$. All such morphisms are equivalent.

The Gaussian operator $L^2(\mathbb{R}^1) \to L^2(\mathbb{R}^0) = \mathbb{R}$ is a 'Gaussian linear functional' (3.17).

3^*. The unit operator $\mathcal{V}_2 \to \mathcal{V}_2$.

The Gaussian operator is the identity operator $L^2(\mathbb{R}^1) \to L^2(\mathbb{R}^1)$, see (3.6);

4^*. The operator $g : \mathcal{V}_2 \to \mathcal{V}_2$ having the matrix $g = \begin{pmatrix} \cosh t & \sinh t \\ \sinh t & \cosh t \end{pmatrix}$, where $t > 0$ The Gaussian operators are the Mehler operators $U(t)$ given by (3.11).

5^*. This type can be represented by linear transformations $g : \mathcal{V}_2 \to \mathcal{V}_2$ with matrices

$$\begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix} \qquad \text{or, equivalently} \qquad \begin{pmatrix} 1 & 0 \\ -t & 1 \end{pmatrix}, \qquad t > 0$$

all such morphisms are equivalent. The corresponding Gaussian operators are respectively the heat operators $V(t)$ (3.9) or the multiplication operators $A(t)$ (3.15).

10.3. Reduction of Theorem 10.1 to a standard problem.

Proposition 10.2 *Each morphism of the category* Sp *admits a decomposition into a direct sum of morphisms of 3 following types*

– *morphisms* $\mathcal{X} \rightrightarrows 0$,

– *morphisms* $0 \rightrightarrows \mathcal{Y}$,

– *invertible operators* $g : \mathcal{Z} \to \mathcal{Z}$ *such that*

$$\Lambda(gv, gv) = \Lambda(v, v); \qquad M(gv, gv) \leqslant (v, v). \tag{10.1}$$

Thus, our problem is reduced to canonical forms of symplectic contractive operators, i.e., elements of the complex symplectic group $\mathrm{Sp}(2n, \mathbb{C})$ contracting the form M, see (10.1). All the possible indecomposable symplectic contractive operators are the 2-dimensional operators 3^*–5^* from the previous subsection.

PROOF. Let $P : \mathcal{Y} \rightrightarrows \mathcal{Z}$ be a morphism, assume $\ker P \neq 0$.

The form M is positive on $\ker P$. By (5.7), the form M is negative on $J \ker P$. Hence $\ker P \cap J \ker P = 0$.

Let us show that

$$\mathcal{Y}_1 := \ker P \oplus J \ker P \quad \text{is a standard symplectic space.}$$

Indeed, the space $\ker P$ is Λ-isotropic; therefore $J \ker P$ is M-orthogonal to $\ker P$. By Corollary 5.4, the form M on $\ker P \oplus J \ker P$ is nondegenerate.

Denote (see Lemma 5.6)

$$\mathcal{Y}_1 := \ker P \oplus J \ker P, \qquad \mathcal{Y}_2 := \mathcal{Y}_1^{\perp\Lambda} = \mathcal{Y}_1^{\perp M}$$

We define the linear relations $P_1 : \mathcal{Y}_1 \rightrightarrows 0$, $P_2 : \mathcal{Y}_2 \rightrightarrows Z$ being

$$P_1 := \ker P, \qquad P_2 := P \cap \big(\mathcal{Y}_2 \oplus \mathcal{Z}\big).$$

Lemma 10.3 *a) The subspace* P_2 *is a morphism* $\mathcal{Y}_2 \to \mathcal{Z}$.

b) $P = P_1 \oplus P_2$

The lemma yields the proposition because we can split the indefiniteness in the same way.

PROOF OF THE LEMMA.Define Q being the M-complement of $\ker P$ in P. The subspacse Q is Λ- and M-orthogonal to $\ker P$, hence Q is M- and Λ-orthogonal to $J \ker P$. Thus $Q \subset P_2$.

Since the form M is nondegenerate on $\ker P$, we get $P = \ker P \oplus Q$ and this implies our statement. □

10.4. The boundedness. Proof of sufficiency. For all the indecomposable morphisms $1^* - -5^*$ the boundedness is trivial.

Direct sums of morphisms correspond to tensor products, see Proposition 9.13.

Thus, the boundedness in L^2 and in $\mathcal{S}$ became obvious.

10.5. Boundedness in L^2. Proof of necessity. Consider a contractive Lagrangian linear relation $P : \mathcal{V}_{2n} \rightrightarrows \mathcal{V}_{2m}$ and the corresponding Gaussian operator $\mathrm{Be}(P) : \mathcal{S}(\mathbb{R}^n) \to \mathcal{S}'(\mathbb{R}^m)$.

Consider an arbitrary Gaussian vector $\mathrm{be}(R) \in L^2(\mathbb{R}^n)$. Then $\mathrm{Be}(P)\,\mathrm{be}(R) \in L^2(\mathbb{R}^m)$ is a Gaussian vector (see Theorem 9.6). By (8.3), for $w \in \mathrm{indef}(P)$, we have

$$\widehat{a}(w)\,\mathrm{Be}(P) = 0 \qquad \Longrightarrow \qquad \widehat{a}(w)\,\mathrm{Be}(P)\,\mathrm{be}(R) = 0$$

By Theorem 6.1, it follows that $M(w,w) > 0$.

Thus, $\mathrm{indef}\, P$ is negative.

Consider the adjoint operator (see Theorem 9.12)

$$\mathrm{Be}(P)^* = \mathrm{Be}(P^\sigma) : L^2(\mathbb{R}^m) \to L^2(\mathbb{R}^n)$$

But $\ker P = J\,\mathrm{indef}\, P^\sigma$. The subspace $\mathrm{indef}\, P^\sigma$ is negative, by Lemma 5.5, $\ker P$ is positive. □

1.11 Bibliographical remarks

11.1. The Weil representation. Apparently, the Weil representation must be attributed to K.O.Friedrichs($\sim$ 1950). His book [43] does not contain the Weil representation of $\mathrm{Sp}(2n,\mathbb{R})$, but evidently he thought that there is nothing to discuss here; he tried to extended the Weil representation to infinite-dimensional case. His book contains correct conjectures concerning this topic supported by 'arguments of physical level'. However, in finite-dimensional case his argument (a reference to the Stone–von Neumann theorem, see Addendum B) definitely is sufficient. His conjectures for the infinite-dimensional case were proved by D.Shale [165] and F.A.Berezin (in several short papers 1960-62 which are not cited in our bibliography and in the book [11])

The construction exposed in Section 1.2 belongs to I.Segal. A.Weil, who is the label of the construction, extended it to finite and p-adic fields [178].

11.2. The Mehler formula was obtain by Mehler [105] in 1866, see also W.Myller-Lebedeff [107].

11.3. The semigroup of compressive symplectic operators. As we seen in Proposition 9.4, the Weil representation of $\mathrm{Sp}(2n,\mathbb{R})$ admits a holomorphic continuation to a subsemigroup in $\mathrm{Sp}(2n,\mathbb{C})$. This fact is a special case of G.I.Olshanski's theorem [132]:

— *A highest weight unitary representation of a semisimple group G admits holomorphic continuation on a certain semigroup $\Gamma \subset G_{\mathbb{C}}$.*

For the groups $G = \mathrm{U}(p,q)$, this phenomenon was met by M.I.Graev in 1958 [58]; in this way he calculated characters of Harish-Chandra holomorphic discrete series.

11.4. The semigroup of Gaussian operators and the strange multiplication (4.2). Apparently, many authors independently met this formula and the corresponding strange algebraic structure. I know definitely about R.Howe and G.I.Olshanski. Both of them found the condition of boundedness and also described canonical forms. For references to earlier papers, see [72]; see also [27]. Howe's paper [72] also is exposed in G.Folland's book [40].

11.5. The multiplication theorem. Theorem 8.5 was obtained in [113], [114].

11.6. Some references on Gaussian operators in real settings: [38], [95], [70].

11.7. Jordan normal form a linear relation. Apparently it was discovered independently several times, for instance [51].

11.8. Pairs of forms. On realizations of classical groups as groups preserving pairs of forms, see [120].