Chapter 2

Pseudo-Euclidean geometry and groups $\mathrm{U}(p,q)$

The standard inner product in an n-dimensional Euclidean space $V = \mathbb{R}^n$ is given by

$$\langle x, y\rangle = \sum_{i=1}^{n} x_j y_j.$$

Replacing this formula by

$$\langle x, y\rangle = -\sum_{i\leqslant p} x_j y_j + \sum_{j>p} x_j y_j$$

we obtain an another geometry, namely the pseudo-Euclidean geometry. It has some similar characteristics with the Euclidean geometry, but also it differs in many details. The basic theorems of Euclidean geometry (as the reduction of quadratic forms to principal axes and the classification of isometries) survive, but not literally; the generalizations are not obvious. Also there arise new phenomena, that do not exist in Euclidean geometry.

There are 7 geometries of this kind (see Addendum A.1), they rather often appear in mathematics. We consider in details two geometries, namely, the complex pseudo-Euclidean geometry (in this chapter) and real symplectic geometry (the next Chapter) The pseudo-Euclidean geometry was a subject of some works by L.S.Pontrjagin, M.G.Krein, M.S.Livshits, V.P.Potapov in the operator theory in 1940-50s. Our point of view is partially influenced by these papers.

Throughout this chapter, a 'linear space' is a finite dimensional linear space over $\mathbb{C}$.

2.1 Geometry of indefinite Hermitian forms

1.1. Definition of Hermitian forms. Let V be a complex finite-dimensional linear space. A *sesquilinear form* on V is a $\mathbb{C}$-valued function $B(v, w)$ on $V \times V$ such that

$$B(v_1 + v_2, w) = B(v_1, w) + B(v_2, w), \quad B(v, w_1 + w_2) = B(v, w_1) + B(v, w_2),$$
$$B(\lambda v, w) = \lambda B(v, w), \qquad B(v, \lambda w) = \overline{\lambda} B(v, w).$$

A sesquilinear form is *Hermitian* if

$$B(w, v) = \overline{B(v, w)}.$$

In particular, $B(v, v) \in \mathbb{R}$; indeed $B(v, v) = \overline{B(v, v)}$

REMARK. Adding the condition $B(v, v) > 0$ for $v \neq 0$, we obtain the definition of an inner product in a complex Euclidean space. □

We say that a vector v is

— *B-positive*, if $B(v, v) > 0$,

— *B-negative*, if $B(v, v) < 0$,

— *B-isotropic*, if $B(v, v) = 0$,

Two vectors v, w are *orthogonal* if $B(v, w) = 0$.

Proposition 1.1 *a) Each Hermitian form on $\mathbb{C}^n$ can be expressed as*

$$B(v, w) = vQw^*,$$

where v, $w \in \mathbb{C}^n$ are vector-rows and Q is an $n \times n$ Hermitian matrix.

b) Under a change of variables $v \mapsto vg$, the matrix Q converts as

$$Q \mapsto gQg^*$$

PROOF. Denote by e_j the standard basis in $\mathbb{C}^n$. We have

$$B(v, w) = B\Big(\sum v_j e_j, \sum w_k e_k\Big) = \sum v_j \overline{w}_k B(e_j, e_k) \qquad \square$$

1.2. Kernels of form3. A *kernel* $\ker B$ of a form B is the subspace $W \subset V$ consisting of vectors w such that $B(w, v) = 0$ for all v.

The *rank* of the form B is the codimension of $\ker B$.

A form B is *nondegenerate* if its kernel is zero.

Lemma 1.2 *A form B induces a nondegenerate Hermitian form on the quotient space $V/\ker B$.*

PROOF. The space $V/\ker B$ consists of vectors defined to within the equivalence

$$v \sim v' \Longleftrightarrow v - v' \in \ker B$$

Let $v' = v + \xi$, $w' = w + \eta$, where ξ, $\eta \in \ker B$, Then

$$B(v', w') = B(v + \xi, w + \eta) = B(v, w) + 0 + 0 + 0$$

Therefore we really obtain a form on the quotient space.

Problem 1.1 Show that the form on $V/\ker P$ is nondegenerate.

A *pseudo-Euclidean space* is a complex linear space equipped with a nondegenerate Hermitian form; in this case we usually denote this form by $\langle \cdot, \cdot \rangle$ and call it a *scalar product.*

1.3. Several preliminary Lemmae.

Lemma 1.3 *Let $h \in V$ be an isotropic vector, $v \notin \ker B$. Then one can find a vector g such that*

$$\langle h, h \rangle = \langle g, g \rangle = 0, \qquad \langle h, g \rangle = 1. \tag{1.1}$$

PROOF. Consider an arbitrary vector $r \in V$ such that $B(h, r) = \sigma \neq 0$. Denoting $r' := \overline{\sigma}^{-1} r$, we get $B(h, r') = 1$. Put $g := r' + sh$, where s ranges in $\mathbb{R}$. Then

$$B(h, g) = 1, \qquad B(g, g) = B(r', r') + 2s$$

We take $s := -B(r', r')/2$ and obtain a desired g. □

Lemma 1.4 *Let a form B be nondegenerate. Then one can find a vector v such that $B(v, v) = \pm 1$*

PROOF. Take an arbitrary vector h. If it is not isotropic, then the conclusion holds. Otherwise we choose a vector g satisfying (1.1) and get $v := (h + g)/\sqrt{2}$. □

Lemma 1.5 *If one can find a positive vector and a negative vector, then there is a non-zero isotropic vector.*

PROOF. Let $B(e, e) = +1$, $B(f, f) = -1$. Then these vectors are linear independent. By the intermediate value theorem, one of the points of the segment $te + (1 - t)f$ is isotropic. □

1.4. Classification of Hermitian forms. For a subspace $S \subset V$, define its *orthocomplement* (or *orthogonal complement*) $Z^{\perp}$ as the set of all the vectors $v \in V$ such that $B(v, z) = 0$ for all $z \in S$.

REMARK. If the form B is degenerate, then $S^{\perp} \supset \ker B$ for all S. □

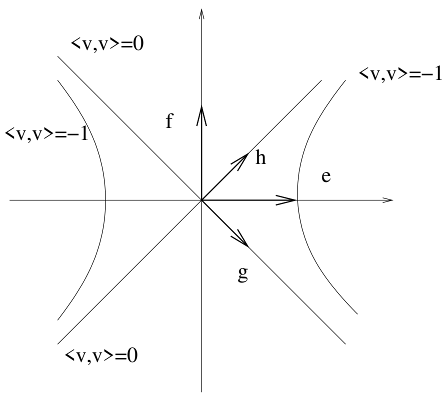

Fig. 2.1: The space $\mathbb{R}^{1,1}$ with the inner product
$\langle (v_1, v_2), (w_1, w_2)\rangle = -v_1 w_1 + v_2 w_2$.
The level curves $\langle v, v\rangle = \text{const}$ are hyperbolas; isotropic vectors v are contained in a pair of lines. We present one of possible bases e, f as in (1.3) and one of bases h, g as in (1.1).

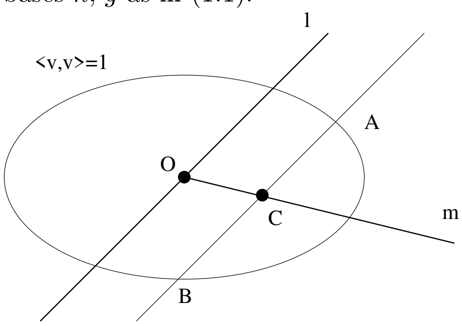

Fig. 2.2: One of interpretations of the orthocomplement. Consider the two dimensional real plane endowed with a positive inner product. The level curve $\langle v, v\rangle = 1$ is an ellipse. Let l be a line. Consider a parallel line to l. Let A, B be points of intersection of the former line and the ellipse. Consider the midpoint C of the segment $[A, B]$. The line OC is the orthocomplement to l.

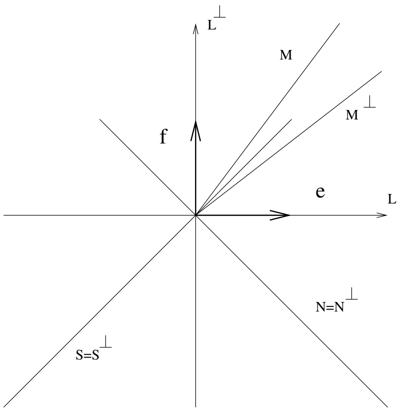

Fig. 2.3: Four examples of orthocomplement in the plane $\mathbb{R}^{1,1}$.

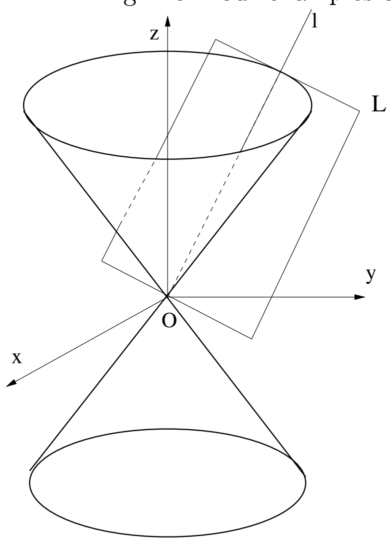

Fig. 2.4: An example of orthocomplements in $\mathbb{R}^{2,1}$. The cone is the cone of isotropic vectors. The orthocomplement $L = l^{\perp}$ of an isotropic line l is the plane tangent to the cone along L.

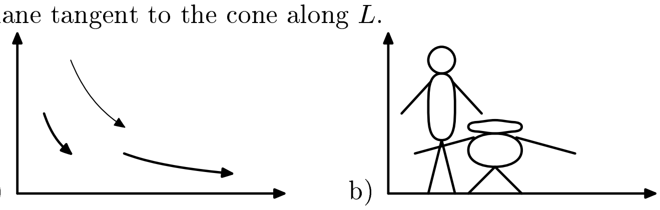

Fig. 2.5: Let us write a scalar product in $\mathbb{R}^{1,1}$ as
$\langle (v_1, v_2), (w_1, w_2)\rangle = v_1w_2 + v_2w_1$.
We instance a pseudo-orthogonal transformation of $\mathbb{R}^{1,1}$, namely a *hyperbolic rotation* $(v_1, v_2) \mapsto (2v_1, v_2/2)$.
Figure a) shows images of some points and their trajectories with respect to the 'rotations' $R_t : (v_1, v_2) \mapsto (2^t v_1, 2^{-t} v_2)$ for $0 \leqslant t \leqslant 1$.
On b) we show a transformation of a geometric configuration.

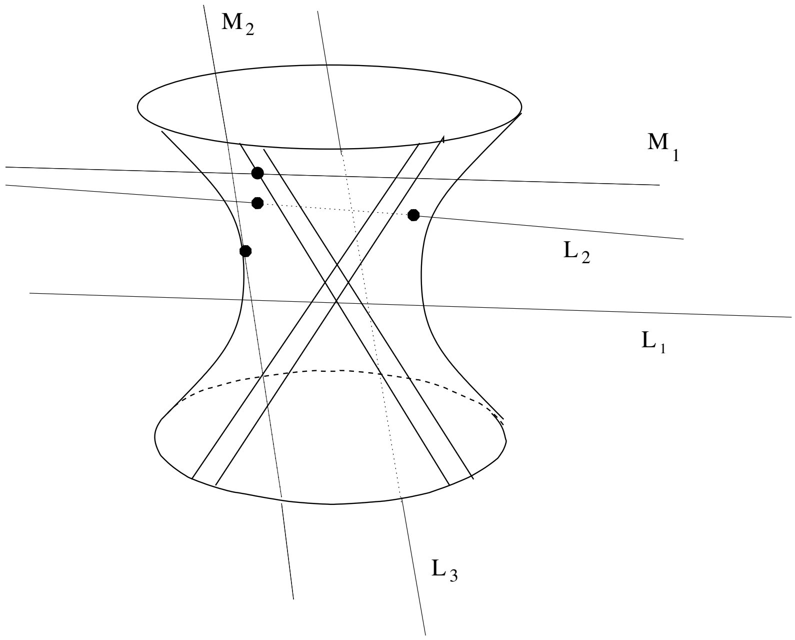

Fig. 2.6: To the Witt theorem. Consider the space $\mathbb{R}^{2,2}$ and its projectivization, i.e., the 3-dimensional real projective space $\mathbb{RP}^3$. The cone of isotropic vectors in $\mathbb{R}^{2,2}$ corresponds to a (toroidal) quadric in $\mathbb{RP}^3$, i.e., to a one-sheet hyperboloid. Lines in $\mathbb{RP}^3$ correspond to two-dimensional planes in $\mathbb{R}^{2,2}$.
1. L_1 corresponds to a positive plane
2. L_2 corresponds to a non-degenerate indefinite plane; points of intersection of L_2 and the hyperboloid correspond to the pair of isotropic lines in the plane.
3. L_3 corresponds to a negative plane
4. M_1 corresponds to a plane with inertia indices (1,0); the point of tangency indicates the kernel of the scalar product on the plane.
5. M_2 corresponds to a degenerate semi-positive plane
6. Linear elements of the hyperboloid correspond to isotropic planes in $\mathbb{R}^{2,2}$. In particular, we observe that the Grassmannian of maximal isotropic subspaces in $\mathbb{R}^{2,2}$ is homeomorphic to a pair of disjoint circles.
Denote by $S(i,j)$ the set of all two-dimensional planes with inertia indices (i,j). The scheme of adjoining of orbits is

$$\begin{tikzcd}
S(2,0) \arrow[dr] & & \\
S(1,1) \arrow[r] \arrow[dr] & S(1,0) \arrow[dr] & \\
 & S(0,1) \arrow[r] & S(0,0) \\
S(0,2) \arrow[ur] & &
\end{tikzcd}$$

Theorem 1.6 *a) Any Hermitian form on an n-dimensional space can be reduced to a sum of squares in some basis, precisely*

$$B(x,y) = -\sum_{j\leqslant p} x_j\overline{y}_j + \sum_{p<j\leqslant p+q} x_j\overline{y}_j \tag{1.2}$$

*b) The numbers p and q (*negative and positive inertia indices*) are independent on a choice of a basis.*

c) $\operatorname{rk} B = p+q$

For instance, in an n-dimensional pseudo-Euclidean space one can choose a basis $e_1,\dots, e_p$, $f_1,\dots, f_q$ (where $p+q=n$) such that

$$\langle e_1, e_1\rangle = \dots = \langle e_p, e_p\rangle = -1, \qquad \langle f_1, f_1\rangle = \dots = \langle f_q, f_q\rangle = +1 \tag{1.3}$$

and all other scalar products of basic elements are zero. We denote by $\mathbb{C}^{p,q}$ the pseudo-Euclidean space equipped with such a basis.

PROOF OF THEOREM. a) Consider an arbitrary complementary subspace Y to $\ker B$. Obviously, the form is nondegenerate on Y, therefore the statement is reduced to the case of nondegenerate forms.

Thus we assume that B is nondegenerate.

By Lemma 1.4, one can find a vector e_1 such that $B(e_1,e_1)=\pm 1$.

Lemma 1.7 $V = \mathbb{C}e_1 \oplus (\mathbb{C}e_1)^\perp$.

PROOF OF THE LEMMA. Let ξ be a vector noncollinear to e_1. Then the vector

$$\eta := \xi - \frac{\langle \xi, e_1\rangle}{\langle e_1, e_1\rangle}\, e_1$$

is orthogonal to e_1. Therefore $\xi \in \mathbb{C}e \oplus (\mathbb{C}e)^\perp$. □

Let us continue the proof of the theorem. Next, we take a vector $e_2 \in (\mathbb{C}e_1)^\perp$ such that $B(e_2,e_2)=\pm 1$, and consider its orthocomplement in $(\mathbb{C}e_1)^\perp$, etc.

c) is obvious.

b) is corollary of Lemma 1.8 given below, which is important by itself. □

1.5. Several definitions. A form B is *positive-definite* iff $B(v,v)>0$ for all nonzero $v\in V$. If the form is positive-definite, we say that the space is *Euclidean*. In this case, (1.2) is reduced to

$$B(x,y) = \sum_{j=1}^{n} x_j\overline{y}_j.$$

Recall also a collection of similar definitions. A form is

— *negative definite* iff $B(v,v)<0$ for all $v\neq 0$

— *nonnegative definite* (or *positive semi-definite*) iff $B(v,v)\geqslant 0$

— *non-positive definite* (or negative semi-definite) iff $B(v, v) \leqslant 0$

— *indefinite* if it has values of different signs.

Let B be an arbitrary Hermitian form in V. We say that a subspace $W \subset V$ is *positive* if the form B is positive definite on W. In a similar way we define *nonnegative, negative, etc.* subspaces.

A subspace W is said to be *regular* if B is nondegenerate on W. Otherwise, W is *singular*. By $zero(W)$ we denote the kernel of the form restricted to W.

Lemma 1.8 *The negative (resp. positive) inertia index of the form B coincides with maximal dimension of a negative (resp. positive) subspace.*

PROOF. Consider the form B given by formula (1.2) Assume that there exists a negative subspace X, whose dimension is $> p$. Denote by Z the linear span of the vectors $e_{p+1}, e_{p+2}, \dots$. Then $\dim X + \dim Z > \dim V$, therefore X and Z have a nonzero intersection. But X is negative and Z is nonnegative. $\square$

1.6. Another canonical form.

Theorem 1.9 *Let B be a Hermitian form with inertia indices $p \leqslant q$,* $\dim \ker B =$ *k. Then its matrix can be reduced to*

$$\begin{pmatrix} 0 & 1_p & 0 & 0 \\ 1_p & 0 & 0 & 0 \\ 0 & 0 & 1_{q-p} & 0 \\ 0 & 0 & 0 & 0_k \end{pmatrix} \qquad \text{or} \qquad \begin{pmatrix} 0 & 0 & 1_p & 0 \\ 0 & 1_{q-p} & 0 & 0 \\ 1_p & 0 & 0 & 0 \\ 0 & 0 & 0 & 0_k \end{pmatrix}. \tag{1.4}$$

Without loss of generality, we can assume that B is nondegenerate.

FIRST PROOF. We consider basis e_i, f_k as in (1.3). Define the new basis

$$h_j = \frac{1}{\sqrt{2}}(e_j + f_j), \quad g_j = \frac{1}{\sqrt{2}}(e_j - f_j), \quad \text{for } j \leqslant p, \qquad f_k, \quad \text{for } k > p.$$

This yields the desired canonical form.

SECOND PROOF. Consider an isotropic vector h_1 and a vector g_1 satisfying (1.1). Then

$$V = (\mathbb{C}h_1 \oplus \mathbb{C}g_1) \oplus (\mathbb{C}h_1 \oplus \mathbb{C}g_1)^{\perp}. \tag{1.5}$$

Indeed, for a vector $\xi \in V$ we put

$$\eta := \xi - \langle \xi, g_1 \rangle f_1 - \langle \xi, f_1 \rangle g_1$$

Then $\eta \in (\mathbb{C}h_1 \oplus \mathbb{C}g_1)^{\perp}$ and this yields (1.5)

Next, we consider vectors $h_2, g_2 \in (\mathbb{C}h_1 \oplus \mathbb{C}g_1)^{\perp}$ being in the same position as above, etc. As result, we obtain a decomposition of the form

$$\Big[\bigoplus_j (\mathbb{C}h_j \oplus \mathbb{C}g_j)\Big] \oplus R,$$

where R does not contains an isotropic vector. By Lemma 1.5, the subspace R is sign-definite. $\square$

1.7. Linear functionals. Let V be a pseudo-Euclidean space. We denote by V° the space of all the linear functionals on V. For a vector $v \in V$, we define the linear functional ℓ_v on V by

$$\ell_v(w) = \langle w, v \rangle$$

Proposition 1.10 *The map* $\mathrm{i} : v \mapsto \ell_v$ *is an antilinear bijection from* V *onto* V°.

Proof. The injectivity is obvious. Since dimensions of V and V° are the same, the map is bijective. $\square$

Constructive proof of bijectivity. For a linear functional ℓ denote $\ell(e_j) = \alpha_j$. Then $\ell(v) = \langle v, h \rangle$, where

$$h = -\sum_{j \leqslant p} \overline{\alpha}_j e_j + \sum_{j > p} \overline{\alpha}_j e_j \qquad \square$$

1.8. Orthogonal complements.

Theorem 1.11 *For any subspace* S *in a pseudo-Euclidean space* V,

a) $(S^\perp)^\perp = S$.

b) $\dim S + \dim S^\perp = \dim V$.

Lemma 1.12 *Let* W *be a finite-dimensional linear space, let* $\ell_1, \dots, \ell_k$ *be linear independent functionals on* W. *Then the subspace*

$$(\ker \ell_1) \cap \dots \cap (\ker \ell_k) \tag{1.6}$$

has codimension k *in* W.

Proof of the lemma. Extend the family ℓ_j to a basis in the dual space. Consider the dual basis $e_1, \dots, e_n \in W$, i.e.,

$$\ell_i(e_j) = \delta_{ij}.$$

The subspace (1.6) is spanned by the vectors $e_{k+1}, \dots, e_n$. $\square$

Problem 1.2 Invent a proof of the lemma avoiding a usage of dual basis. For instance, the lemma remains valid in Bahach spaces.

Proof of Theorem. b) Let $e_1, \dots, e_k$ be a basis in S. Consider the functionals $\ell_j(v) := \langle v, e_j \rangle$. By definition, $S^\perp = \bigcap \ker \ell_j$, therefore $\operatorname{codim} S^\perp = \dim S$

a) Obviously, $(S^\perp)^\perp \supset S$. But these spaces have the same dimension. $\square$

Proposition 1.13 *In a pseudo-Euclidean space, the following statements holds.*

a) The intersection $S \cap S^{\perp}$ coincides with zero(S) and zero$(S^{\perp})$.

b) In particular, if S is regular, then $V = S \oplus S^{\perp}$

c) $(S \cap T)^{\perp} = S^{\perp} + T^{\perp}$

d) $(S + T)^{\perp} = S^{\perp} \cap T^{\perp}$

e) If $S \supset T$, then $S^{\perp} \subset T^{\perp}$.

PROOF. Obvious.

1.9. Isotropic subspaces. Consider a pseudo-Euclidean space V equipped with a form $\langle \cdot, \cdot \rangle$. A subspace $W \subset V$ is

$$\textit{isotropic if} \qquad \langle v, w \rangle = 0 \qquad \textit{for all } v,\, w \in W.$$

Equivalently, W is isotropic if $W \subset W^{\perp}$.

Theorem 1.14 *a) The maximal possible dimension of an isotropic subspace is* $\min(p, q)$*, where (p, q) are the inertia indices.*

b) Any isotropic subspace is contained in an isotropic subspace of the maximal possible dimension

PROOF. a) For definiteness, assume $p \leqslant q$.

Our second canonical form (Theorem 1.9) implies the existence of a p-dimensional isotropic subspace, it is sufficies to consider the linear span of the first p basic vectors.

Assume that there exists an isotropic subspace of dimension $> p$. Then its intersection with a q-dimensional positive subspace is nonzero. This is a contradiction.

b) Let T be an m-dimensional isotropic subspace. Bearing in mind $T^{\perp} \supset T$ and $(T^{\perp})^{\perp} = T$, we observe that $T = \mathit{zero}(T^{\perp})$. Denote by Z a complementary subspace to T in $T^{\perp}$; then $\dim Z = \dim V - 2m$ and $V = Z \oplus Z^{\perp}$.

Consider two cases.

First, let Z be indefinite, $h \in Z$ an isotropic vector. Then $T \oplus \mathbb{C}h$ is an isotropic subspace of dimension $m + 1$.

Second, let Z be sign-definite. Then $\dim Z^{\perp} = 2m$, $T \subset Z^{\perp}$ is an isotropic subspace, and $\dim Z = m$. Hence the inertia indices of our form on $Z^{\perp}$ are (m, m). Therefore the inertia indices of our form on V are $(m, m + \dim Z)$ or $(m + \dim Z, m)$ (according to the sign of Z). Thus, T has maximal possible dimension. ☐

ANOTHER PROOF OF B). Let us repeat the second proof of Theorem 1.9, choosing $h_1, \dots h_m \in T$. This produce a decomposition

$$\Bigl[\bigoplus_{j=1}^{m} (\mathbb{C}h_j \oplus \mathbb{C}g_j)\Bigr] \oplus R,$$

If R is sign-definite, then the isotropic subspace T has the maximal possible dimension. Otherwise, we consider an isotropic vector $\xi \in R$; then $T \oplus \mathbb{C}\xi$ is an isotropic subspace. $\square$

In fact, our second proof implies the following statement (which is a stroger form of Theorem 1.9).

Proposition 1.15 *For an isotropic subspace T in a pseudo-Euclidean space V, one can find a decomposition*

$$V = T \oplus S \oplus Z, \quad \text{where } S \text{ is isotropic, } \dim S = \dim T, \text{ and } (T \oplus S) \perp Z$$

Moreover, each basis $h_1, \dots, h_m \in T$ there is a unique basis $g_1, \dots, g_m \in S$ such that $\langle h_i, g_j \rangle = \delta_{ij}$.

1.10. Duality of isotropic subspaces.

Problem 1.3 Assume $p = q$, let W_1, W_2 be complementary isotropic subspaces. For a vector $h \in W_1$, we define a linear functional on W_2 by $\ell_h(v) = \langle v, h \rangle$. Then the map $h \mapsto \ell_h$ is a bijection of W_1 onto the space of linear functionals on W_2.

1.11. Digression. Similar geometries.

Problem 1.4 a) Convince yourself, that you can repeat the considerations of this section for symmetric bilinear forms over $\mathbb{R}$. For instance, each nondegenerate symmetric bilinear form can be reduced to

$$\langle x, y \rangle = -\sum\nolimits_{i \leqslant p} x_i y_i + \sum\nolimits_{j \geqslant p} x_j y_j$$

We denote by $\mathbb{R}^{p,q}$ the space equipped with this form.Figures 2.1–2.6 are drawn in the spaces $\mathbb{R}^{1,1}$, $\mathbb{R}^{1,2}$, $\mathbb{R}^2$.

b) Solve the same problem for symmetric bilinear forms over $\mathbb{C}$. For instance, each nondegenerate form of this type can be reduced to

$$\langle x, y \rangle = \sum x_k y_k$$

Maximal dimension of an isotropic subspace in $\mathbb{C}^{2n}$ is n, in $\mathbb{C}^{2n+1}$ it is also n.

2.2 Pseudo-unitary group $\mathrm{U}(p, q)$

2.1. The pseudo-unitary groups $\mathrm{U}(p, q)$**.** Let $V \simeq \mathbb{C}^{p,q}$ be a pseudo-Euclidean space. We say that an operator g in the space V is *pseudo-unitary* if it preserves our form, i.e.,[1]

$$\langle vg, wg \rangle = \langle v, w \rangle. \tag{2.1}$$

[1] We can apply an operator g to a matrix-row v, then we write vg, we also can apply g to a vector-column u, then we write gu.

The *pseudo-unitary group*

$$\mathrm{U}(p,q) \qquad \text{or} \qquad \mathrm{U}[V]$$

is the group of all the operators satisfying this condition.

If the scalar product is positive definite (i.e., $p = 0$), then $\mathrm{U}(0,q)$ is the usual *unitary group* $\mathrm{U}(q)$, i.e., the group of unitary operators in a p-dimensional complex Euclidean space. This group is compact[2].

Also $\mathrm{U}(p,0) \simeq \mathrm{U}(p)$.

Proposition 2.1 *If an invertible operator satisfies*

$$\langle vg, vg\rangle = \langle v,v\rangle \qquad \textit{for all } v,$$

then g is pseudo-unitary

PROOF.
$$2\,\mathrm{Re}\langle v,w\rangle = \langle v+w, v+w\rangle - \langle v,v\rangle - \langle w,w\rangle;$$
$$2i\,\mathrm{Im}\langle v,w\rangle = \langle v-iw, v-iw\rangle - \langle v,v\rangle - \langle w,w\rangle.$$

Thus operators preserving scalar squares preserve scalar products. □

2.2. Equations. If $\langle v,w\rangle = vQw^*$, then the condition (2.1) changes to

$$gQg^* = Q. \tag{2.2}$$

In particular, assuming that $\langle\cdot,\cdot\rangle$ has the canonical form (1.2), and representing g as a block matrix $g = \begin{pmatrix} a & b \\ c & d\end{pmatrix}$, we get the equation

$$\begin{pmatrix} a & b \\ c & d\end{pmatrix}\begin{pmatrix} -1_p & 0 \\ 0 & 1_q\end{pmatrix}\begin{pmatrix} a & b \\ c & d\end{pmatrix}^* = \begin{pmatrix} -1_p & 0 \\ 0 & 1_q\end{pmatrix}, \tag{2.3}$$

or

$$aa^* - bb^* = 1, \qquad cc^* - dd^* = -1, \qquad ac^* - bd^* = 0. \tag{2.4}$$

We can also write out these conditions as a system of equations for matrix elements. Denote by r_j the rows of the matrix g. Then the conditions are

— $\langle r_j, r_j\rangle$ is 1 if $j \leqslant p$ and -1 if $j > p$. In particular (if $p \neq 0$),

$$-|a_{11}|^2 - \dots - |a_{1p}|^2 + |b_{11}|^2 + \dots + |b_{1q}|^2 = -1$$

— $\langle r_i, r_j\rangle = 0$ for $i \neq j$. For instance, if $i = 1$, $j = p+1$,

$$-a_{11}\overline{c}_{11} - \dots - a_{1p}\overline{c}_{1p} + b_{11}\overline{d}_{11} + \dots + b_{1q}\overline{d}_{1q} = 0$$

[2]because the equation $gg^* = 1$ implies that all the matrix elements g_{ij} satisfy $|g_{ij}| \leqslant 1$.

Lemma 2.2 *For the standard scalar product, the conditions $g \in \mathrm{U}(p,q)$ and $g^* \in \mathrm{U}(p,q)$ are equivalent, i.e., the equation (2.3) is equivalent to*

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^* \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{2.5}$$

PROOF. We write the condition (2.3) as

$$\left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \right\} \cdot \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix}^* \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \right\} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Next, we transpose two factors in the left-hand side, and reduce our condition to the desired form. □

The conditions (2.5) can be written out as

$$a^*a - c^*c = -1, \qquad b^*b - d^*d = 1; \qquad a^*b - c^*d = 0. \tag{2.6}$$

Lemma 2.3 *Let S be a subspace invariant with respect to a pseudo-unitary operator g. Then the subspace $S^\perp$ also is g-invariant.*

We leave this as an exercise.

Problem 2.1 a) If $g \in \mathrm{U}(p,q)$, then $|\det g| = 1$.

b) $\mathrm{U}(p,q) \simeq \mathrm{U}(q,p)$.

2.3. Self-adjoint operators. Consider a linear operator A in a pseudo-Euclidean space V. An operator $A^\circledast$ is *adjoint* to A if

$$\langle vA, w \rangle = \langle v, wA^\circledast \rangle \qquad \text{for all } v,\, w \in V.$$

REMARK. We use the notation $A^\circledast$ keeping the symbol A^* for adjoint operators in Euclidean spaces and for adjoint matrices in the usual sense. □

Problem 2.2 Prove the existence and uniqueness of the adjoint operator.

An operator A is *self-adjoint* if $A = A^\circledast$.

Problem 2.3 Let the scalar product be given by the standard formula. Then

$$\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \text{ is self-adjoint} \qquad \Longleftrightarrow \qquad \alpha = \alpha^*,\ \delta = \delta^*,\ \beta = -\gamma^* \tag{2.7}$$

Problem 2.4 Let S be a subspace invariant with respect to a self-adjoint operator A. Then $S^\perp$ also is A-invariant.

2.4. The Lie algebra. Now we are going to desribe the Lie algebra $\mathfrak{u} = \mathfrak{u}(p,q)$ of the group $\mathrm{U}(p,q)$. Let $X \in \mathfrak{u}$. By definition of Lie algebra (see Addendum C.3),

$$\exp(\varepsilon X) \in \mathrm{U}(p,q) \text{ for real } \varepsilon$$

Hence

$$\big\langle v\exp(\varepsilon X), w\exp(\varepsilon X)\big\rangle = \big\langle v, w\big\rangle.$$

The left hand side equals

$$\big\langle v(1+\varepsilon X+O(\varepsilon)^2), w(1+\varepsilon X+O(\varepsilon)^2)\big\rangle = \big\langle v,w\big\rangle+\varepsilon\big\langle vX,w\big\rangle+\varepsilon\big\langle v,wX\big\rangle+O(\varepsilon)^2.$$

Therefore X satisfies the condition

$$\langle vX, w\rangle + \langle v, wX\rangle = 0 \qquad \text{or} \qquad X = -X^{\circledast}, \tag{2.8}$$

i.e., the operator X is *anti-self-adjoint*. Equivalently, $\frac{1}{i}X$ is self-adjoint.

Conversely, let X satisfies (2.8). Then $\langle vX^n, w\rangle = \langle v, w(-X)^n\rangle$ and hence

$$\big\langle v\exp(\varepsilon X), w\exp(\varepsilon X)\big\rangle = \Big\langle v\exp(\varepsilon X), w\big(\sum \frac{1}{j!}\varepsilon^j X^j\big)\Big\rangle =$$
$$= \Big\langle v\exp(\varepsilon X)\big(\sum \frac{(-1)^j}{j!}\varepsilon^j X^j\big), w\Big\rangle = \big\langle v\exp(\varepsilon X)\exp(-\varepsilon X), w\big\rangle = \big\langle v,w\big\rangle.$$

Thus we obtained the following theorems.

Theorem 2.4 *The Lie algebra $\mathfrak{u}$ of $\mathrm{U}[V]$ consist of operators X satisfying the anti-self-adjointness condition (2.8).*

Theorem 2.5 *Let X be an anti-self-adjoint operator in a pseudo-Euclidean space. Then its exponent is pseudo-unitary*

Problem 2.5 Let $g \in \mathrm{U}(p,q)$, and $\|g-1\| < 1$, where $\|\cdot\|$ denotes the usual matrix norm. Show directly, that

$$X = \ln g := \sum_{j=0}^{\infty} \frac{(-1)^{j-1}}{j}(g-1)^j \tag{2.9}$$

satisfies (2.8).

If the Hermitian form is written as $\langle v, w\rangle = vQw^*$, then the condition (2.8) of anti-self-adjointness is

$$XQ + QX^* = 0. \tag{2.10}$$

Next, let the form Q be standard, $Q = \begin{pmatrix} -1 & 0 \\ 0 & 1\end{pmatrix}$. Represent $X \in \mathfrak{u}(p,q)$ in a block form $X = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta\end{pmatrix}$. The condition (2.10) changes to

$$\begin{pmatrix} \alpha & \beta \\ \gamma & \delta\end{pmatrix}\begin{pmatrix} -1 & 0 \\ 0 & 1\end{pmatrix} + \begin{pmatrix} -1 & 0 \\ 0 & 1\end{pmatrix}\begin{pmatrix} \alpha^* & \gamma^* \\ \beta^* & \delta^*\end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0\end{pmatrix}$$

Thus, the Lie algebras $\mathfrak{u}(p,q)$ consists of the matrices

$$X = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}, \qquad \alpha + \alpha^* = 0, \qquad \gamma = \beta^*, \qquad \delta + \delta^* = 0 \tag{2.11}$$

Problem 2.6 Verify directly that for each $X, Y \in \mathfrak{u}(p,q)$ their commutator $[X,Y] = XY - YX$ also is an element of $\mathfrak{u}(p,q)$.

Problem 2.7 Write similar conditions for the form defined by the matrix

$$\begin{pmatrix} 0 & 0 & 1_p \\ 0 & 1_{q-p} & 0 \\ 1_p & 0 & 0 \end{pmatrix}$$

Consider individually the special case $p = q$.

2.5. The Witt Theorem.

Theorem 2.6 (Witt Theorem) *Let V be a pseudo-Euclidean space, T, S two subspaces having the same dimension. Let $A : T \mapsto S$ be a bijective operator preserving the scalar product, i.e., $\langle Av, Aw\rangle = \langle v, w\rangle$ for all v, $w \in T$. Then there exists a pseudo-unitary operator $\widetilde{A} : V \to V$ such that the restriction of $\widetilde{A}$ on T is A.*

Problem 2.8 Show that the Witt Theorem is not valid for spaces endowed with degenerate Hermitian forms.

First proof. We consider two cases.

Case-1. Let T be regular, and its inertia indices be (k,l). Then the spaces $T^\perp$, $S^\perp$ have the same inertia indices $(p-k, q-l)$. Therefore one can find a pseudo-unitary bijection $B : T^\perp \to S^\perp$, and we put $\widetilde{A} = A \oplus B$.

Case-2. Let T be singular. Denote by Z some complementary subspace to *zero*(T) in T. Fix a vector $r \in$ *zero*(T) One can find $h \in W$ such that

$$\langle r, h\rangle = 1 \tag{2.12}$$

Next, consider a basis $r_1 = r$, $r_2, \cdots \in$ *zero*(T) such that

$$\langle r_2, h\rangle = \langle r_3, h\rangle = \cdots = 0 \tag{2.13}$$

Let us vary a vector h as $h \to h - q$, where $q \in Z$. Then we can obtain a vector $h^* = h^*_T$ such that $\langle h, z\rangle = 0$ for all $z \in Z$. Indeed, h determines a linear functional

$$\ell(\zeta) = \langle \zeta, h\rangle, \qquad \zeta \in Z$$

on Z. By Proposition 1.10, this functional can be represented as $\ell(\zeta) = \langle \zeta, q\rangle$, $q \in Z$. Therefore the vector $h^* = h - q$ is orthogonal to Z. Since q is also orthogonal to *zero*(T), our variation does not break the conditions (2.12) and (2.13).

Next, for the subspace S we construct a vector $h*_S$ by the same algorithm. Now we can extend $A : T \to S$ by the assumption $\widetilde{A}h*_T = h^*_S$.

Repeating this step-by-step, we obtain a desired extension of A or (more usually) come to Case-1. $\square$

Thus, the Witt theorem is a rather simple statement. Nevertheless, below (in Chapters 9-10) we need in a proof that is valid 'for arbitrary fields'. Our considerations in the Case-1 essentially use the classification of Hermitian forms and we intend to avoid it.

PROOF-II. We need the following lemma which also is a special case of the Witt theorem.

Lemma 2.7 *Fix $\alpha \neq 0$ The group $\mathrm{U}(p,q)$ acts transitively on the set of vectors v satisfying $\langle v, v\rangle = \alpha$.*

PROOF OF LEMMA.[3] Let $\langle v,v\rangle = \langle w,w\rangle = \alpha$. Without a loss of generality, we can assume $\langle v,w\rangle \in \mathbb{R}$. Otherwise, we multiply v by an appropiate constant $e^{i\theta}$.

We are going to construct an operator A taking v to w. If $w = \lambda v$, then we can take A being a multiplication by a scalar. Thus let v, w span a plane. There are two cases.

First, let the plane $\mathbb{C}v \oplus \mathbb{C}w$ be regular. Then we define the following reflection operator A:

$$Av = w, \qquad Aw = v, \qquad Ah = h, \text{ for } h \in (\mathbb{C}v \oplus \mathbb{C}w)^\perp$$

Second, let this plane be singular. Then

$$\det \begin{pmatrix} \langle v,v\rangle & \langle v,w\rangle \\ \langle w,v\rangle & \langle w,w\rangle \end{pmatrix} = 0$$

Hence $\langle v,w\rangle = \pm\alpha$. Since v and $(-v)$ are contained in one $\mathrm{U}(p,q)$-orbit, we can assume $\langle v,w\rangle = +\alpha$. Then $\mathit{zero}(\mathbb{C}v \oplus \mathbb{C}w) = \mathbb{C}(v-w)$. Choose an isotropic vector $h \in \mathbb{C}(v+w)^\perp$ such that $\langle h, v-w\rangle \neq 0$. Then the 3-dimensional space R spanned by v, w, h is regular. We define A by

$$Av = w, \quad Aw = v, \quad Ah = -h, \quad A\zeta = \zeta \quad \text{for } \zeta \in R^\perp \qquad \square$$

Let us continue a proof of the Witt theorem. Let the subspaces T, S be regular. Let $r \in T$ be a non-isotropic vector. Take an operator $U : V \to V$ such that $Ur = Ar$. Let us consider the space $S' = U^{-1}S$ and the operator $A' : T \to S'$ given by $A' = U^{-1}A$. Then $r \in T, S'$ and $A'r = r$.

Now we can work in $(\mathbb{C}r)^\perp$, and dimension of this space is $\dim W - 1$. $\square$

2.6. Grassmannians. Recall that a *Grassmannian* $\mathrm{Gr}_{k,n-k}(V)$ is the space of all k-dimensional subspaces in an n-dimensional linear space V.

[3] Recall, that we intend to avoid a reference to the classification of Hermitian forms. Otherwise, there is nothing to prove.

Theorem 2.8 *a)* $\mathrm{Gr}_{k,l}$ *is a compact* kl*-dimensional complex manifold.*

b) $\mathrm{Gr}_{k,l}(\mathbb{C}^n)$ *is the homogeneous space*

$$\mathrm{Gr}_{k,l} \simeq \mathrm{U}(k+l)/(\mathrm{U}(k) \times \mathrm{U}(l))$$

PROOF. Let us endow the space $\mathbb{C}^{k+l}$ with the standard inner product. Obviously, $\mathrm{Gr}_{k,l}$ is $\mathrm{U}(k+l)$-homogeneous. If g stabilizes a subspace $L \in \mathrm{Gr}_{k,l}$, then it also must stabilize its orthocomlement $L^\perp$. Thus, $g \in \mathrm{U}[L] \times \mathrm{U}[L^\perp]$. This proves b).

A homogeneous space G/H of a Lie group G with respect to a closed subgroup H is a smooth submanifold whose dimension is $(\dim G - \dim H)$. This also proves a), the real dimension of $\mathrm{Gr}_{k,l}$ is

$$\dim \mathrm{U}(k+l) - \dim \mathrm{U}(k) - \dim \mathrm{U}(l) = (k+l)^2 - k^2 - l^2 = 2kl$$

The structure of a complex manifold is not visible in this reasoning. However, the space $\mathrm{Gr}_{k,l}$ admits an action of a larger group $\mathrm{GL}(k+l, \mathbb{C})$. Consider the subspace M generated by the first k basis vectors. Its stabilizer P consists of matrices of the form.

$$\begin{pmatrix} A & B \\ 0 & D \end{pmatrix}$$

Both Lie groups G, P are complex, and hence $\mathrm{Gr}_{k,l} = \mathrm{GL}(k+l, \mathbb{C})/P$ is a complex manifold. □

ANOTHER PROOF. COORDINATES ON $\mathrm{Gr}_{k,l}$. Decompose $\mathbb{C}^{k+l}$ as a direct sum $\mathbb{C}^n = T \oplus S$ of k-dimensional and l-dimensional subspaces. Consider the set Ω_S of all k-dimensional subspaces L having a zero intersection with S. Each $L \in \Omega_S$ is a graph of a linear operator $A_L : T \to S$. Thus, there is one-to-one correspondence between Ω_S and the space of linear operators $T \to S$. Considering all the possible T, S, we obtain an atlas on the manifold $\mathrm{Gr}_{k,l}$. □

To simplify the imagination, we present also the following descriptions of the topology in $\mathrm{Gr}_{k,l}$.

A) We can treat points of the Grassmannian as $(k-1)$-dimensional subspaces in the projective space $\mathbb{CP}^{k+l-1}$. For a given subspace $N \subset \mathbb{CP}^{k+l-1}$ consider some its open neighborhood $\mathcal{O} \subset \mathbb{CP}^{k+l-1}$. Then all the subspaces $L \subset \mathcal{O}$ form a neigborhood of the point N, see Fig.2.7.a.

B) Denote by $B_\rho(v)$ the ball in $\mathbb{C}^{k+l}$ of radius ρ and center v. A sequence $L_j \in \mathrm{Gr}_{k,l}$ converges to L, if for each point $v \in L$ each ball $B_\varepsilon(v)$ has a nonempty intersection with all the L_j starting a certain number N. See Fig.2.7.b.

C) Certainly, the descriptions of a structure of manifold on $\mathrm{Gr}_{k,l}$ also produce descriptions of the topology. For instance, a sequence L_j converges to L if there is a sequence of operators g_j convergent to 1 such that $g_j L = L_j$. This is a rephrasing of defininition of a topology in a quotient space.

2.7. Orbits of $\mathrm{U}(p,q)$ on Grassmannians. A description of orbits of $\mathrm{U}(p,q)$ on the Grassmannians is given by the Witt theorem.

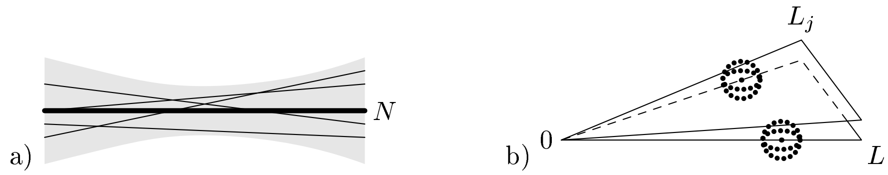

Fig. 2.7: To the definitions of topology in the Grassmannian.
a) A neigborhood of a line N on projective plane and some lines closed to N
b) Convergence of planes in $\mathbb{R}^3$

Problem 2.9 a) Consider a subspace having the inertia indices (r,s). Find the closure of its $\mathrm{U}(p,q)$-orbit.

b) Find all the closed orbits in of $\mathrm{U}(p,q)$ in the Grassmannians.

Now we wish to discuss stabilizers of subspaces in the most important cases.

A) *The stabilizer of a regular k-dimensional subaspace L is* $\mathrm{U}[L]\times\mathrm{U}[L^\perp]$. Indeed, an element of $\mathrm{U}(p,q)$ stabilizing L stabilizes $L^\perp$

Therefore the set of all the regular k-dimensional subspaces in $\mathbb{C}^{p,q}$ is a disjoint union of the homogeneous spaces

$$\bigcup_{r_1+s_1=k,\, r_1+r_2=p,\, s_1+s_2=q} \mathrm{U}(p,q)\Big/\big(\mathrm{U}(r_1,s_1)\times\mathrm{U}(r_2,s_2)\big)$$

One of these spaces, namely $\mathrm{U}(p,q)/\mathrm{U}(p)\times\mathrm{U}(q)$, is the subject of our discussion in the next Section and an important hero of the whole book.

B) Now let ut us desribe the stabilizer $\mathcal{P}$ of an isotropic subspace T. Decompose our space $V=T\oplus Z\oplus S$ as in Proposition 1.15, here S is isotropic, Z is regualar, and $Z=(T\oplus S)^\perp$.

An operator $g\in\mathrm{U}(p,q)$ preserving T, also preserves $T^\perp=T\oplus Z$. As we seen in Proposition 1.15, the spaces T and S are dual. Thus the stabilizer $\mathcal{P}$ of L consists of pseudo-unitary matrices having the following block decomposition with respect to $T\oplus Z\oplus S$:

$$g=\begin{pmatrix} A & B & C\\ 0 & D & E\\ 0 & 0 & A^{*-1}\end{pmatrix}, \qquad \text{where } A\in\mathrm{GL}[L],\ D\in\mathrm{U}[M]$$

2.8. Exercises.

Problem 2.10 Convince yourself, that you can extend all theorems given above to the case of symmetric bilinear forms over $\mathbb{R}$.

2.9. Miscellanea. A) $\mathrm{U}(p,q)$ is a smooth manifold, $\dim U(p,q)=(p+q)^2$, this follows from abstract considerations, see Addendum C.3.

B) The space $\mathrm{U}(p,q)$ is homeomorphic to $\mathrm{U}(p)\times\mathrm{U}(q)\times\mathbb{R}^{2pq}$. Moreover, there is a canonical diffeomorphism (Cartan coordinates) between these spaces, see below Theorem 3.4.

In particular, $\mathrm{U}(p,q)$ is connected, and its fundamental group $\pi_1(\mathrm{U}(p,q))$ is $\mathbb{Z}\oplus\mathbb{Z}$, see Proposition 3.6.2.

C) The center $\mathcal{Z}$ of $\mathrm{U}(p,q)$ consists of matrices $e^{i\theta}\cdot 1$.

D) There are two followng close relatives of $\mathrm{U}(p,q)$ (but we prefer the group $\mathrm{U}(p,q)$ itself as a basic object). The *special pseudo-unitary group* $\mathrm{SU}(p,q)$ is defined as group of pseudo-unitary matrices whose determinant is 1.

The *projective pseudo-unitary group* is $\mathrm{PU}(p,q) := \mathrm{U}(p,q)/\mathcal{Z}$

For instance, the group effectively acting on the Grassmannians is $\mathrm{PU}(p,q)$.

Problem 2.11 a) The group $\mathrm{SU}(2)$ consists of matrices

$$g = \begin{pmatrix} a & b \\ \overline{b} & \overline{a} \end{pmatrix}, \qquad |a|^2 + |b|^2 = 1$$

b) The group $\mathrm{SU}(1,1)$ consists of matrices

$$g = \begin{pmatrix} a & b \\ \overline{b} & \overline{a} \end{pmatrix}, \qquad |a|^2 - |b|^2 = 1$$

c) Define the scalar product in $\mathbb{C}^{1,1}$ by the Hermitian matrix $\begin{pmatrix} 0 & i \\ -i & 0 \end{pmatrix}$. Then the group $\mathrm{SU}(1,1)$ changes to the group of matrices

$$g = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \qquad a,b,c,d \in \mathbb{R},\ \det g = 1$$

that is, $\mathrm{SU}(1,1) \simeq \mathrm{SL}(2,\mathbb{R})$.

2.3 Cartan matrix balls

To be definite, we assume $p \leqslant q$.

3.1. Space of negative subspaces. Consider the pseudo-Euclidean space $V \simeq \mathbb{C}^{p,q}$. Fix a decomposition

$$V = V_- \oplus V_+, \tag{3.1}$$

where V_- is negative, V_+ is positive. It is convenient (but not necessary) to think that

$$V_- = \mathbb{C}^p \oplus 0, \qquad V_+ = 0 \oplus \mathbb{C}^q.$$

Denote $\mathcal{M}_{p,q}$ the set of all p-dimensional negative subspaces in V.

By the Witt theorem, the set $\mathcal{M}_{p,q}$ is $\mathrm{U}(p,q)$-homogeneous. If $g \in \mathrm{U}(p,q)$ stabilizes the subspace $V_- = \mathbb{C}^p \oplus 0$, then it also stabilizes its orthogonal complement $V_+ = 0\oplus\mathbb{C}^q$. Hence the stabilizer K of V_- is isomorphic to $\mathrm{U}(p)\times\mathrm{U}(q)$. Thus $\mathcal{M}_{p,q}$ is a homogeneous space,

$$\mathcal{M}_{p,q} = G/K = \mathrm{U}(p,q)/\mathrm{U}(p)\times\mathrm{U}(q).$$

3.2. Coordinates in $\mathcal{M}_{p,q}$. Denote by $\mathrm{B}_{p,q}$ the space of all $p\times q$ matrices z with norm < 1.

Equivalently, $\mathrm{B}_{p,q}$ is the set of all operators $V_- \to V_+$ with norm < 1. The space V_+ is Euclidean by definition. The space V_- is Euclidean with respect to the inner product $(-\langle\cdot,\cdot\rangle)$. Thus, the norm of an operator $V_- \to V_+$ is well defined.

Theorem 3.1 *a) For $z \in \mathrm{B}_{p,q}$, its graph* $\mathrm{graph}(z) \subset V_- \oplus V_+$ *is a negative p-dimensional subspace.*

b) Each negative p-dimensional subspace in V is a graph of an operator $z : V_- \to V_+$ with $z \in \mathrm{B}_{p,q}$

PROOF. a) Let z in $\mathrm{B}_{p,q}$. Its graph consists of all vectors having the form $v_- \oplus v_- z$. Then

$$\langle v_- \oplus v_- z, v_- \oplus v_- z\rangle_V = \langle v_- \oplus 0, v_- \oplus 0\rangle_V + \langle 0 \oplus v_- z, 0 \oplus v_- z\rangle_V = \\ = -\|v_-\|^2_{V_-} + \|v_+ z\|^2_{V_+}. \quad (3.2)$$

Since $\|z\| < 1$, the latter expression is negative.

b) Let $P \in \mathcal{M}_{p,q}$. Since P is negative, its intersection with a positive subspace V_+ is zero. Hence P is a graph of an operator $z : V_- \to V_+$. Formula (3.2) implies the required statement. □

The group $\mathrm{U}(p,q)$ acts on $\mathcal{M}_{p,q}$ and therefore it acts on $\mathrm{B}_{p,q}$.

Theorem 3.2 *For $g = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$, the corresponding transformation of* $\mathrm{B}_{p,q}$ *is*

$$z \mapsto z^{[g]} := (a + zc)^{-1}(b + zd) \quad (3.3)$$

PROOF. Let $w \in V_-$. The operator g takes a (negative) vector $w \oplus wz \in \mathrm{graph}(z)$ to

$$h := w \oplus wz \begin{pmatrix} a & b \\ c & d \end{pmatrix} = w(a + zc) \oplus w(b + zd)$$

which also must be negative. First of all $w(a + zc) \neq 0$. Otherwise, $h = 0 \oplus w(b + zd)$ and this vector is non-negative.

Next, denote $y = w(a + zc)$ and rewrite

$$w(a + zc) \oplus w(b + zd) = y \oplus y(a + zc)^{-1}(b + zd).$$

Thus the subspace $\mathrm{graph}(z) \cdot g$ is a graph of (3.3). □

Problem 3.1 Derive unvertibility of $(a + zc)$ from (2.4) and (2.6).

EXAMPLE. Let $p = q = 1$. The Grassmannian in $\mathbb{C}^{1,1}$ is nothing but the complex projective line $\mathbb{CP}^1 = \mathbb{C} \cup \infty$. The space $\mathrm{B}_{1,1}$ is the unit disc $|z| < 1$ on $\mathbb{C}$. The group $\mathrm{U}(1,1)$ (see its description in Problem 2.11) acts on $\mathrm{B}_{1,1}$ by the Möbius linear-fractional transformations $z \mapsto (dz + b)/(az + c)$. The action is not effective, since the center of $\mathrm{U}(1,1)$ acts trivially. □

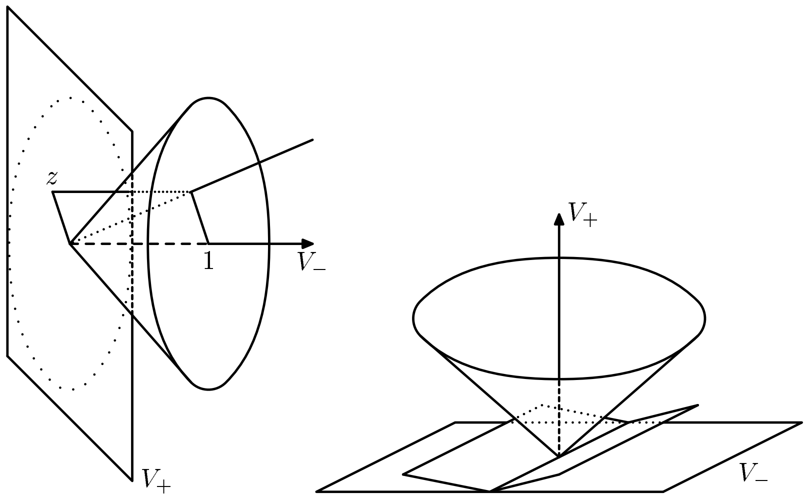

Fig. 2.8: 'Matrix balls' for $\mathbb{R}^{1,2}$ and $\mathbb{R}^{2,1}$. In the first case, negative subspaces are lines inside the isotropic cone. The corresponding 'ball' of 1×2 matrices is a disk. In the second case negative subspaces are planes outside the cone. Another example is given on Fig.2.6, points of the matrix ball for $\mathbb{R}^{2,2}$ are lines lying inside the hyperboloid.

EXAMPLE. Let $p = 1$. The Grassmannian of 1-dimensional subspaces in $\mathbb{C}^{1,q}$ is the complex projective space $\mathbb{CP}^q$. The space $\mathrm{B}_{1,q}$ is the usual ball $|z| < 1$ in $\mathbb{C}^q \subset \mathbb{CP}^q$. □

Problem 3.2 It seems that $\mathrm{B}_{p,q} \simeq \mathrm{B}_{q,p}$. Interpret this on language of Grassmannians. What is the transposition?

3.3. Digression. Different forms of linear-fractional transformations. Consider the Grassmannian $\mathrm{Gr}_{p,q}$ of all p-dimensional subspaces in $\mathbb{C}^{p+q}$. A subspace $P \in \mathrm{Gr}_{p,q}$ of a general position is a graph of an operator $z : \mathbb{C}^p \oplus 0 \to 0 \oplus \mathbb{C}^q$. Clearly, the action of $\mathrm{GL}(p+q, \mathbb{C})$ on the Grassmannian can be written by the same formula (3.3). Note that $(a + xc)^{-1}$ can be non-invertible, therefore the formula is discontinuous.

Problem 3.3 a) Denote

$$g^{-1} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} =: \begin{pmatrix} P & -Q \\ -R & T \end{pmatrix}$$

Then

$$(a + zc)^{-1}(b + zd) = (Pz + Q)(Rz + T)^{-1} \tag{3.4}$$

b) For $g \in \mathrm{U}(p, q)$, simplify the right hand-side of (3.4) as

$$(a + zc)^{-1}(b + zd) = (a^*z + c^*)(b^*z + d^*)^{-1} \tag{3.5}$$

c) In Theorem 3.2, we begin with the right action of the group on vector-rows. Repeat the same considerations for the left action of the group on vector-columns.

There is third (Krein–Shmul'yan) way to write linear fractional transformations, namely

$$z \mapsto K + L(1 - zN)^{-1} zM = K + Lz(1 - Nz)^{-1} M$$

Indeed, we write out

$$\begin{aligned}(a + zc)^{-1}(b + zd) &= a^{-1}(1 + zca^{-1})^{-1}[b + zd] = \\ &= a^{-1}(1 + zca^{-1})^{-1}\big[(1 + zca^{-1})b + z(d - ca^{-1}b)\big] = \\ &= a^{-1}b + a^{-1}(1 + \boxed{z}ca^{-1})^{-1}\boxed{z}(d - ca^{-1}b) \end{aligned} \tag{3.6}$$

Below, in Section 2.8, we will observe that the latter form is more general than the two former forms.

3.4. Transitivity again. As we observe above, the transitivity of the action $\mathrm{U}(p, q)$ on $\mathrm{B}_{p,q}$ follows from the Witt theorem. In what follows we present another proof of transitivity.

For a matrix C with norm < 1 and $\alpha \in \mathbb{C}$, we define

$$(1 + C)^\alpha = 1 + \frac{\alpha}{1!}C + \frac{\alpha(\alpha - 1)}{2!}C^2 + \dots \tag{3.7}$$

Problem 3.4 a) If $\|A\|$, $\mathrm{B}\| < 1$, then

$$(1 - AB)^\alpha A = A(1 - BA)^\alpha \tag{3.8}$$

(expand both sides in Taylor series).

b) If n is positive integer, then (3.8) holds for all A, B. If n is negative integer then (3.8) is valid if $\det(1 - AB) \neq 0$ (apply the analytic continuation).

Denote

$$S(z) = \begin{pmatrix} (1 - zz^*)^{-1/2} & (1 - zz^*)^{-1/2} z \\ z^*(1 - zz^*)^{-1/2} & (1 - z^*z)^{-1/2} \end{pmatrix} \tag{3.9}$$

Proposition 3.3 *a) $S(z) \in \mathrm{U}(p, q)$.*

b) The transformation $S(z)$ transfers 0 to z.

c) $S(z)^* = S(z)$.

d) $S(z)$ is positive definite.

Proof. Statements a), b) can be easily checked by straightforward calculations, c) is obvious. To prove d), we observe

$$S(z) = \begin{pmatrix} (1 - zz^*)^{-1/4} & 0 \\ 0 & (1 - z^*z)^{-1/4} \end{pmatrix} \begin{pmatrix} 1 & z \\ z^* & 1 \end{pmatrix} \begin{pmatrix} (1 - zz^*)^{-1/4} & 0 \\ 0 & (1 - z^*z)^{-1/4} \end{pmatrix}$$

Since $\|z\| < 1$, the middle factor is positive definite. $\square$

Problem 3.5 a) Solve directly the system of equations

$$g \in \mathrm{U}(p,q), \qquad 0^{[g]} = z.$$

and show that g_z is one of its solutions.

b) Find a (unique) solution satisfying $g^2 = 1$.

3.5. The Cartan decomposition.

Theorem 3.4 *a) Each $r \in \mathrm{U}(p,q)$ admits a unique decomposition*

$$r = hS, \quad \text{where } h \in \mathrm{U}(p) \times \mathrm{U}(q),\ S = S^* > 0 \tag{3.10}$$

b) Each $r \in \mathrm{U}(p,q)$ admits a unique decomposition

$$r = h\exp\{X\}, \qquad h \in \mathrm{U}(p) \times \mathrm{U}(q),\ X = X^* \in \mathfrak{u}(p,q) \tag{3.11}$$

REMARK. Having in mind (see (2.11)), we see that X has the form

$$X = \begin{pmatrix} 0 & \beta \\ \beta^* & 0 \end{pmatrix} \qquad \square$$

In fact, we wish to prove a stronger variant of theorem

Theorem 3.5 *Each $r \in \mathrm{U}(p,q)$ admits a unique decomposition*

$$r = hS(z), \quad \text{where } h \in \mathrm{U}(p) \times \mathrm{U}(q),\ z \in \mathrm{B}_{p,q} \tag{3.12}$$

where $S(z)$ is the form (3.9).

PROOF OF THEOREM 3.5. Put $z = 0^{[r]}$, then $h := S(z)^{-1}r$ stabilizes the point 0, and hence $h \in \mathrm{U}(p) \times \mathrm{U}(q)$. $\square$

PROOF OF THEOREM 3.4. a) Consider the polar decomposition $r = hS$ in the usual sense, see Theorem 1.1.3 and its proof. We have $h \in \mathrm{U}(p+q)$, $S = S^* > 0$. The polar decomposition is unique, but a priory, it is not obvious that h, $S \in \mathrm{U}(p,q)$.

On another hand, Theorem 3.5 provides such decomposition with positive $S(z)$. Therefore h in (3.12) and (3.10) are the same. Thus $S = S(z)$.

b) A positive self-adjoint matrix S has a canonically defined logarithm $\ln S$,

$$\ln S = \ln \left\{ U \begin{pmatrix} \lambda_1 & 0 & \dots \\ 0 & \lambda_2 & \dots \\ \vdots & \vdots & \ddots \end{pmatrix} U^{-1} \right\} := U \begin{pmatrix} \ln\lambda_1 & 0 & \dots \\ 0 & \ln\lambda_2 & \dots \\ \vdots & \vdots & \ddots \end{pmatrix} U^{-1}$$

where U is a unitary matrix.

Our $S(z) \in \mathrm{U}(p,q)$ is positive. If it is sufficiently close to 1, then $\ln S \in \mathfrak{u}(p,q)$. Since logarithm is analytic, we get $\ln S \in \mathfrak{u}(p,q)$ for all positive self-adjoint $S \in \mathrm{U}(p,q)$. $\square$

Problem 3.6 It is interesting to prove Theorem 3.4 avoiding Theorem 3.5. By Lemma 2.2, we have $r^* \in \mathrm{U}(p,q)$. Since $r^*r > 0$, it has positive eigenvalues and no Jordan blocks. Why $\sqrt{r^*r} \in \mathrm{U}(p,q)$? See a discussion of congugacy classes in Section 2.6.

Problem 3.7 Each element of $\mathrm{U}(p,q)$ admits a representation in the form

$$g = h_1 A(t) h_2$$

where h_1, $h_2 \in \mathrm{U}(p) \times \mathrm{U}(q)$ and

$$A(t) := \left(\begin{array}{ccc|cccccc}
\cosh t_1 & 0 & \dots & \sinh t_1 & 0 & \dots & 0 & 0 & \dots \\
0 & \cosh t_2 & \dots & 0 & \sinh t_2 & \dots & 0 & 0 & \dots \\
\vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots\ddots & \\
\hline
\sinh t_1 & 0 & \dots & \cosh t_1 & 0 & \dots & 0 & 0 & \dots \\
0 & \sinh t_2 & \dots & 0 & \cosh t_2 & \dots & 0 & 0 & \dots \\
\vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots\ddots & \\
0 & 0 & \dots & 0 & 0 & \dots & 1 & 0 & \dots \\
0 & 0 & \dots & 0 & 0 & \dots & 0 & 1 & \dots \\
\vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots & \vdots & \ddots
\end{array}\right)$$

where

$$t_1 \geqslant t_2 \geqslant \ldots \geqslant t_p \geqslant 0$$

The numbers t_j are uniquely determined by the matrix g. Are h_1, h_2 unique?

It is more pleasant to write the Hermitian form as (1.4), then $A(t)$ becomes diagonal.

3.6. The boundary of $\mathrm{B}_{p,q}$. Let $p \leqslant q$. Denote by $\overline{\mathcal{M}}_{p,q}$ the set of all the semi-negative subspaces in $\mathbb{C}^{p,q}$. Denote by $\mathcal{N}_\alpha$ the set of semi-negative subspaces P, such that the rank of $\langle \cdot, \cdot \rangle$ on P is α. By definition, $\mathcal{N}_p = \mathcal{M}_{p,q}$.

By the Witt Theorem, the sets $\mathcal{N}_\alpha$ are precisely the orbits of $\mathrm{U}(p,q)$ on $\overline{\mathcal{M}}_{p,q}$. Obviously, the closure $\overline{\mathcal{N}}_\alpha$ of an orbit $\mathcal{N}_\alpha$

$$\overline{\mathcal{N}}_\alpha = \cup_{\beta \leqslant \alpha} \mathcal{N}_\beta$$

Denote by $\overline{\mathrm{B}}_{p,q}$ the set of complex $p \times q$ matrices having the norm $\leqslant 1$ In Subsection 3.2 we constructed a bijection $\mathcal{M}_{p,q} \leftrightarrow \mathrm{B}_{p,q}$. Obviously, it can be extended to a bijection $\overline{\mathcal{M}}_{p,q} \leftrightarrow \overline{\mathrm{B}}_{p,q}$.

Denote by $\mathcal{O}_\alpha$ the subset in $\overline{\mathrm{B}}_{p,q}$ defined by the condition

$$z \in \mathcal{O}_\alpha \qquad \text{if} \qquad \mathrm{rk}(1 - zz^*) = \alpha \tag{3.13}$$

Proposition 3.6 *The* $\mathrm{U}(p,q)$*-orbits* $\mathcal{N}_\alpha$ *correspond to the subsets* $\mathcal{O}_\alpha \subset \overline{\mathrm{B}}_{p,q}$*. In particular, the sets* $\mathcal{O}_{p,q}$ *are* $\mathrm{U}(p,q)$*-invariant.*

PROOF. Let a vector $h \in \mathbb{C}^p$. Since the operator $(1 - zz^*)$ is semipositive,

$$\begin{aligned} h(1 - zz^*) = 0 \iff 0 = \langle h, h(1 - zz^*) \rangle_{\mathbb{C}^p} = \\ = \langle h, h \rangle_{\mathbb{C}^p} - \langle zh, zh \rangle_{\mathbb{C}^q} \iff h \oplus hz \in \mathbb{C}^{p,q} \text{ is isotropic} \end{aligned}$$

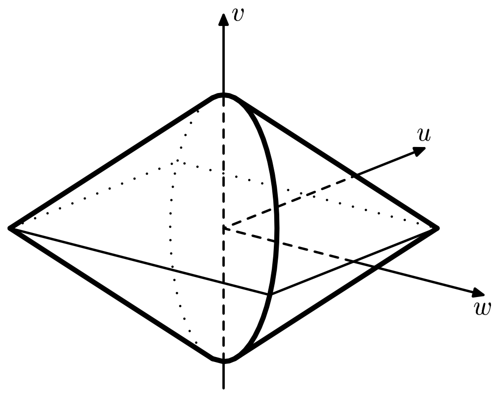

Fig. 2.9: Boundaries of a matrix balls have a rather complicated structure. Unfortunately it is difficult to imagine it because $\dim \mathrm{B}_{pq} = 2pq$ (and the simplest nontrivial case is $p = q = 2$). On the figure we draw the set of 2×2 real symmetric matrices $z = \begin{pmatrix} u & v \\ v & w \end{pmatrix}$ satisfying the condition $\|z\| \leqslant 1$.
The conic surfaces are the sets of matrices having an eigenvalue 1 or (-1) respectively.
The vertices of the cones correspond to the matrices $z = \pm\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$
The circle in the intersection of cones corresponds to matrices z, whose eigenvalues are 1 and (-1).

Let Pbe the corresponding semi-positive subspace. Since h is isotropic, it follows that $h \in \mathit{zero}\, P$. Thus, we obtain a bijection between $\ker(1 - zz^*)$ and $\mathit{zero}(P)$. $\square$

Problem 3.8 Prove directly, starting formula (3.3), that the sets $\mathrm{rk}(1 - zz^*) = \alpha$ are $\mathrm{U}(p, q)$-invariant.

Problem 3.9 (exotic) a) Show directly, that $\mathrm{rk}(z - u)$, where $z,\ u \in \overline{\mathrm{B}}_{p,q}$, is an invariant of $\mathrm{U}(p, q)$-action, i.e., $\mathrm{rk}(z^{[g]} - u^{[g]}) = \mathrm{rk}(z - u)$.

b) The *integer distance* $\rho(z, u) = \mathrm{rk}(z - u)$ satisfies the axioms of metric spaces.

c) Interpret this metric in the terms of $\overline{\mathcal{M}}_{p,q}$.

3.7. Digression. Biholomorphic automorphisms of $\mathrm{B}_{p,q}$.

Theorem 3.7 (Elie Cartan) *If $p \neq q$, then the group of biholomorphic[4] transformations of $\mathrm{B}_{p,q}$ coincides with $\mathrm{U}(p, q)$. If $p = q$, then each biholomorphic transformation is an element of $\mathrm{U}(p, q)$ or a composition of a such map with $z \mapsto z^t$.*

Sketch of proof. Let us denote by Γ the group of all biholomorphic transformations of $\mathrm{B}_{p,q}$. Let Γ_0 be the stabilizer of the point 0. It sufficies to show that Γ_0 coincides with $\mathrm{U}(p) \times \mathrm{U}(q)$.

[4]Let $\Omega \subset \mathbb{C}^N$ be an open domain. A map $\psi : \Omega \to \Omega$ is *biholomorphic* if ψ is holomorphic and bijective (and hence ψ^{-1} also is holomorphic)

A) Each partial derivative $\partial^\alpha \gamma_\mu(0)$ is uniformly bounded for all $\gamma \in \Gamma_0$ (since it can be evaluated as a Cauchy integral, see any textbook in the complex analysis).

B) Let $\gamma \in \Gamma_0$. Denote by R_θ the transformation $z \mapsto e^{i\theta} z$. If γ is not a linear transformation, then $\gamma_\theta := R_\theta \circ \gamma \circ R_\theta^{-1} \neq \gamma$. Denote $\nu := \gamma \circ \gamma_\theta^{-1}$.

C) Expand $\nu(z)$ in Taylor series,

$$\nu(z) = z + \sigma_2(z) + \sigma_3(z) + \dots$$

here σ_j are homogeneous forms of order j. Then

$$\underbrace{\nu \circ \nu \cdots \circ \nu}_{n \text{ times}} = z + n\sigma_2(z) + \dots$$

By A), it follows that $\sigma_2(z) = 0$. Writing out $\gamma \circ \cdots \circ \gamma$ again, we obtain $\sigma_3(z) = 0$ etc.

Therefore $\nu(z) = z$ and γ is a linear transformation.

D) It remains to describe linear transformations of $\mathbb{C}^{pq}$ preserving the compact set $\overline{\mathrm{B}}_{p,q}$. There are many ways to do this, our approach is based on examination of the boundary of $\mathrm{B}_{p,q}$.

E) The cases $p = 1$ or $q = 1$ are trivial (because $\mathrm{B}_{1,q}$ is the usual ball). Thus, we assume p, $q > 1$. Consider the $\mathrm{U}(p,q)$-orbit $\mathcal{O}_1 \subset \overline{\mathrm{B}}_{p,q}$, see (3.13).

Lemma 3.8 *The orbit $\mathcal{O}_1 \subset \mathrm{B}_{P,q}$ is a disjoint union of flat 1-dimensional (open) complex disks lying on affine lines of the form $a + s \cdot H$, where H is a rank 1 matrix and s ranges in $\mathbb{C}$. Moreover, all the rank 1 matrices H arise in this context.*

PROOF. Let $z \in \mathcal{O}_1$. Then there are bases $e_k \in \mathbb{C}^p$, $f_l \in \mathbb{C}^q$ such that

$$ze_1 = \zeta e_1, \qquad ze_m = e_m \text{ for } m > 1$$

the corresponding line consists of operator of same form with arbitrary ζ; the complex disk is $|\zeta| < 1$ (its boundary is contained in $\mathcal{O}_0$). □

Our linear transformation must take the lines described in Lemma to lines having the same form.

F) Thus, our linear transformation γ preserves the cone of rank one matrices.

Lemma 3.9 *If $p \neq q$, then all the linear transformations of* $\mathrm{Mat}(p,q)$ *preserving the cone of rank 1 matrices have the form $z \mapsto azb$, where $a \in \mathrm{GL}(p,\mathbb{C})$, $b \in \mathrm{GL}(q,\mathbb{C})$. If $p = q$, then also transforms $z \mapsto az^t b$ arise.*

PROOF. Denote our cone by Ξ. Each point of the cone is of the form $s \cdot t$ where s is a matrix-column, t is a matrix-row. Consider the projectivized cone $\mathbb{P}\Xi$ (that is, we consider points of Ξ to within a proportionality). Obviously, $\mathbb{P}\Xi$ is a product of projective spaces.

$$\mathbb{P}\Xi = \mathbb{CP}^{p-1} \times \mathbb{CP}^{q-1}$$

The group of biholomorphic maps of $\mathbb{CP}^{k-1}$ is $\mathrm{PGL}(k,\mathbb{C})$. We leave the remaining details for an enthusiastic reader. □

G) The set of limit points of $\mathcal{O}_1$ is $\mathcal{O}_0$, the latter is the set of isometric embeddings $\mathbb{C}^p \to \mathbb{C}^q$. The transformation $z \mapsto azd$ must preserve $\mathcal{O}_0$. More-or-less evidently, $a \in \mathrm{U}(p)$, $b \in \mathrm{U}(q)$. □

Problem 3.10 Each $\mathrm{U}(p,q)$-orbit $\mathcal{O}_\alpha$ is a disjoint union of flat pieces. Describe them. In fact all these pieces are copies of the ball $\mathrm{B}_{\alpha,q-p+\alpha}$.

3.8. Siegel matrix wedges[5]**.** For $p = q$, there is an another nice way to introduce coordinates on the set $\mathcal{M}_{q,q}$.

Consider the pseudo-Euclidean space $H := \mathbb{C}^{q,q}$. Take a basis $e_1, \dots, e_q$, $f_1, \dots, f_q$ such that

$$\langle e_i, e_j\rangle = \langle f_i, f_j\rangle = 0, \qquad \langle e_i, f_j\rangle = -\delta_{ij} \tag{3.14}$$

We denote by H^- (resp. H^+) the space generated by e_j (resp. by f_j). By definition, these two spaces are isotropic.

A matrix $g = \begin{pmatrix} A & B \\ C & D\end{pmatrix}$ is an element of $\mathrm{U}[H] = \mathrm{U}(q,q)$ if

$$\begin{pmatrix} A & B \\ C & D\end{pmatrix}\begin{pmatrix} 0 & 1 \\ 1 & 0\end{pmatrix}\begin{pmatrix} A & B \\ C & D\end{pmatrix}^* = \begin{pmatrix} 0 & 1 \\ 1 & 0\end{pmatrix} \tag{3.15}$$

Proposition 3.10 *A subspace P is an element of $\mathcal{M}_{q,q}$ iff it is a graph of an operator $T : H^- \to H^+$ satisfying the condition*

$$(T + T^*) > 0$$

Proof. Since H^+ is isotropic and P is negative, it follows that $H^+ \cap P = 0$. Therefore P is a graph of an operator $T : H^- \to H^+$. Consider a point $v_- \oplus v_- T \in P$. Then

$$\langle v_- \oplus v_- T, v_- \oplus v_- T\rangle = -v_- T^* v_-^* - v_- T v_- = -v_-(T^* + T)v_-^* -$$

This expressions is negative iff $T + T^* > 0$ $\square$

We denote by W_q the wedge of all $q \times q$ matrices T satisfying $\mathrm{Re}\, T > 0$.

The group $\mathrm{U}(q,q)$ acts on W_q by the transformations

$$T \mapsto (A + TC)^{-1}(B + TD)$$

Since $\mathrm{B}_{q,q}$ and W_q are identified with the same set $\mathcal{M}_{q,q}$, one can find a (non-canonical) identification $\mathrm{B}_{q,q} \simeq \mathrm{W}_q$. One of possible maps is the *Cayley transform*

$$u = (1 - T)/(1 + T) \quad \Longleftrightarrow \quad T = (1 - u)/(1 + u), \quad \text{where } T \in \mathrm{W}_q,\ u \in \mathrm{B}_{q,q}.$$

Remark. The closure $\overline{\mathrm{W}}_q$ of W_q is non-compact, and hence the map $\mathcal{M}_{q,q} \to \mathrm{W}_q$ can not be extended to a bijection[6] $\overline{\mathcal{M}}_{q,q} \to \overline{\mathrm{W}}_q$.

[5] Another term is the *Siegel half-plane*

[6] In fact, Gaussian vectors in Section 1.1.3 are enumerated by points of a matrix wedge and remaining Gaussian distributions correspond to remaining points of boundary of a matrix ball.

Problem 3.11 Thus, $\mathrm{U}(q,q)$ acts on the boundary of W_q discontinuously. Inspite of this, describe 'orbits'.

3.9. Digression. Embeddings $\mathrm{B}_{p,q} \to \mathrm{U}(p,q)$**.**

A) The map $g \mapsto g^*$ is an antiinvolution [7] of the group $\mathrm{U}(p,q)$. Therefore the group $\mathrm{U}(p,q)$ acts on itself by the transformations $g : r \mapsto g^* r g$.

Problem 3.12 The stabilizer of the point $r = 1$ is the group $\mathrm{U}(p) \times \mathrm{U}(q) \subset \mathrm{U}(p,q)$.

Thus the $\mathrm{U}(p,q)$-orbit $\mathcal{R}$ of the point 1 is the homogeneous space $\mathrm{U}(p,q)/\mathrm{U}(p) \times \mathrm{U}(q)$. Hence it must be a $\mathrm{U}(p,q)$-equivariant map $\mathrm{B}_{p,q} \to \mathrm{U}(p,q)$.

Problem 3.13 a) Show that the map $\mathrm{B}_{p,q} \to \mathrm{U}(p,q)$ given by $z \mapsto S(z)^2$ commutes with the action of $\mathrm{U}(p,q)$ (for instance, see proof of Theorem 3.4).

b) Show that its image consists of pseudo-unitary matrices that are self-adjoint positive in the usual sense.

B) Now let $\mathrm{U}(p,q)$ acts on itself by conjugations $r \mapsto g^{-1} r g$. Consider the manifold $\mathfrak{S} \subset \mathrm{U}(p,q)$, whose elements are matrices satisfying $g^2 = 1$. Eigenvalues of such matrices are ± 1.

Lemma 3.11 *The space $\mathfrak{S}$ is in one-to-one correspondence with the set of all the regular subspaces $W \subset \mathbb{C}^{p,q}$.*

PROOF. For a given W define the map g (*reflection in* W) given by

$$g(w+y) = w - y, \quad \text{where } w \in W,\ y \in W^{\perp}$$

Conversely, for a given $g \in \mathfrak{S}$ we consider the subspace $W = \ker(g-1)$ and $W^{\perp} = \ker(g+1)$. □

Now we are ready to construct the canonical embedding $\mathcal{M}_{p,q} \to \mathrm{U}(p,q)$. Let W be a negative subspace of the maximal possible dimension ($\dim W = p$). Then the corresponding $g \in \mathrm{U}(p,q)$ is the reflection in W.

Thus the manifold $\mathfrak{S}$ is a disjoint union of the homogeneous spaces

$$\mathfrak{S} = \bigcup_{k \leqslant p,\, l \leqslant q} \mathrm{U}(p,q) \big/ \mathrm{U}(k,l) \times \mathrm{U}(p-k, q-l)$$

All these spaces are interesting (and may be very interesting) objects[8] but our hero $\mathrm{U}(p,q)/\mathrm{U}(p) \times \mathrm{U}(q)$ is distinguished and the most interesting among them.

Problem 3.14 Find a relation between the constructions A) and B) (in fact, they more-or-less coincide).

[7] An antiinvolution is a map satisfying $g^{**} = g$, $(hg)^* = g^* h^*$

[8] they are examples of pseudo-Riemannian symmetric spaces, see Subsection 12.1

2.4 The space $\mathrm{U}(n)$. Cell decomposition

4.1. Coordinates on the Grassmannian of maximal isotropic subspaces. First, let $p = q$. We keep the notation of Subsection 3.8. Recall that H^+ and H^- are complementary maximal isotropic subspaces in $\mathbb{C}^{p,p}$. Let R be an p-dimensional isotropic subspace. *Assume* that $R \cap H^+ = 0$, under this condition R is a graph of an operator $T : H^- \to H^+$.

Theorem 4.1 *The following conditions are equivalent*

1) R is isotropic

2) $T = -T^*$ (i.e., T is anti-Hermitian)

PROOF. See the proof of Proposition 3.10. □

4.2. Isotropic subspaces and unitary matrices. Next, decompose the space $\mathbb{C}^{p,p}$ into an orthogonal sum of a negative subspace V_- and a positive subspace V_+.

Theorem 4.2 *The following conditions are equivalent*

1) R is a p-dimensional isotropic subspace

2) R is a graph of an operator $z : V_- \to V_+$ satisfying the condition $zz^* = 1$ (in other words, $z \in \mathrm{U}(n)$).

We formulate this theorem due to its importance, but it was proved above in Subsection 1.7.7).

Corollary 4.3 *The group* $\mathrm{U}(q, q)$ *acts on the* space $\mathrm{U}(q)$ *by the transformations*

$$g : z \mapsto z^{[g]} = (a + zc)^{-1}(b + zd), \qquad z \in \mathrm{U}(n), \textit{ and } g \textit{ satisfies (2.4)}$$

REMARK. In this context, behaviors of the groups $\mathrm{U}(q)$ and $\mathrm{SU}(q)$ are completely different. The group $\mathrm{SU}(q) \times \mathrm{SU}(q)$ acts on $\mathrm{SU}(q)$ by left and right multiplication. But it is a maximal Lie group acting on $\mathrm{SU}(q)$ and there is no analog of the previous corollary. □

4.3. Stiefel manifolds and their groups of symmetries. Let $p \leqslant q$. Recall that a complex *Stiefel manifold* $\mathrm{Sti}(p, q)$ is the space, whose points are orthonormal frames $e_1, \dots, e_p$ in $\mathbb{C}^q$. We are going to describe 3 groups acting transitively on this space.

A) The group $\mathrm{U}(q)$ acts on $\mathbb{C}^q$, therefore it acts also on $\mathrm{Sti}(p, q)$. Evidentely the action is transitive and a stabilizer of a point is isomorphic to $\mathrm{U}(q - p)$, i.e.,

$$\mathrm{Sti}(p, q) \simeq \mathrm{U}(q)/\mathrm{U}(q - p)$$

B) Equivalently, $\mathrm{Sti}(p, q)$ is the space of isometric embeddings $z : \mathbb{C}^p \to \mathbb{C}^q$. The group $\mathrm{U}(p)$ acts on $\mathbb{C}^p$ and the group $\mathrm{U}(q)$ acts on $\mathbb{C}^q$. Therefore $\mathrm{U}(p) \times \mathrm{U}(q)$ acts on $\mathrm{Sti}(p, q)$. Thus, $\mathrm{Sti}(p, q)$ is a homogeneous space

$$\mathrm{Sti}(p, q) = \Big[\mathrm{U}(p) \times \mathrm{U}(q)\Big] \Big/ \Big[\mathrm{U}(p) \times \mathrm{U}(q - p)\Big]$$

where the stabilizer $\mathrm{U}(p) \times \mathrm{U}(q-p)$ consists of pairs

$$\left\{ h\,,\ \begin{pmatrix} h & 0 \\ 0 & r \end{pmatrix} \right\} \in \mathrm{U}(p) \times \mathrm{U}(q)$$

C) Thus we enlarged the group of symmetries of $\mathrm{Sti}(p,q)$. But our previous considerations show that $\mathrm{Sti}(p,q)$ admits a group of symmetries that is larger than $\mathrm{U}(p) \times \mathrm{U}(q)$. Indeed, $z \in \mathrm{Sti}(p,q)$ satisfies $zz^* = 1$. This is point of the $\mathrm{U}(p,q)$-orbit $\mathcal{O}_0$ on the boundary in the matrix ball, see Proposition 3.6. Thus, there is an one-to-one correspondence

$$\mathrm{Sti}(p,q) \quad \longleftrightarrow \quad \text{set of all maximal isotropic subspaces in } \mathbb{C}^{p,q}. \tag{4.1}$$

Proposition 4.4 *The group* $\mathrm{U}(p,q)$ *acts on Stiefel manifold* $\mathrm{Sti}(p,q)$ *by the transformations* $z \mapsto (a+zc)^{-1}(b+zd)$.

4.4. Digression. The Dynkin cell decomposition of $\mathrm{U}(n)$. In Theorem 4.1 we defined a coordinate system on $\mathrm{U}(n)$, but the this system does not cover the whole group. Evidentely, we also get an atlas on $\mathrm{U}(n)$, whose maps are enumerated by pairs of transversal maximal isotropic subspaces.

Nevertheless it remains a problem of description (and parametrization) of the complement to the image of our map. In what follows we briefly discuss this question.

Let $\mathbb{C}^n$ be the standard complex Euclidean space with the standard basis $e_1, \dots, e_n$. Denote by $\mathbb{C}^k$ the subspace spanned by the first k basic vectors $e_1, \dots, e_k$.

We say that a *complex reflection*[9] in a vector v is a linear map in $\mathbb{C}^n$ of the form

$$S[v; e^{i\varphi}]x = x + (e^{i\varphi} - 1)\langle x, v\rangle \cdot v \tag{4.2}$$

where a vector v satisfies $\langle v, v\rangle = 1$, and $\varphi \in (0, 2\varphi)$; note that we forbid $e^{i\varphi} = 1$, otherwise the map is identical.

REMARK. Let us complete a vector v to an orthonormal basis v, $f_2, \dots, f_n$. Then

$$S[v; e^{i\varphi}]\, v = \lambda v, \qquad S[v, e^{i\varphi}]\, f_j = f_j, \quad \text{for } j \geqslant 2 \qquad \square$$

Thus a reflection is a unitary operator having a unique eigenvalue $\neq 1$.

Problem 4.1 a) Let $\langle x, x\rangle = \langle y, y\rangle = 1$, let $y \neq x$. Then there is a unique reflection S such that $Sx = y$.

b) The inverse operator to a reflection is a reflection.

Denote by $\mathcal{R}$ the set of all the reflections. Let $\theta \in \mathbb{C}$, $|\theta| = 1$. Evidentely, $S[\theta v, e^{i\varphi}] = S[v, e^{i\varphi}]$. Therefore $\mathcal{R}$ is diffeomorphic to the space

$$\mathcal{R} \simeq \mathbb{CP}^{n-1} \times (0, 2\pi)$$

where $\mathbb{CP}^{n-1}$ is the complex projective space.

[9] For instance, a 'reflection' in $\mathbb{C}^1$ is a rotation in a human language. But the term 'reflection' is generally accepted.

Denote by $\mathcal{R}_k$ the set of all the reflections in vectors $v \in \mathbb{C}^k \setminus \mathbb{C}^{k-1}$. By definition $\mathcal{R}$ is a disjoint union

$$\mathcal{R} = \bigcup_{j=1}^{n} \mathcal{R}_j$$

Each set $\mathcal{R}_j$ is (canonically) diffeomorphic to

$$\mathcal{R}_j \simeq \left(\mathbb{CP}^{j-1} \setminus \mathbb{CP}^{j-2}\right) \times (0, 2\pi) \simeq \mathbb{C}^{j-1} \times (0, 2\pi)$$

Now, we are going to show that *each $g \in \mathrm{U}(n)$ admits a canonical decomposition in a product of $\leqslant n$ reflections.*

Denote $g_n := g$. Let $e_1, \dots, e_n$ denote the standard basis in $\mathbb{C}^n$ as above. First, suppose $ge_n \neq e_n$. Consider the unique reflection S_n such that $S_n e_n = g_n e_n$. Put $g_{n-1} := S_n^{-1} g_n$. If $g_n e_n = e_n$, we simply put $g_{n-1} := g_n$.

Next, we consider $g_{n-1}e_{n-1}$. If $g_{n-1}e_{n-1} \neq e_{n-1}$, then we take the reflection S_{n-1} such that $S_{n-1}e_{n-1} = g_{n-1}e_{n-1}$. It is important, that $S_{n-1}e_n = e_n$ (because e_n is orthogonal to both e_{n-1}, $g_{n-1}e_{n-1}$). Put $g_{n-2} := S_{n-2}^{-1} g_1$. If $g_{n-1}e_{n-1} = e_{n-1}$, we put $g_{n-2} := g_{n-1}$. Etc.

For an element $g \in \mathrm{U}(n)$ of general position, we obtain

$$g = S_1 S_2 \dots S_n$$

Moreover, $S_j \in \mathcal{R}^j$. For exceptional elements g some of factors S_j can be omitted. More precisely, we obtained the following theorem.

Theorem 4.5 *For each $g \in \mathrm{U}(n)$ there is a unique collection $k_1 < k_2 < \dots < k_l$ and unique $S_{k_j} \in \mathcal{R}_{k_j}$ such that*

$$g = S_{k_1} S_{k_2} \dots S_{k_l} \tag{4.3}$$

(the empty collection $\{\varnothing\}$ corresponds to $g = 1$).

Thus, we get a representation of $\mathrm{U}(n)$ as a disjoint union of sets

$$\mathrm{U}(n) = \bigcup_{k_1 < \dots < k_l} \mathcal{R}_{k_1} \times \dots \times \mathcal{R}_{k_l} =: \bigcup_{k_1 < \dots < k_l} \mathcal{D}_{k_1, \dots, k_l} \tag{4.4}$$

Each set $\mathcal{D}_{k_1,\dots,k_l}$ is (canonically) diffeomorphic to $\mathbb{C}^{\sum k_j} \times (0, 2\pi)^l$.

It is easy to show that this is indeed a CW-complex, and all the boundary operators are zero. This imples the following theorem.

Theorem 4.6 *The manifolds (4.4) form a basis of homologies of the topological space* $\mathrm{U}(n)$.

4.5. Digression. Dynkin cells. Continuation. Denote

$$\widetilde{\mathbb{C}}^j := \mathbb{C}e_{n-j+1} \oplus \dots \oplus \mathbb{C}e_n \subset \mathbb{C}^n$$

For $g \in \mathrm{U}(n)$, denote by $\mathrm{Fix}(g)$ the space of fixed vectors of g, i.e., $x \in \mathrm{Fix}(g)$ if $gx = x$. For a collection of integers

$$0 \leqslant \alpha_1 \leqslant \alpha_2 \leqslant \dots \leqslant \alpha_n \leqslant n$$

denote by $\mathcal{B}_{\alpha_1,\dots,\alpha_n}$ the set of all $g \in \mathrm{U}(n)$ such that

$$\dim \mathrm{Fix}(g) \cap \widetilde{\mathbb{C}}^j = \alpha_j$$

(by definition, $\alpha_{j+1} - \alpha_j$ is 0 or 1).

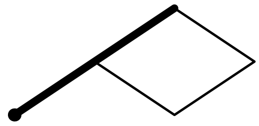

Fig. 2.10: A flag in $\mathbb{R}^3$

Theorem 4.7 *The decompositions*

$$\mathrm{U}(n) = \bigcup \mathcal{D}_{k_1,\dots,k_l} \qquad \textit{and} \qquad \mathrm{U}(n) = \bigcup \mathcal{B}_{\alpha_1,\dots,\alpha_n}$$

coincide. The bijection of the sets of indices $\{k_j\} \leftrightarrow \{\alpha_\tau\}$ *is given by*

$$\forall \tau: \qquad \tau - \alpha_\tau \quad \textit{equals number of } k_j > n - \tau + 1$$

Lemma 4.8 *Decompose* $g \in \mathrm{U}(n)$ *as a product of reflections* $S_{k_1} \dots S_{k_l}$, *where* S_{k_j} *is a reflection in* $v_j \in \mathbb{C}^{k_j} \setminus \mathbb{C}^{k_j-1}$. *Then*

$$\mathrm{Fix}(g) = \bigcap_j \mathrm{Fix}(S_{k_j}) = \bigcap_j (\mathbb{C}v_j)^\perp \tag{4.5}$$

Proof of Lemma. The inclusion $\supset$ is obvious, because

$$\mathrm{Fix}(AB) \supset \mathrm{Fix}(A) \cap \mathrm{Fix}(B)$$

Next, note that for $S_{k_j} \in \mathcal{R}_{k_j}$,

$$\forall x \in \mathbb{C}^n: \qquad S_{k_j}x - x \neq 0 \quad \Longrightarrow \quad S_{k_j}x - x \in \mathbb{C}^{k_j} \setminus \mathbb{C}^{k_j-1},$$

Thus, if x was moved by S_{k_l}, all the other reflections S_{k_i} can not return it to the initial place. Let $x \in \mathrm{Fix}(S_{k_l})$. If it was moved by $S_{k_{l-1}}$, the all the remaining reflections can not return it to home. Etc. □

Proof of Theorem 4.7. Let g be a product of l reflections (4.3),

$$\mathrm{Fix}(g) \cap \widehat{\mathbb{C}}^\tau = \Big(\bigcap_{j \leqslant l} (\mathbb{C}v_j)^\perp\Big) \cap \widehat{\mathbb{C}}^\tau = \Big(\big(\bigoplus_{j \leqslant l} \mathbb{C}v_j\big) + \mathbb{C}^{n-\tau+1}\Big)^\perp$$

All the vectors v_j are linear independent, but v_j with small numbers (i.e., $k_j \leqslant n-\tau+1$) are in $\mathbb{C}^{n-\tau+1}$. This remark gives us the desired dimension. □

4.6. Digression. Schubert cells in the Grassmannian. Recall that a *flag* in a linear space V is a family of subspaces

$$0 \subset W_1 \subset W_2 \subset \dots \subset W_k \subset V$$

A *complete flag* is a flag containing subspaces of all dimensions.

We are going to describe the standard cell decomposition of the Grassmannian $\mathrm{Gr}_{p,q}$. Fix a complete a flag

$$0 \subset W_1 \subset W_2 \subset \dots \subset W_{p+q-1} \subset \mathbb{C}^{p+q} \tag{4.6}$$

in $\mathbb{C}^{p+q}$. For each collection of integers

$$0 \leqslant \alpha_1 \leqslant \alpha_2 \leqslant \ldots \leqslant \alpha_{p+q-1} \leqslant p$$

denote by $\mathcal{S}_{\alpha_1,\alpha_2,\ldots}$ (a *Schubert cell*) the set of all subspaces $P \in \mathrm{Gr}_{p,q}$ such that

$$\dim P \cap W_\tau = \alpha_\tau$$

(by definition, $\alpha_{j+1} - \alpha_j$ is 0 or 1). It is not difficult to show that each $\mathcal{S}_{\alpha_1,\alpha_2,\ldots}$ is homeomorphic to some linear space $\mathbb{C}^{N(\alpha)}$. Thus we obtain a CW-complex with zero boundary operators (since dimensions of all the cells are even). This implies

Theorem 4.9 *Manifolds $\mathcal{S}_{\alpha_1,\alpha_2,\ldots}$ form a basis of homolgies of the manifold* $\mathrm{Gr}_{p,q}$

Now let our flag (4.6) be the standard flag of coordinate subspaces, $W_j := \mathbb{C}^j$. The stabilizer $\mathcal{P}$ of this flag in $\mathrm{GL}(p+q,\mathbb{C})$ is the upper-triangular group $\mathcal{P}$ consisting of matrices

$$g = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \ldots \\ 0 & a_{22} & a_{23} & \ldots \\ 0 & 0 & a_{33} & \ldots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}$$

Problem 4.2 Schubert cells are precesily the orbits of $\mathcal{P}$ on $\mathrm{Gr}_{p,q}$.

4.7. The Dynkin cells again. Now let us translate this cell decomposition into a language of isotropic Grassmannians. Above (Theorem 4.2) we identified $\mathrm{U}(n)$ with the Grassmannian of maximal isotropic subspaces in $\mathbb{C}^{n,n}$. We wish to define analogs of Schubert cells in this Grassmannian.

Recall that we decompose $\mathbb{C}^{n,n} = \mathbb{C}^n \oplus \mathbb{C}^n$ into an orthogonal sum of negative and positive subspaces and treat maximal isotropic subspaces as graphs of operators from the negative subspace to the positive one.

Consider the (maximal isotropic) subspace $F \subset \mathbb{C}^{n,n}$ consisting of vectors $v \oplus v$. Then

$$v \in \mathrm{Fix}(g) \qquad \Longleftrightarrow \qquad v \oplus v \in \operatorname{graph} g \cap F$$

Denote by $F_\tau \subset F$ the subspace consisting of vectors $v \oplus v$ whose coordinates satisfy $v_{\tau+1} = \cdots = v_n = 0$. By construction, such subspaces form a flag

$$0 \subset F_1 \subset F_2 \subset \cdots \subset F_n = F \tag{4.7}$$

Now we can rephrase the definition of the cells $\mathcal{B}$ in the following form.

Proposition 4.10 *A cell $\mathcal{B}_{\alpha_1,\ldots,\alpha_{n-1}}$ is the set of $g \in \mathrm{U}(n)$ such that*

$$\dim \operatorname{graph}(g) \cap F_j = \alpha_j$$

We observe a precise analogy with Schubert cells.

Denote by $\mathcal{P}$ the subgroup [10] in $\mathrm{U}(n,n)$ preserving the flag (4.7). An operator preserving F_τ must preserve also $F_\tau^\perp$ and hence our operators preserve the long flag

$$0 \subset F_1 \subset F_2 \subset \cdots \subset F_n = F_n^\perp \subset \cdots \subset F_2^\perp \subset F_1^\perp \subset \mathbb{C}^{n,n}$$

[10] It is called a minimal parabolic subgroup in $\mathrm{U}(n,n)$; a *parabolic subgroup* is a stabilizer of an arbitrary isotropic flag (which can be shorter than (4.7)). Parabolic subgroups have fundamental importance in the representation theory, but do not too often appear in this book. For instance, the Stifel manifold $\mathrm{Sti}(p,q)$ is a quotient-space of $\mathrm{U}(p,q)$ by a parabolic subgroup.

Theorem 4.11 *The cells $\mathcal{B}_{\alpha_1,\dots,\alpha_{n-1}}$ are the orbits of P on the Grassmannian of maximal isotropic subspaces.*

Obviously, the cells are $\mathcal{P}$-invariant. We omit a proof of the transitivity.

2.5 Jordan angles and Hua double ratio

5.1. Jordan angles in Euclidean space. Consider the Euclidean space $\mathbb{R}^n$, and two k-dimensional subspaces L_1 and L_2.

Theorem 5.1 *a) There are an orthonormal basis $e_1,\dots, e_k$ and an orthonormal basis $f_1,\dots, f_k$ such that*

$$\langle e_i, f_j \rangle = \mu_j \delta_{ij}, \quad \text{where } 0 \leqslant \mu_j \leqslant 1$$

b) The numbers

$$\mu_1 \geqslant \mu_2 \geqslant \dots \geqslant \mu_k \geqslant 0$$

are uniquely determined by L_1, L_2

We say that the numbers

$$\varphi_j(L_1, L_2) = \arccos \mu_j(L_1, L_2)$$

are *angles* between L_1 and L_2.

Remark. There is lot of terms for this geometric structure, namely *Jordan angles*, *stationary angles*, *complex distance*, *compound distance.*Unfortunately, it is almost impossible to visualize this geometric structure, since a pair of subspaces in $\mathbb{R}^3$ has only one nontrivial angle. The first nontrivial case is a pair of two-dimensional planes in $\mathbb{R}^4$. $\square$

By $\mathrm{O}(n)$ we denote the group of all orthogonal transformations of the real Euclidean space $\mathbb{R}^n$.

Theorem 5.2 *a) For $g \in \mathrm{O}(n)$*

$$\mu_j(gL_1, gL_2) = \mu_j(L_1, L_2)$$

b) If L_1, L_2 and M_1, M_2 are k-dimensional subspaces, and

$$\mu_j(L_1, L_2) = \mu_j(M_1, M_2) \qquad \text{for all } j$$

then there is an orthogonal transformation g such that

$$gL_1 = M_1, \qquad gL_2 = M_2$$

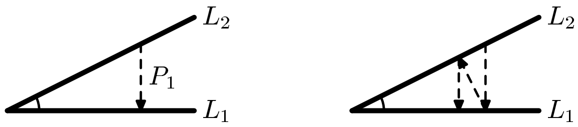

Fig. 2.11: To subsection 5.2. Definition of angles through projectors

PROOF OF THEOREM 5.1. Consider an orthonormal basis $u_\alpha \in L_1$ and an orthonormal basis $v_\beta \in L_2$. Consider the matrix $A_{ij} = \langle u_i, v_j \rangle$. We have a possibility to choose the both bases, therefore the matrix A is defined to within the equivalence

$$A \sim \xi A \eta, \qquad \xi, \eta \in \mathrm{O}(k)$$

Using these transformations, we can reduce the matrix A to the diagonal form

$$A = \begin{pmatrix} \mu_1 & 0 & \dots \\ 0 & \mu_2 & \dots \\ \vdots & \vdots & \ddots \end{pmatrix}, \qquad \mu_j \geqslant 0$$

The numbers μ_j are are the singular numbers of the matrix A, therefore they are uniquely determined.

Finally, the matrix A coincides with the matrix of the projection $L_1 \to L_2$. Hence $\|A\| \leqslant 1$ and $\mu_l \leqslant 1$. $\square$

PROOF OF THEOREM 5.2. Choose a system of vectors e_α, f_β for L_1, L_2. Choose an orthonormal basis r_σ in $(L_1 + L_2)^\perp$. Choose a similar system e'_α, f'_β, r'_σ for M_1, M_2. It remains to take the orthogonal operator g such that $ge_\alpha = e'_\alpha$, $gf_\beta = f'_\beta$, $gr_\sigma = r'_\sigma$. $\square$

REMARK. There is a minor difference between the cases $2k \leqslant n$ and $2k > n$. In the first case, for subspaces in general position $\mu_j \neq 0$. In the second case, $L_1 \cap L_2 \neq 0$, therefore at least $2k - n$ angles are zero by definition. $\square$

Corollary 5.3 *Let $2k \leqslant n$. Then there are k pairwise-orthogonal two-dimensional planes $W_1, \dots, W_k$, such that each W_j intersects both L_1, L_2 and is perpendicular to both of them.*

Problem 5.1 a) $\mu_j(L_1, L_2) = \mu_j(L_2, L_1)$.

b) Angles $\varphi_j(L_1^\perp, L_2^\perp)$ coincides with angles $\varphi_j(L_1, L_2)$ modulo some collection of zeros.

c) Describe possible stabilizers in the group $\mathrm{O}(n)$ of pairs L_1, L_2.

5.2. Reformulations of the definition. Denote by P_1 (resp. P_2) the orthogonal projection on L_1 (resp., L_2).

Proposition 5.4 *a) The numbers $\cos\varphi_j$ are the singular values of the operator*

$$P_1\Big|_{L_2} : L_2 \to L_1.$$

b) The eigenvalues of the operator $P_1P_2P_1$ are

$$\cos^2\varphi_1,\ \dots,\cos^2\varphi_k, 0, 0, \dots$$

PROOF is an exercise.

Problem 5.2 Show that the definition of angles survives for a pair of subspaces having different dimensions.

5.3. The minimax principle.

Theorem 5.5 *Let A be a self-adjoint operator in an n-dimensional Euclidean space W, denote by*

$$\lambda_1 \geqslant \lambda_2 \geqslant \dots \geqslant \lambda_n$$

its eigenvalues. Then

$$\lambda_l = \max_{\dim V = l}\left\{\min_{v\in V,\ \|v\|=1} \langle Av, v\rangle\right\}, \tag{5.1}$$

$$\lambda_l = \min_{\dim V = n-l+1}\left\{\max_{v\in V,\ \|v\|=1} \langle Av, v\rangle\right\} \tag{5.2}$$

where V ranges in the space of all the subspaces of W.

PROOF. Prove the first identity. Let e_1, e_2, $\dots$, be an eigenbasis for A. Consider the subspace H spanned by e_l, e_{l+1}, $\dots$ (its dimension is $(n-l+1)$). We have $\langle Av, v\rangle \leqslant \lambda_l\langle v, v\rangle$ for all $v\in H$.

Let V be an arbitrary l-dimensional subspace of W. The intersection $V\cap H$ is not zero (since the sum of their dimensions is $> n+1$). Let h be a unit vector in $V\cap H$. We have $\langle Ah, h\rangle \leqslant \lambda_l$, and hence

$$\min_{v\in V,\ \|v\|=1}\langle Av, v\rangle \leqslant \lambda_l$$

Therefore $\max\{\dots\}$ in (5.1) is $\leqslant \lambda_l$.

On another hand, $\max\{\dots\} \geqslant \lambda_l$ since we can choose V being the subspace spanned by the vectors $e_1, \dots e_l$.

Changing A to $(-A)$, we obtain (5.2). □

Problem 5.3 (Rayleigh Theorem) Consider an ellipsoid in $\mathbb{R}^3$ with semi-axes $a_1 \geqslant a_2 \geqslant a_3$. Consider its elliptic section by a plane through 0. Show that semiaxes of this ellipses satisfy the condition of interlacing

$$a_1 \geqslant b_1 \geqslant a_2 \geqslant b_2 \geqslant a_3$$

Problem 5.4 Let V, W be Euclidean spaces. Let $B : V \to W$ be an operator, denote by $\mu_1 \geqslant \mu_2 \geqslant \dots$ its singular values. Then

$$\mu_l = \max_{\dim S = l}\left\{\min_{v\in S,\ \|v\|=1} \|Av\|\right\} \tag{5.3}$$

Now we formulate the corresponding principle for angles (it is similar to common definitions of angles in $\mathbb{R}^3$.)

Proposition 5.6 *Let L_1, L_2 be k-dimensional subspaces in a Euclidean space.*

$$\varphi_k(L_1, L_2) = \max_{V \subset L_2,\, \dim V = k} \left\{ \min_{h \in V} \angle(h, L_1) \right\}$$

where $\angle(h, L_1)$ is the angle [11] *between a vector h and the subspace L_1.*

This follows from the previous problem and Proposition 5.4a).

REMARK. Below we use also a more complicated form of the same statement

$$\varphi_k = \max_{V \subset L_2,\, \dim V = k} \left\{ \min_{u_2 \in V} \left(\max_{u_1 \in L_1} \angle(u_1, u_2) \right) \right\} \tag{5.4}$$

where $\angle(u_1, u_2)$ is the angle between h and u. □

5.4. Angles in pseudo-Euclidean space.. Let $p \leqslant q$. Consider the space $\mathbb{C}^{p,q}$ and two points of L_1, $L_2 \in \mathcal{M}_{p,q}$ (i.e., maximal negative subspaces, see Section 2.3). We can regard such subspaces as Euclidean spaces with respect to the inner product $-\langle \cdot, \cdot \rangle$.

Theorem 5.7 *a) There are orthonormal bases $e_i \in L_1$, $f_j \in L_2$ such that*

$$-\langle e_k, e_l \rangle = \delta_{kl}, \ -\langle f_k, f_l \rangle = \delta_{kl}, \ -\langle e_i, f_j \rangle = \mu_i \delta_{ij} \tag{5.5}$$

The numbers

$$\mu_1 \geqslant \mu_2 \geqslant \ldots \geqslant 1$$

are uniquely determined by the subspaces L_1, L_2.

b) For negative subspaces L_1, L_2, M_1, M_2, the following conditions are equivalent

— $\mu_j(L_1, L_2) = \mu_j(M_1, M_2)$ *for all* j

— *there is* $g \in \mathrm{U}(p, q)$ *such that* $gL_1 = M_1$, $gL_2 = M_2$.

PROOF is the same as for Theorem 5.1, we only must show that the singular values μ_j are $\geqslant 1$.

For this, consider the subspace $H_\alpha := L_1 + \mathbb{C} f_\alpha$. It is an orthogonal direct sum of the $(k-1)$-dimensional subspace $\oplus_{j \neq \alpha} \mathbb{C} e_j$ and the plane (or line) $\mathbb{C} e_\alpha + \mathbb{C} f_\alpha$. If $0 \leqslant \mu_\alpha < 1$, then $\mathbb{C} e_\alpha + \mathbb{C} f_\alpha$ is a strictly negative two-dimensional plane. Hence H_α is a $(p+1)$-dimensional negative subspace in $\mathbb{C}^{p,q}$. This contradiction concludes the proof. □

Further, we define the numbers (*hyperbolic angles*) as

$$\psi_\alpha = \operatorname{arccosh}(\mu_j) = \ln\bigl(\mu_j + \sqrt{\mu_j^2 - 1}\bigr)$$

[11] i.e., the angle between h and its projection to L_1.

Problem 5.5 Convince yourself that all the considerations of previous subsections survive in our case.

5.5. Angles and coordinates. Now identify the set $\mathcal{M}_{p,q}$ of negative subspaces with the matrix ball $\mathrm{B}_{p,q}$ as above.

Theorem 5.8 *Let L, $M \in \mathcal{M}_{p,q}$, let z, u be their matrix coordinates.*

a) $\cosh \psi_j(L,M)$ *coincide with the singular values of the matrix*

$$\Lambda(z,u) = (1-zz^*)^{-1/2}(1-zu^*)(1-uu^*)^{-1/2} \tag{5.6}$$

b) The numbers $\cosh^2 \psi_j(L,M)$ *coincide with the eigenvalues of the matrix*

$$\Xi(z,u) = (1-zz^*)^{-1}(1-zu^*)(1-uu^*)^{-1}(1-uz^*) \tag{5.7}$$

PROOF. a) Denote $V_- := \mathbb{C}^p \oplus 0$ as above, (Subsection 3.1). Consider the map $R_z : V_- \to \mathbb{C}^{p,q}$ given by

$$v \mapsto v(1-zz^*)^{-1/2} \oplus v(1-zz^*)^{-1/2}z$$

We claim that R_z is an isometric embedding $V_{\to}\mathbb{C}^{p,q}$. Indeed,

$$\langle vR_z, vR_z\rangle = -v(1-zz^*)v^* + v(1-zz^*)zz^*(1-zz^*)^{-1}v^* = -vv^*$$

The image of this map is the subspace L.

Therefore for an orthonormal basis $e_j \in V_-$, the system of vectors $a_j := e_j R_z$ is an orthonormal basis in L.

In the same way, the system of vectors $b_j := e_j R_u$ is an orthonormal basis in M. We write out the matrix $\langle a_i, b_j\rangle$ as in proof of Theorem 5.1 and get

$$\begin{aligned}\langle a_i, b_j\rangle = -a_i(1-zz^*)^{-1/2}(1-uu^*)^{-1/2}b^* + a_i(1-zz^*)^{-1/2}zu^*(1-uu^*)^{-1/2}b^* = \\ = -a_i(1-zz^*)^{-1/2}(1-zu^*)(1-uu^*)^{-1/2}b_j^*\end{aligned}$$

This is the statement a).

b) Squares of singular values of $\Lambda(z,u)$ are the eigenvalues of

$$\begin{aligned}\Lambda(z,u)\Lambda(z,u)^* = (1-zz^*)^{-1/2}(1-zu^*)(1-uu^*)^{-1}(1-uz^*)(1-zz^*)^{-1/2} = \\ = (1-zz^*)^{1/2}\Xi(z,u)(1-zz^*)^{-1/2}\end{aligned}$$

We observe that the matrices $\Lambda\Lambda^*$ and Ξ are conjugate. Therefore they have the same eigenvalues. $\square$

Problem 5.6 By construction, for any $g \in \mathrm{U}(p,q)$, the matrices $\Xi(z,u)$ and $\Xi(z^{[g]}, u^{[g]})$ have the same eigenvalues. Show directly, that these two matrices are conjugate.

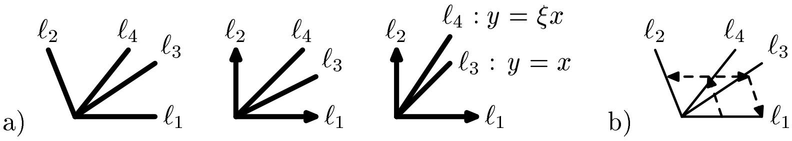

Fig. 2.12: To subsection 5.6.
a) Having 4 lines, we put ℓ_1 to the horizontal axis and ℓ_2 to the vertical axis. There is the one-parametric group of hyperbolic rotations preserving ℓ_1, ℓ_2. Using this group, we put ℓ_3 to the line $y = x$. Now the position of the line $\ell_4 : y = \xi x$ is fixed and $\xi = \xi(\ell_1, \ell_2, \ell_3, \ell_4)$ is the double ratio
b) Let us consider pair of lines ℓ_1, ℓ_2 as a coordinate system, then ℓ_3, ℓ_4 are graphs of certain operators $B, C : \ell_1 \to \ell_2$. We consider the operator $B^{-1}C : \ell_1 \to \ell_1$. It is an operator of multiplication by some canonically defined constant ξ.

5.6. The classical double ratio. Let t_1, t_2, t_3, t_4 be four distinct points on the projective line. Recall that their *double ratio* is the number

$$\xi(t_1, t_2, t_3, t_4) = \frac{t_1 - t_3}{t_1 - t_4} : \frac{t_2 - t_3}{t_2 - t_4}$$

see two its geometric definitions on Fig. 2.12.

Problem 5.7 There are 24 permutations of the points t_1, t_2, t_3, t_4.

a) The permutations $(12)(34)$, $(13)(24)$, $(14)(23)$ do not change double ratio.

b) There are 6 numbers that can be obtained from $\xi = \xi(t_1, t_2, t_3, t_4)$ by permutations of the arguments t_j, namely

$$\xi, \quad \xi^{-1}, \quad 1-\xi, \quad 1-\xi^{-1}, \quad (1-\xi)^{-1}, \quad -\xi(1-\xi)^{-1}$$

5.7. The Hua double ratio. Now, let W be a $(p+q)$-dimensional linear space. Let K_1, K_2 be subspaces of dimension q, let L_1, L_2 be subspaces of dimension p. Suppose this quadruple is in a general position. For us 'general position' means

$$K_2 \cap L_1 = 0, \qquad K_1 \cap L_2 = 0 \tag{5.8}$$

Then K_2 is a graph of an operator

$$P : K_1 \to L_1$$

and L_2 is graph of an operator

$$Q : L_1 \to K_1$$

The *Hua double ratio* $\Xi(K_1, L_1, K_2, L_2)$ of the quadruple is the linear operator

$$\Xi(K_1, L_1, K_2, L_2) := QP : \quad K_1 \to K_1$$

Theorem 5.9 *a) The eigenvalues $\lambda_1, \dots, \lambda_p$ of Ξ are invariants of a quadruple of subspaces under linear transformations of W*

b) Let $p \leqslant q$. Consider two quadruples (K_1, L_1, K_2, L_2) and (K_1', L_1', K_2', L_2'). Let the corresponding collections of eigenvalues $(\lambda_1, \dots, \lambda_p)$, $(\lambda_1', \dots, \lambda_p')$ coincide and λ_j are pairwise distinct nonzero numbers.[12] *Then one can find an invertible operator g in the ambient space W that maps one quadruple to another.*

REMARK. If $p > q$, then $\operatorname{rk} \Xi(\cdot) \leqslant q$, therefore $\Xi(\cdot)$ has at least $(q-p)$ zero eigenvalues. □

PROOF. a) is obvious, since the operator $\Xi(\cdot)$ is canonically defined.

b) Consider an eigenbasis $e_1, e_2, \dots, e_p$ of the operator $\Xi(K_1, L_1, K_2, L_2)$. Put $f_j := Pe_j$. Denote by Y_j the plane spanned by e_j and f_j. In this plane, we have 4 distinguished lines

$$\ell_1 = \mathbb{C}e_j = Y_j \cap K_1, \qquad \ell_2 = \mathbb{C}f_j = Y_j \cap L_1$$
$$\ell_3 = \mathbb{C}(e_j + f_j) = Y_j \cap K_2, \qquad \ell_4 = \mathbb{C}(f_j + \lambda e_j) = Y_j \cap L_2$$

Also,

$$W = Y_1 \oplus \dots \oplus Y_p \oplus (L_1 \cap L_2)$$

For another quadruple, one can find a similar decomposition with the same picture in each summand. □

Corollary 5.10 *Consider four n-dimensional subspaces V_j in a general position in a $2n$-dimensional linear space W. There are n two-dimensional planes $Y_1, \dots Y_n$, such that $W = \oplus Y_j$, and each Y_j intersects all the V_j*

EXAMPLE. Consider 4 two-dimensional planes in a general position in $\mathbb{C}^4$. Equivalently, consider 4 pairwise skew lines ℓ_1, ℓ_2, ℓ_3, ℓ_4 in the projective space $\mathbb{CP}^3$. For each point $z \in \mathbb{CP}^3$ one can find a unique line $m = m_z$ containing z and intersecting ℓ_1 and ℓ_2. All lines intersecting ℓ_1, ℓ_2, ℓ_3 sweep a quadric in $\mathbb{CP}^3$ (a hyperboloid in our real life). A generic line ℓ_4 intersects this quadric in two points. Thus, there are two lines m_1, m_2 intersecting four lines ℓ_1, ℓ_2, ℓ_3, ℓ_4. We have a configuration of 4 points in each line m_1, m_2. Thus we have the double ratio on each line m_1, m_2. This pair is the Hua double ratio. □

Problem 5.8 Interpret in the terms of the Hua ratio the following positions of 4 lines

— 4 skew lines in a hyperboloid

— line l_4 is tangent to the hyperboloid containing the lines ℓ_1, ℓ_2, ℓ_3.

5.8. Matrix expression for double ratio. Fix a decomposition of the ambient space W into a direct sum $W = X \oplus Y$ of a p-dimensional subspace X and a q-dimensional subspace Y. Our subspaces K_1, K_2, consist of vectors

$$K_1 : \ x \oplus A_1 x, \qquad K_2 : \ x \oplus A_2 x,$$

[12] this is a property that holds in a general position

where A_1, A_2 are linear operators $X \to Y$ and x ranges in X. Likewise, L_1, L_2 consist of vectors

$$L_1 : \ B_1 y \oplus y, \qquad L_2 : \ B_2 y \oplus y$$

where y ranges in Y.

Theorem 5.11 *The double ratio of this quadruple is*

$$\Xi(K_1, L_1, K_2, L_2) = (1 - B_2 A_1)^{-1}(B_2 - B_1)(1 - A_1 B_2)^{-1}(A_2 - A_1); \tag{5.9}$$

in this formula, we identify X and K_1, and the operator Ξ is written as an operator $X \to X$

PROOF. We must write out the operators P and Q from the previous subsection. Let us decompose a vector $h \oplus A_2 h \in K_2$ as sum of vectors lying in K_1, L_1,

$$h \oplus A_2 h = \big[x \oplus A_1 x\big] + \big[B_1 y \oplus y\big]$$

i.e.,

$$h = x + B_1 y, \qquad A_2 h = A_1 x + y$$

Excluding h, we obtain a formula for the operator P,

$$y = (1 - A_2 B_1)^{-1}(A_2 - A_1)x$$

Decomposing a vector $B_2\eta \oplus \eta \in L_2$ as a sum of vectors lying in K_1, L_1, we obtain a formula for the operator Q:

$$x = (1 - B_2 A_1)^{-1}(B_2 - B_1)y$$

The last two formulae provide the desired expression. □

Now, let the dimension of the ambient space W be $2n$, and let all subspaces of our quadruple K_1, K_2, K_3, K_4 have the same dimension n. Decompose the space into a direct sum of n-dimensional subspaces, $W = X \oplus Y$. Now we can write out all the subspaces K_j in a uniform way. Thus, let K_j consist of vectors $x \oplus S_j x$, where $j = 1, 2, 3, 4$. Then

$$\Xi(K_1, K_2, K_3, K_4) = (S_4 - S_1)^{-1}(S_4 - S_2)(S_3 - S_2)^{-1}(S_3 - S_1); \tag{5.10}$$

we merely substitute $A_1 \mapsto S_1$, $B_1 \mapsto S_2^{-1}$, $A_2 \mapsto S_2$, $B_2 \mapsto S_4^{-1}$ to (5.9.

Proposition 5.12 *Denote by ξ_j the eigenvalues of the matrix (5.10). Under permutations of S_1, S_2, S_3, S_4 we can obtain one of the six following collections of numbers*

$$\{\xi_j\}, \quad \{\xi_j^{-1}\}, \quad \{1 - \xi_j\}, \quad \{1 - \xi_j^{-1}\}, \quad \{(1 - \xi_j)^{-1}\}, \quad \{-\xi_j(1 - \xi_j)^{-1}\},$$

PROOF. A straightforward verification of the behavior of (5.10) under permutations of arguments is a pleasant exercise. But the statement is trivial, it follows from Corollary 5.10 and Problem 5.7,

5.9. The Jordan angles and the Hua ratio. Let us return to the situation discussed at the beginning of this Section. We have two subspaces L_1, L_2 in an Euclidean space having angles φ_j. Also, we have a quadruple L_1, L_2, $L_1^\perp$, $L_2^\perp$.

Theorem 5.13 *Hua double ratio of this quadruple coincides with the numbers* $\cos^2 \varphi_j$

We leave this as an exercise.

5.10. Digression. An example of vizualization of an angle-type structure.

Problem* 5.9 Consider the 3-dimensional sphere $S^3 := \mathbb{R}^3 \cup \infty$ with the action of the conformal group $\mathrm{O}(4,1)$. Find invariants of pairs of circles C_1, C_2 under this action. Use the following idea. Points of S^3 are in one-to-one correspondence with isotropic lines in the pseudo-Euclidean space $\mathbb{R}^{3,1}$. Circles are in one-to-one correspondence with planes having the inertia indices $(1,1)$. Such pair of planes together with their orthocomplements have Hua double ratios. Interpret this double ratio in the terms of geometry of our $\mathbb{R}^3$.

2.6 Classification of pseudo-unitary and pseudo-self-adjoint operators

6.1. Formulations of problems. We intend to solve the following 3 problems.

1°. *Find canonical forms of self-adjoint operators in pseudo-Euclidean spaces.* More precisely, the group $\mathrm{U}(p,q)$ acts on the set of self-adjoint operators (we also use the term *pseudo-Hermitian* operator) by conjugations $g: A \mapsto g^{-1}Ag$. We wish to classify orbits and to find canonical forms

2°. *Find canonical forms of pseudo-unitary operators.*

In other words, we intend to classify conjugacy classes in the group $\mathrm{U}(p,q)$.

3°. *Find canonical forms of pairs of Hermitian forms on a given linear space* V *under the action of the complete linear group* $\mathrm{GL}[V]$.

6.2. Self-adjoint operators in general position. Let A be a self-adjoint operator in V having distinct eigenvalues.

Theorem 6.1 *a) The spectrum of* A *is symmetric about the real axis, i.e., it has the form*

$$\lambda_1, \lambda_2, \dots, \mu_1, \overline{\mu}_1, \mu_2, \overline{\mu}_2, \dots \quad \textit{where } \lambda_j \in \mathbb{R},\ \mu_k \in \mathbb{C} \setminus \mathbb{R}.$$

b) There is a basis e_1, $e_2, \dots$, f_1, f_1', f_2, $f_2', \dots$ *in* V *such that*

$$\langle e_j, e_j \rangle = \pm 1, \qquad \langle f_k, f_k' \rangle = 1 \tag{6.1}$$

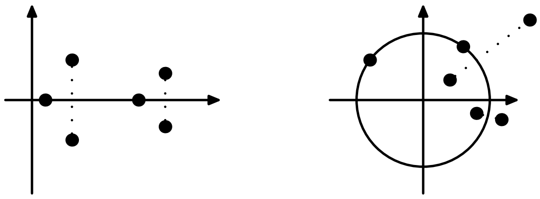

Fig. 2.13: The spectrum of a pseudo-Hermitian operator is symmetric about the real axis. The spectrum of a pseudo-unitary operator is symmetric with respect to the circle $|\lambda| = 1$, i.e., it is stable under the inversion $\lambda \mapsto 1/\overline{\lambda}$.

(all other pairwise scalar products equal zero) and

$$Ae_j = \lambda_j e_j,$$
$$Af_k = \mu_k f_k, \qquad Af'_k = \overline{\mu}_k f'_k.$$

PROOF. Denote by v_ν an eigenvector corresponding to an eigenvalue ν. We have

$$\langle Av_\nu, v_\mu\rangle = \langle v_\nu, Av_\mu\rangle \qquad \Longrightarrow \qquad \nu\langle v_\nu, v_\mu\rangle = \overline{\mu}\langle v_\nu, v_\mu\rangle$$

Thus we have an alternative:

$$\langle v_\nu, v_\mu\rangle = 0 \quad \text{or} \quad \nu = \overline{\mu}$$

In particular, if $\nu \notin \mathbb{R}$, then $\langle v_\nu, v_\nu\rangle = 0$.

But our form is nondegenerate, Therefore for each v_ν one can find v_μ such that $\langle v_\nu, v_\mu\rangle \neq 0$. The only possibility is $\mu = \overline{\nu}$.

For real ν this implies $\langle v_\nu, v_\nu\rangle \neq 0$.

If ν is not real, then $\mu = \overline{\nu}$ must be an eigenvalue. □

Problem 6.1 Let $g \in \mathrm{U}(p,q)$ be a pseudo-unitary operator having a simple spectrum.

a) Show that the spectrum is symmetric with respect to the circle $|\lambda| = 1$, i.e., for $\lambda \in \mathrm{Spec}(g)$, also $\overline{\lambda}^{-1} \in \mathrm{Spec}(g)$.

b) Show that there exists a basis $e_1, e_2, \ldots, f_1, f'_1, f_2, f'_2, \ldots$ in V as above (6.1) such that g is

$$Ae_j = \lambda_j e_j, \qquad |\lambda_j| = 1$$
$$Af_k = \gamma_k, \qquad Af'_k = \overline{\gamma}_k^{-1} f'_k, \qquad |\gamma_k| \neq 1.$$

6.3. Reformulation of the problem. Often (and even usually) Theorem 6.1 is sufficient for understanding situation. However, we start an examination of of arbitrary self-adjoint operators in V.

We say that a self-adjoint operator A in V is *decomposable*, if there are A-invariant subspaces W, Y such that $W \oplus Y = V$ and $W = V^\perp$. In particular, V and W must be regular.

Problem 6.2 An operator A is decomposable iff there exists a proper A-invariant regular subspace W.

Thus the problem of description of canonical forms can be formulated as follows.

To describe all the indecomposable self-adjoint operators in all the pseudo-Euclidean spaces.

6.4. The classification theorem.

Theorem 6.2 *There are two types of indecomposable self-adjoint operators:* $\mathrm{I}_n^{\pm}(\lambda)$, *where* $\lambda \in \mathbb{R}$, $n \in \mathbb{N}$, *and* $\mathrm{II}_n(\mu)$, *where* $\mu \in \mathbb{C} \setminus \mathbb{R}$*:*

— *Type* $\mathrm{I}_n^{\pm}(\lambda)$. *There is a basis* $e_1, \ldots, e_n$ *with scalar products* [13]

$$\langle e_k, e_l \rangle = \pm \delta_{k+l,n+1}.$$

The operator A *is*

$$\begin{aligned} Ae_j &= \lambda e_j + e_{j+1}, \qquad j \neq n, \\ Ae_n &= \lambda e_n. \end{aligned}$$

— *Type* $\mathrm{II}_n(\mu)$. *There is a basis* $f_1, \ldots, f_k$, $f_1', \ldots, f_k'$ *with scalar products*

$$\langle f_i, f_j \rangle = \langle f_i', f_j' \rangle = 0; \qquad \langle f_i, f_j' \rangle = \delta_{i+j,n+1}$$

The operator A *is given by*

$$\begin{aligned} Af_j &= \mu f_j + f_{j+1}; \qquad (j \neq n), \qquad & Af_k' &= \overline{\mu} f_k' + f_{k-1}', \qquad (k \neq 1), \\ Af_n &= \mu e_n, & Af_1' &= \mu f_1'. \end{aligned}$$

Since the description looks as intricate, we wish to clarify it.

1) An indecomposable operator with a real eigenvalue is a Jordan box. Moreover each Jordan box A in a linear space admits a unique to within a scalar factor pseudo-Euclidean scalar product making A self-adjoint.

2) Let us present a verbal description of an operator of type II. Consider the subspaces W, W' spanned respectively by $f_1, \ldots, f_n$ and $f_1', \ldots, f_n'$. They are isotropic and dual one to another (see Subsection 1.10). Further W, W' are invariant, and the restrictions of A to W amd W' are Jordan boxes. These Jordan boxes are dual operators in dual spaces. Theorem claims that nothing else can happened for non-real eigenvalues.

3) Now let us discuss $\pm$. This sign arises as an invariant even for operators with simple spectrum (see Theorem 6.1) It is easy to distinguish $\mathrm{I}_n^+(\lambda)$ and $\mathrm{I}_n^-(\lambda)$ if $n = 2k+1$ is odd, since the ambient pseudo-Euvlidean spaces have different inertia indices(namely, $(k+1, k)$ and $(k, k+1)$). Generally, consider a cyclic vector[14] h of the Jordan box $A - \lambda$. Then $\operatorname{sgn}\langle (A-\lambda)^{n-1} h, h \rangle$ is an invariant of a Jordan box A.

6.5. The symmetry of the spectrum.

[13] δ_{kl} denotes the *Kronecker symbol*, it equals 1 if $k = l$ and zero otherwise.

[14] A vector v is *cyclic* for an operator A, if the vectors v, Av, $A^2v, \ldots$ span the whole space.

Lemma 6.3 *The spectrum of a self-adjoint operator is symmetric with respect to the real axis, i.e., if $\lambda \in \operatorname{Spec} A$, then $\overline{\lambda}$ also is an element of the spectrum having the same multiplicity.*

PROOF. The condition $A^{\circledast} = A$ of self-adjointness is $JA = A^*J$, where A^* denotes the usual adjoint matrix and $J = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$. Let λ be an eigenvalue of A. Then

$$0 = \det(A - \lambda) = \det(J^{-1}) \det(A - \lambda) \det(J) = \det(J^{-1}AJ - \lambda) = \det(A^* - \lambda)$$

Therefore the spectrum of A coincides with the spectrum of A^*. $\square$

6.6. The root decomposition. Consider a self-adjoint operator A in a space V. Denote by V_ν the root subspace corresponding to an eigenvalue ν,

$$V_\nu = \ker(A - \nu)^N \qquad \text{for sufficiently large } N$$

Lemma 6.4 *The space V is an orthogonal direct sum of the regular subspaces of the forms*

— *V_λ with $\lambda \in \mathbb{R}$*

— *$V_\mu + V_{\overline{\mu}}$ with $\mu \in \mathbb{C} \setminus \mathbb{R}$; the summands V_μ, $V_{\overline{\mu}}$ are isotropic and have the same dimension.*

PROOF. Consider $v \in V_\nu$, $w \in V_\mu$. Then

$$0 = \langle (A - \nu)^N v, w \rangle = \langle v, (A - \overline{\nu})^N w \rangle$$

The operator $(A - \overline{\nu})^N$ is invertible in all the root spaces V_σ except $V_{\overline{\nu}}$. Hence

$$V_\nu \perp V_\sigma \quad \text{for } \sigma \neq \overline{\nu}$$

Therefore V is an orthogonal direct sum of subspaces of the forms V_λ and $V_\mu \oplus V_\mu$; recall that $\lambda \in \mathbb{R}$ and $\mu \notin \mathbb{C}$. Since the form $\langle \cdot, \cdot \rangle$ is nondegenerate, it follows that all the summands must be regular.

The spaces V_μ and $V_{\overline{\mu}}$ are isotropic. Since their sum is regular, they have the same dimensions (since dimension of an isotropic subspace is $\leqslant$ one-half of a dimension of ambient space).[15] $\square$.

REMARK. Let y_k be a basis in V_μ and y_k° be a dual basis in $V_{\overline{\mu}}$. Then the matrix of our operator in $V_\mu \oplus V_{\overline{\mu}}$ has a form $\begin{pmatrix} H & 0 \\ 0 & H^* \end{pmatrix}$. $\square$

Now Theorem 6.2 about canonical forms is reduced to two following cases:

a) an operator having a unique real eigenvalue λ,

b) an operator having two complex conjugate eigenvalues.

These cases are examined in Subsections 6.7–6.9

[15] This gives also another proof of Lemma 6.3.

Problem 6.3 Prove a counterpart of Lemma 6.4 for pseudo-unitary operators.

6.7. Indecomposable operators. Complex eigenvalues. Let $\mu \in \mathbb{C} \setminus \mathbb{R}$. Consider a self-adjoint operator whose eigenvalues are $\{\mu, \overline{\mu}\}$. Recall that the subspaces V_μ and $V_{\overline{\mu}}$ are isotropic and dual with respect to the scalar product.

Assume that V_μ is not a Jordan box, i.e., V_μ decomposes into a sum $Y \oplus Z$ of two A-invariant subspaces. Choose a basis $y_j \in Y$ and a basis $z_k \in Z$. Denote by y_j°, z_k° the dual basis in V_μ. Then the vectors of the form y_j, y_i° span a regular A-invariant subspace.

Thus, an indecomposable operator A has one Jordan box in V_μ. Choosing a Jordan basis in V_μ and the dual basis in $V_{\overline{\mu}}$, we obtain the desired canonical form.

6.8. Indecomposable operators. Real eigenvalues.

Lemma 6.5 *An indecomposable self-adjoint operator having a real eigenvalue is a Jordan box.*

Let A be a self-adjoint operator in a space V having a unique eigenvalue $\lambda \in \mathbb{R}$. To simplify a notation, we assume $\lambda = 0$, otherwise, we pass to the operator $\widetilde{A} := A - \lambda$.

First, we formulate two simple problems.

Problem 6.4 Let A be a nilpotent[16] Jordan box in a space W, h a cyclic vector. Then all the A-invariant subspaces have the form $W_k = \mathbb{C} \cdot A^k h + \mathbb{C} \cdot A^{k+1} h + \dots$, or equivalently $W_k = A^k W$.

Problem 6.5 Let Z be a space equipped with a degenerate Hermitian form S, and B self-adjoint with respect to S linear operator. Show that the kernel of S is B-invariant.

Proof of Lemma 6.5. Denote by N the maximal exponent such that $A^N \neq 0$.

Problem 6.6 $(\ker A^N)^\perp = \operatorname{im} A^N$

Let us define the form

$$\mathcal{B}(v, w) := \langle A^N v, w \rangle, \qquad v,\, w \in V \tag{6.2}$$

Since A is self-adjoint, the form $\mathcal{B}$ is Hermitian. Obviously, $\ker \mathcal{B} = \ker A^N$. Thus, we can regard $\mathcal{B}$ as a nondegenerate Hermitian form on $V / \ker A^N$.

By the previous remark, there is a vector $h \in V$ such that

$$\mathcal{B}(h, h) \neq 0 \tag{6.3}$$

Lemma 6.6 *The subspace W spanned by the vectors h, $Ah, \dots$, $A^N h$ is regular.*

[16] An operator A is nilpotent, if $A^N = 0$ for some N. Equivalently, all the eigenvalues of A are zero.

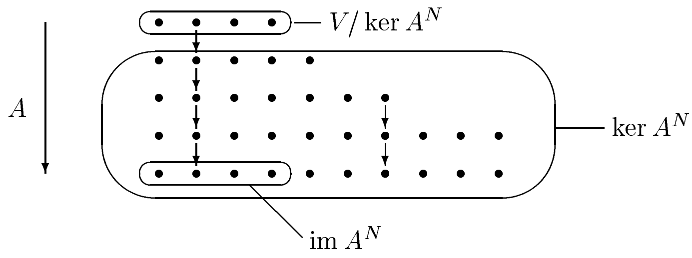

Fig. 2.14: Black circles correspond to elements of a Jordan basis. The operator A is the operator of the shift down. A vertical chain spans an invariant subspace.

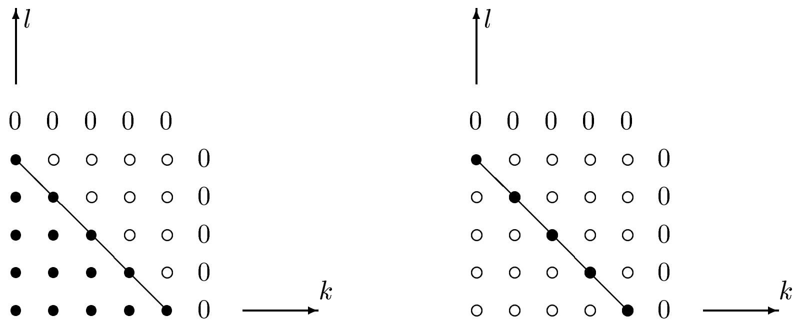

Fig. 2.15: The left-hand side. The coefficients c_{kl} are constant along the lines $k+l = \text{const}$. For positions marked by white circles, $c_{kl} = 0$. All the coefficients on the black positions except of the marked diagonal can be 'killed'. We obtain a picture on the right-hand side.

Proof. By construction, W is a Jordan box. Suppose W is a singular. By Problem 6.5, the kernel $zero(W)$ is an A-invariant subspace. By Problem 6.4 $zero(W)$ contains $A^N h$. But this contradicts to the choice of the vector h, see (6.3). This completes the proof of the latter lemma. □

Thus we constructed a regular invariant subspace for A. This finishes the proof of Lemma 6.5. □

6.9. The end of proof of Theorem 6.2. Nilpotent Jordan boxes. Due to Lemma 6.5, it remains to examine scalar products compatible with a nilpotent Jordan box.

Thus, let A be a self-adjoint Jordan box such that $A^N \neq 0$, $A^{N+1} = 0$, let e_0 be a cyclic vector. Denote

$$e_k := A^k e_0, .$$

Also, denote

$$c_{kl} := \langle e_k, e_l \rangle,$$

see Fig. 2.15. In particular, if

$$\big\{k > N \quad \text{or} \quad l > N\big\} \Rightarrow c_{kl} = 0 \tag{6.4}$$

(since $e_{N+1} = 0$). By the self-adjointness,

$$c_{k(l+1)} = \langle e_k, e_{l+1}\rangle = \langle e_k, Ae_l\rangle = \langle Ae_k, e_l\rangle = c_{(k+1)l}. \tag{6.5}$$

We denote

$$\sigma_{k+l} := c_{kl}.$$

By (6.4), it follows that $\sigma_p = 0$ for $p > N$. Also, $c_{kl} = \overline{c}_{lk}$. Hence $\sigma_{k+l} \in \mathbb{R}$.

The cyclic vector e_0 is not unique, we can replace it by an arbitrary vector of the form

$$\widetilde{e}_0 = s_0 e_0 + s_1 e_1 + s_2 e_2 + \ldots, \qquad s_0 \neq 0.$$

Choosing $\widetilde{e}_0 = s_0 e_0$, we can get $\sigma_N = \pm 1$.

Next, choose $\widetilde{e}_0 = e_0 + s_1 e_1$. Then

$$\widetilde{\sigma}_{N-1} = \langle \widetilde{e}_0, \widetilde{e}_{N-1}\rangle = \langle e_0 + s_1 e_1, e_{N-1} + s_1 e_N\rangle = \sigma_{N-1} + (s_1 + \overline{s}_1)\sigma_N$$

Choosing s_1 in an appropriate way, we can assume $\widetilde{\sigma}_{N-1} = 0$.

Thus, assume $\sigma_{N-1} = 0$, and vary e_0 as $\widetilde{e}_0 = e_0 + s_2 e_2$. Then

$$\widetilde{\sigma}_{N-2} = \langle e_0 + s_2 e_2, e_{N-1} + s_2 e_N\rangle = \sigma_{N-1} + (s_2 + \overline{s}_2)\sigma_N$$

Choosing s_2 in an appropritae way, we 'kill' σ_{N-2}. Etc.

This completes the proof of Theorem 6.2.

6.10. Conjugacy classes in the group $\mathrm{U}(p,q)$. We say that $g \in \mathrm{U}[V]$ is *decomposable*, if V decomposes into a direct sum of two orthogonal invariant subspaces. Equivalently, an operator is decomposable if there is a g-invariant regular subspace $W \subset V$.

The classification of conjugacy classes is reduced to description of all the indecomposable pseudo-unitary operators.

Theorem 6.7 *There are two types of indecomposable pseudo-unitary operators.*

— *An operator* $\exp(iA)$*, where* A *is of the form* $\mathrm{I}_n^{\pm}(\lambda)$*, where* $\lambda \in [0, 2\pi)$

— *An operator* $\exp(iB)$*, where* B *has the form* $\mathrm{II}_n(\mu)$*.*

Proof. 1. The spectrum is symmetric with respect to the circle $|\lambda| = 1$, i.e., $\lambda \in \mathrm{Spec}(g)$ implies $\overline{\lambda}^{-1} \in g$ and these eigenvalues have the same multiplicity.

2. Denote by V_λ the root space corresponding an eigenvalue λ. The space V is an orthogonal direct sum of subspaces having the following forms:

— V_λ, where $|\lambda| = 1$,

— $V_\gamma \oplus V_{\overline{\gamma}^{-1}}$, where $|\gamma| \neq 1$.

3. Let $|\gamma| \neq 1$. The considerations of Subsection 6.7 can be applied in our case. We obtain a Jordan block J in a space V_γ and the operator J^{*-1} in the dual subspace $V_{\overline{\gamma}^{-1}}$ (this canonical form differs from the form described in our Theorem 6.7).

4. Consider the case $|\lambda| = 1$. We can think that $V = V_\lambda$. The operator $g' = \lambda^{-1} g$ also is pseudo-unitary. Its unique eigenvalue is 1. We define its logarithm as

$$\ln g' := -\sum_{j=0} \frac{1}{j}(1-g')^j$$

Obviously, this sum is finite.

Problem 6.7 Show that $\frac{1}{i}\ln g'$ is a self-adjoint operator (this is not totally trivial).

This problem finishes the proof. □

6.11. A pair of Hermitian forms. Now let B_1, B_2 be a pair of Hermitian forms on a complex linear space V. We wish to find canonical forms of such pairs with respect to the action of the complete linear group $\mathrm{GL}[V]$.

Assume that one of the forms, say B_1 is nondegenerate.

Proposition 6.8 *There is a unique operator T such that*

$$B_2(v, w) = B_1(v, Tw).$$

Moreover, T is B_1-pseudo-Hermitian.

Thus, our problem is equivalent to the description of canonical forms of pseudo-Hermitian operators. The latter problem is solved above.

6.12. Invariants of two Hermitian forms. Now let Q_1 and Q_2 be matrices of the forms B_1, B_2. Changing a basis, we transform these two matrices as

$$Q_1 \mapsto g^* Q_1 g, \qquad Q_2 \mapsto g^* Q_2 g \tag{6.6}$$

Theorem 6.9 *The roots of the equation*

$$\det(Q_1 - \lambda Q_2) = 0$$

are invariants under the transformation (6.6).

PROOF. $\det(g^* Q_1 g - \lambda g^* Q_2 g) = \det(g^*)\det(Q_1 - \lambda Q_2)\det(g)$. □

6.13. Some similar problems.

Problem 6.8 Convince yourself that you can solve (some of) the following problems

a) Classification of conjugacy classes in the complex orthogonal group $O(n, \mathbb{C})$

b) Canonical forms of a pair of symmetric bilinear forms on a complex linear space (if one of these forms is non-degenerate)

c) Canonical forms of a pair $\{$ symmetric bilinear form, skew-symmetric bilinear form $\}$ on a complex linear space (if one of these forms is non-degenerate).

2.7 Indefinite contractions

7.0. Contractive operators in Euclidean spaces. First, we wish to recall several definitions and trivial statements concerning operators in in Euclidean spaces.

An invertible operator g in a complex Euclidean space is a *contraction* if

$$\langle vg, vg\rangle \leqslant \langle v, v\rangle \qquad \text{for all } v \in V$$

or equivalently $\|g\| \leqslant 1$.

An operator X in a Euclidean space is *dissipative*, if $\mathrm{Re}\langle Xv, v\rangle \leqslant 0$.

Proposition 7.1 *An operator X is dissipative iff it can be represented as $X = A + B$, where $A = -A^*$ is anti-Hermitian, and $B = B^*$ is Hermitian and negative semi-definite.*

Theorem 7.2 *The following conditions for an operator X are equivalent*
– X is dissipative
– the semigroup $\exp(tX)$ consists of contractions.

Theorem 7.3 *Each contractive invertible operator g admits a unique decomposition $g = U\cdot\exp(-X)$, where U is unitary and X is self-adjoint semi-positive.*

Problem 7.1 Verify these statements.

For instance, to prove Theorem 7.3, we consider the polar decomposition $g = US$ and put $X := -\ln S$. Since S is positive definite, it follows that its logarithm is well defined. Since $\|S\| \leqslant 1$, the eigenvalues of S satisfy $0 < \lambda_j \leqslant 1$. Therefore $\ln\lambda_j \leqslant 0$ and hence X is negative semi-definite.

REMARK. Let g be an arbitrary operator. Then $s \cdot g$ is contractive for some $s \neq 0$. This is not the case for pseudo-Euclidean contractions defined below.

7.1. Indefinite contractions. Let $V \simeq \mathbb{C}^{p,q}$ be a pseudo-Euclidean space. An invertible operator $g : V \to V$ is an *indefinite contaction* if

$$\langle vg, vg\rangle \leqslant \langle v, v\rangle \qquad \text{for all } v \in V$$

We denote the set of all the indefinite contractions in V by $\Gamma\mathrm{U}[V]$ or $\Gamma\mathrm{U}(p,q)$. Obviously, $\Gamma\mathrm{U}(p,q)$ a semigroup, i.e., for each $g_1, g_2 \in \Gamma\mathrm{U}(p,q)$, we have $g_1g_2 \in \Gamma\mathrm{U}(p,q)$.

We say that g is a *strict contraction* if

$$\langle vg, vg\rangle < \langle v, v\rangle \qquad \text{for all nonzero } v \in V$$

Let a scalar product in V be $\langle v, w\rangle := vQw^*$.

Proposition 7.4 $g \in \Gamma\mathrm{U}(p,q)$ *iff* $Q - gQg^* \geqslant 0$.

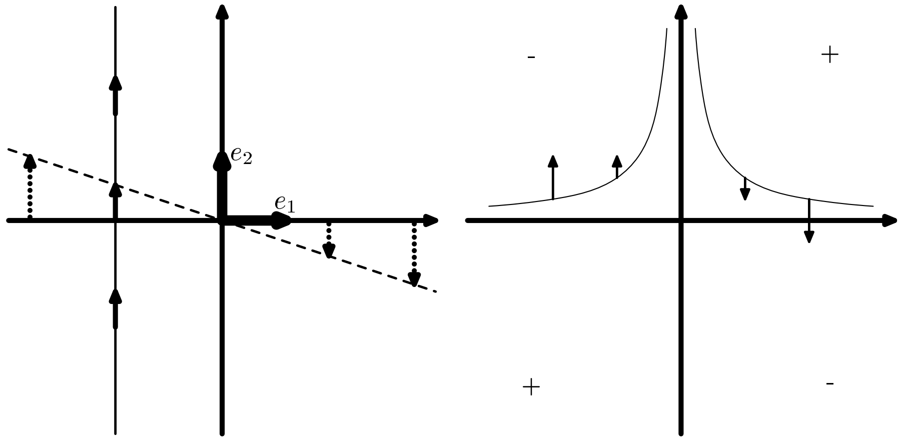

Fig. 2.16: To Example 3. We draw $\mathbb{R}^{1,1}$ instead of $\mathbb{C}^{1,1}$. The axes are isotropic and level lines $\langle v, v\rangle = \text{const}$ are hyperbolas. The signs $\pm$ indicate quadrants with positive/negative scalar squares. The operator A_t fixes the vertical axis and shifts lines $x = c$ by $(-ct)$ (see the left figure). The contractivity of A_t is observable from the right figure.

PROOF is obvious. □

7.2. Several examples.

EXAMPLE 1. A pseudo-unitary operator is a contraction. □

EXAMPLE 2. Let $A : \mathbb{C}^p \to \mathbb{C}^p$, $B : \mathbb{C}^q \to \mathbb{C}^q$ be invertible linear operators, $\|A\| \leqslant 1$, $\|B\| \leqslant 1$. Then $\begin{pmatrix} A^{-1} & 0 \\ 0 & B \end{pmatrix}$ is an indefinite contraction. □

EXAMPLE 3. Consider a two-dimensional space with basis e_1, e_2 such that

$$\langle e_1, e_1\rangle = \langle e_2, e_2\rangle = 0, \qquad \langle e_1, e_2\rangle = 1 \tag{7.1}$$

and the operator (a Jordan box)

$$A_t e_1 = e_1 - te_2, \qquad A_t e_2 = e_2, \qquad \text{where } t > 0 \tag{7.2}$$

(this operator is $\mathrm{I}_2^+(1)$ in the classification of Subsection 6.4). For $v = \alpha e_1 + \beta e_2$,

$$\langle v, v\rangle - \langle A_t v, A_t v\rangle = \langle \alpha e_1 + \beta e_2, \alpha e_1 + \beta e_2\rangle - \langle \alpha e_1 + (\beta - t\alpha) e_2, \alpha e_1 + (\beta - t\alpha) e_2\rangle = \\ = 2t|\alpha|^2 \geqslant 0$$

Thus A_t is a contraction, but not a strict contraction. □

REMARK. Theorem 7.7 obtained below, shows that operators of these 3 types generate the semigroup $\Gamma\mathrm{U}(p, q)$. □

REMARK. Denote by $C_- \subset V$ the cone consisting of nonpositive vectors v. A contractive operator g transfers C_- to C_-. For an inverse statement, see Theorem 7.13. □

The next example illustrates our argument that is used below for proofs of non-contractivity.

EXAMPLE 4. In the notation of Theorem 6.2, consider the indecomposable pseudo-self-adjoint operator of the type $I_4^+(\lambda)$. In other words, we have a basis $e_1, \ldots, e_4$ with scalar products $\langle e_k, e_l \rangle = \delta_{k+l,5}$ and the matrix

$$A := \begin{pmatrix} 1 & 0 & 0 & 0 \\ \lambda & 1 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & \lambda & 1 \end{pmatrix}$$

Let $v = \sum \alpha_j e_j$. Then

$$\begin{aligned} &\langle v, v \rangle - \langle Av, Av \rangle = \\ &= 2\,\mathrm{Re}(\alpha_1 \overline{\alpha}_4 + \alpha_2 \overline{\alpha}_3) - 2\,\mathrm{Re}(\lambda\alpha_1(\overline{\alpha}_4 + \lambda\overline{\alpha}_3) + (\lambda\alpha_2 + \alpha_1)(\overline{\alpha}_3 + \lambda\overline{\alpha}_2)) = \\ &= \begin{pmatrix} \alpha_1 & \alpha_2 & \alpha_3 & \alpha_4 \end{pmatrix} \left\{ \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} - \begin{pmatrix} 0 & 1 & 2\lambda & \lambda^2 \\ 1 & 2\lambda & \lambda^2 & 0 \\ 2\lambda & \lambda^2 & 0 & 0 \\ \lambda^2 & 0 & 0 & 0 \end{pmatrix} \right\} \begin{pmatrix} \overline{\alpha}_1 \\ \overline{\alpha}_2 \\ \overline{\alpha}_3 \\ \overline{\alpha}_4 \end{pmatrix} \end{aligned}$$

The matrix in the curly brackets is not semi-positive. Indeed, for a Hermitian semi-positive matrix R, the following inequality for the minors holds

$$\det \begin{pmatrix} r_{ii} & r_{ij} \\ r_{ji} & r_{jj} \end{pmatrix} = r_{ii} r_{jj} - |r_{ij}|^2 \geqslant 0$$

If $r_{jj} = 0$, then $r_{ij} = 0$. Thus, *our matrix has too much zeros on the diagonal* therefore it is not semi-positive. Thus A is not contractive. □

7.3. Dissipative operators. We say that an operator X in a pseudo-Euclidean space is *(pseudo)-dissipative*, if

$$\mathrm{Re}\langle vX, v \rangle \leqslant 0 \qquad \text{for all } v$$

or equivalently

$$\langle v(X + X^{\circledast}), v \rangle \leqslant 0$$

Evidently, *the set of all the dissipative operators is a convex cone (wedge).* We denote it by $\mathrm{Diss} = \mathrm{Diss}(p,q)$.

Represent X in the form

$$X = A + B, \qquad \text{where } A = \tfrac{1}{2}(X - X^{\circledast}),\ B = \tfrac{1}{2}(X + X^{\circledast})$$

Then $A = -A^{\circledast} \in \mathfrak{u}(p,q)$ and $B^{\circledast} = B$ is self-adjoint dissipative, i.e., $\langle vB, v \rangle \leqslant 0$. We denote the cone of pseudo-self-adjoint dissipative operators by $\mathrm{SDiss}(p,q)$,

$$\mathrm{Diss}(p,q) = \mathfrak{u}(p,q) \oplus \mathrm{SDiss}(p,q)$$

7.4. Exponents.

Theorem 7.5 *The following conditions are equivalent*

— $\exp\{tX\} \in \Gamma\mathrm{U}(p,q)$ *for all* $t > 0$

— $X \in \mathrm{Diss}(p,q)$

PROOF. Let $\exp\{tX\} \in \Gamma\mathrm{U}(p,q)$, then

$$0 \geqslant \langle v\exp\{\varepsilon X\}, v\exp\{\varepsilon X\}\rangle - \langle v,v\rangle = \varepsilon\big(\langle vX,v\rangle + \langle v,vX\rangle\big) + O(\varepsilon^2) \qquad (7.3)$$

The left-hand side is nonpositive, hence X is dissipative.

Conversely, let X be dissipative. Then

$$\langle v\exp\{tX\}, v\exp\{tX\}\rangle - \langle v,v\rangle =$$

$$= \sum_{j=0}^{N-1}\Big[\langle v\exp\{\frac{t(j+1)}{N}X\}, v\exp\{\frac{t(j+1)}{N}X\}\rangle - \langle v\exp\{\frac{tj}{N}X\}, v\exp\{\frac{tj}{N}X\}\rangle\Big] =$$

$$= \sum_{j=0}^{N-1}\Big[\langle v_j\exp\{\frac{t}{N}X\}, v_j\exp\{\frac{t}{N}X\}\rangle - \langle v_j,v_j\rangle\Big], \qquad \text{where } v_j = v\exp\{\frac{tj}{N}X\}$$

By (7.3), each summand $\leqslant \mathrm{const}(X,t,v)N^{-2}$. Passing to the limit as $N \to \infty$, we obtain that our expression is non-positive. □

Proposition 7.6 *Consider a smooth curve* $\gamma(t) \in \mathrm{GL}(p+q,\mathbb{C})$, *where* $t \geqslant 0$ *and* $\gamma(0) = 1$.

a) If $\gamma(t) \in \Gamma\mathrm{U}(p,q)$, *then* $\gamma'(0)$ *is dissipative.*

b) If $\gamma(t)^{-1}\gamma'(t)$ *is dissipative for all* $t > 0$, *then* $\gamma(t) \in \Gamma\mathrm{U}(p,q)$.

PROOF is same. □

The statement a) means that the wedge of dissipative operators is the tangent cone to the semigroup $\Gamma\mathrm{U}(p,q)$ at unit.

Problem 7.2 Curves $\gamma(t)$ as in Proposition 7.6 sweep out the whole semigroup $\Gamma\mathrm{U}(p,q)$ (see Theorem 7.7 below).

7.5. The Potapov-Olshanski decomposition.

Problem 7.3 Let g be an invertible operator in a pseudo-Euclidean space. Is it correct that g admits a unique decomposition $g = UR$, where U is pseudo-unitary, and R is pseudo-Hermitian? ANSWER: NO.

Theorem 7.7 *Each element* $g \in \Gamma\mathrm{U}(p,q)$ *admits a unique decomposition*

$$g = \exp(X)\cdot h, \qquad (7.4)$$

where $h \in \mathrm{U}(p,q)$, *and* $X \in \mathrm{SDiss}(p,q)$. *In particular,*

$$\Gamma\mathrm{U}(p,q) \simeq \mathrm{U}(p,q)\times\exp(\mathrm{SDiss}(p,q)) \text{ canonically as topological space}$$

We use the following Lemma that will become obvious in the next section, see Theorem 8.2.

Lemma 7.8 *The semigroup* $\Gamma\mathrm{U}(p,q)$ *is connected.*

PROOF OF THEOREM. Let us try to repeate the proof of Theorem 1.1.3 about polar decomposition. We have

$$\langle v,v\rangle \geqslant \langle vg,vg\rangle = \langle v, vgg^{\circledast}\rangle \tag{7.5}$$

The element $gg^{\circledast}$ is pseudo-self-adjoint ($(gg^{\circledast})^{\circledast} = gg^{\circledast}$). We must show the existence of a square root $\sqrt{gg^{\circledast}}$.

Denote by $\mathfrak{R}$ the set of all the (invertible) self-adjoint operators in a pseudo-Euclidean space satisfying the condition $\langle vr,v\rangle \leqslant \langle v,v\rangle$. We wish to list indecomposable elements of this set.

Lemma 7.9 *Indecomposable operators* $r\in\mathfrak{R}$ *exist only in spaces* W *of dimension 1 and 2 and have one of the following forms*

1°. $W=\mathbb{C}^{0,1}$ $wr=\lambda w$, $|\lambda|\leqslant 1$, $\lambda\neq 0$

2°. $W=\mathbb{C}^{1,0}$, $wr=\lambda w$, $\lambda\geqslant 1$.

3°. $W=\mathbb{C}^{1,1}$, *the operator is the Jordan box* $(\pm A_t)$, *where* A_t *is given by*[17] *(7.1), (7.2).*

PROOF OF LEMMA 7.9. Let us filtrate canonical forms of pseudo-self-adjoint operators. For instance, consider the operator $\mathrm{I}_3^+(\lambda)$ (in notation of Theorem 6.2). For $v=\sum\alpha_j e_j$,

$$\langle v,v\rangle - \langle vr,v\rangle = \begin{pmatrix}\alpha_1 & \alpha_2 & \alpha_3\end{pmatrix}\left\{\begin{pmatrix}0&0&1\\0&1&0\\1&0&0\end{pmatrix} - \begin{pmatrix}0&1&\lambda\\1&\lambda&0\\\lambda&0&0\end{pmatrix}\right\}\begin{pmatrix}\alpha_1\\\alpha_2\\\alpha_3\end{pmatrix}$$

The matrix in the curly brackets is not positive definite (indeed it has too much zeros on the diagonal), see Example 4 from Subsection 7.1).

The last argument is valid in all the other cases. For operators of the type II_n the diagonal consists of zeros. For operators of the type $\mathrm{I}_n^\pm$ our argument is valid for all $n\geqslant 3$. The case $n=2$ require a straightforward calculation (see Example 3 in the beginning of this section). □

Lemma 7.9 itself does not allow to extract a canonical square root, since r admits negative eigenvalues.

Lemma 7.10 *If* $r=gg^{\circledast}$, *then all the eigenvalues of* r *are positive.*

PROOF All the eigenvalues of an operator $r\in\mathfrak{R}$ are real. Therefore the set $\mathfrak{R}$ is disconnected. But the semigroup $\Gamma\mathrm{U}(p,q)$ is connected, hence the image of

[17] for all $t>0$ the operators A_t are equivalent

the map $g \mapsto gg^{\circledast}$ is connected. Since $1 \cdot 1^{\circledast}$ has positive eigenvalues, an arbitrary matrix $gg^{\circledast}$ also has positive eigenvalues. $\square$

END OF PROOF OF THEOREM 7.7. Thus, the operator $gg^{\circledast}$ is a direct sum of several operators enumerated in Lemma 7.9 with positive eigenvalues (i.e., $\lambda > 0$ and $(-A_t)$ is forbiden). In all these cases, there is a canonical square root and a canonical logarithm; for the two dimensional Jordan box A_t, we have $\sqrt{A_t} = A_{t/2}$, $\ln A_t = \begin{pmatrix} 0 & -t \\ 0 & 0 \end{pmatrix}$.

Hence, there is a canonical square root $y := \sqrt{gg^{\circledast}}$ that is pseudo-contractive and has positive eigenvalues.

$$y^{\circledast} = y, \qquad y \in \Gamma\mathrm{U}(p,q), \qquad \langle vy, v\rangle \leqslant \langle v, v\rangle$$

It remains to repeat the proof of Theorem 1.3 about polar decomposition. $\square$

ANOTHER SEMI-PROOF OF THEOREM. For a strict contractions g, the theorem admits a simple proof avoiding a reference to canonical forms.

By the Krein Theorem formulated in the next subsection, one can find a negative r-invariant p-dimensional subspace H. Since r is self-adjoint, the positive subspace $H^{\perp}$ also is r-invariant. Therefore r has the form $\begin{pmatrix} A & 0 \\ 0 & B \end{pmatrix}$, where A, B are self-adjoint in the usual sense (the both spaces H, $H^{\perp}$ are Euclidean). Now $A > 1$, $-1 < B < 1$. We again apply the argument with connectedness and extract a square root. $\square$

REMARK. It seems that after the last considerations we can apply a continuity arguments. This way is possible but not simple. A behavior of a matrix square root under a limit passage is bad (for instance, the set of solutions of the equation $A^2 = 1$ is unbounded). $\square$

We also get the following corollary

Proposition 7.11 *Each element $g \in \Gamma\mathrm{U}(p,q)$, admits a decomposition $g = h_1 B h_2$, where h_1, $h_2 \in \mathrm{U}(p,q)$ and B is a block-diagonal matrix, whose blocks (having size 1×1 or 2×2) are matrices* with positive eigenvalues *enumerated in Lemma 7.9.*

For a strict contraction g, the operator B is diagonal; the eigenvalues μ_i corresponding negative eigenvectors are > 1, the eigenvalues λ_j corresponding positive eigenvectors satisfy $0 < \lambda_j < 1$.

Problem 7.4 Equip $\mathbb{C}^{p,q}$ with the standard scalar product. Then $g \in \Gamma\mathrm{U}(p,q)$ implies $g^* \in \Gamma\mathrm{U}(p,q)$.

Problem* 7.5 a) The semigroup of contractions of $\mathbb{R}^{p,q}$ has 4 connected components.

b) In spite of this, Theorem 7.7 survives in this case.

7.6. Action of indefinite contractions on a matrix ball. Let $S \in \mathcal{M}_{p,q}$ be a negative subspace, let $g \in \Gamma\mathrm{U}(p,q)$. Obviously, the subspace Sg also is negative.

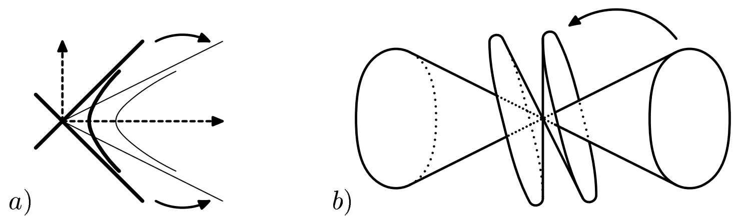

Fig. 2.17: a) The operator $(x,y) \mapsto (2x,y)$ in $\mathbb{R}^{1,1}$.
b) A contractive operator in $\mathbb{R}^{2,1}$.

Thus the semigroup $\Gamma\mathrm{U}(p,q)$ acts on $\mathcal{M}_{p,q}$ and on the matrix ball $\mathrm{B}_{p,q}$; the latter action is given by the same formula (3.3).

Theorem 7.12 (Krein) *Each element $g \in \Gamma\mathrm{U}(p,q)$ has a semi-negative invariant subspace of dimension p. If g is strictly contractive, then it has a negative invariant subspace.*

PROOF. We consider the transformation $z \mapsto z^{[g]}$ of the closed matrix ball $\overline{\mathrm{B}}_{p,q}$. By the Brouwer fixed point theorem[18], this transformation has a fixed point. This point corresponds to an element of $\mathcal{M}_{p,q}$.

If g is a strict contraction, then g-image of a semi-positive subspace is a negative subspace. Hence a fixed point corresponds to a negative subspace. □

7.7. A reformulation of contractivity. Denote by $C_- \subset \mathbb{C}^{p,q}$ the cone of non-positive vectors.

Theorem 7.13 *Let $AC_- \subset C_-$. Then there is a positive constant s such that $s \cdot A$ is a contraction.*

Lemma 7.14 *Under the conditions of Theorem, let e be a negative vector such that $\langle e,e\rangle = -1$ and let f be a positive vector such that $\langle f,f\rangle = +1$. Then*

$$-\langle Ae, Ae\rangle \geqslant \langle Af, Af\rangle$$

PROOF OF LEMMA. The vectors $e \pm f$ are isotropic, therefore $A(e \pm f) \in C_-$. Hence

$$0 \geqslant \langle A(e+f), A(e+f)\rangle + \langle A(e-f), A(e-f)\rangle = 2(\langle Ae, Ae\rangle + \langle Af, Af\rangle) \qquad \square$$

PROOF OF THEOREM. We consider a constant λ separating all $(-\langle Ae, Ae\rangle)$ from all $\langle Af, Af\rangle = 1$. Next, consider $\lambda^{-1}A$. □

7.8. Digression. The cone of self-adjoint dissipative operators.

[18]Each continuous map from the closed m-dimensional ball to itself has a fixed point

Proposition 7.15 *a) Each indecomposable self-adjoint dissipative matrix has one of the following forms (we present the space, the inner product, and the operator).*

— *the space is* $\mathbb{C}^{0,1}$ *with a basis* e

$$\langle e, e\rangle = 1, \qquad Ae = -s\cdot e, \qquad \text{with } s \geqslant 0.$$

— *the space is* $\mathbb{C}^{1,0}$ *with a basis* e

$$\langle e, e\rangle = -1, \qquad Ae = s\cdot e, \qquad \text{with } s \geqslant 0.$$

— *the space is* $\mathbb{C}^{1,1}$ *with a basis* e, f,

$$\langle e, f\rangle = 1, \quad e,\ f \text{ are isotropic, } Ae = -tf,\ Af = 0, \text{ where } t > 0$$

(operators with different t *are equivalent).*

b) Each operator $B \in \mathrm{SDiss}$ *is a direct sum of indecomposable operators.*

PROOF. a) We must merely sort out the canonical forms of self-adjoint operators A from Theorem 6.2.

For the type II_n, the Hermitian form $\langle Av, v\rangle$ has a matrix of the shape $\begin{pmatrix} 0 & J \\ J^* & 0\end{pmatrix}$. It is not negative semi-definite, since it has zeros on the diagonal.

For the type $\mathrm{I}_n^{\pm}$, the form $\langle Av, v\rangle$ is

$$\pm\begin{pmatrix} 0 & 0 & \dots & 1 & \lambda \\ 0 & 0 & \dots & \lambda & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & \lambda & \dots & 0 & 0 \\ \lambda & 0 & \dots & 0 & 0 \end{pmatrix}$$

It has too much zeros on the diagonal. The only cases then it can be non-negative definite are $n = 1$ or $n = 2$, $\lambda = 0$. $\square$

Problem 7.6 a) Describe the boundary of $\mathrm{SDiss}(p, q)$

b) Find extreme points of the cone $\mathrm{SDiss}(p, q)$.

2.8 Potapov coordinates

8.1. The Potapov transform. Let $V = V_1 \oplus V_2$ be a linear space, $\dim V_1 = p$, $\dim V_2 = q$. Consider a linear transformation in V,

$$\begin{pmatrix} w_1 \\ w_2 \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \tag{8.1}$$

Write out these relations in the form

$$\begin{aligned} w_1 &= av_1 + bv_2 \\ w_2 &= cv_1 + dv_2 \end{aligned}$$

Let us regard these equations as equations for the variables v_1, w_2. The first equation implies

$$v_1 = a^{-1}w_1 - a^{-1}bv_2$$

Substituting this v_1 to the second equation, we obtain

$$\begin{pmatrix} v_1 \\ w_2 \end{pmatrix} = \begin{pmatrix} -a^{-1}b & a^{-1} \\ d - ca^{-1}b & ca^{-1} \end{pmatrix} \begin{pmatrix} v_2 \\ w_1 \end{pmatrix} \tag{8.2}$$

We say that the matrix

$$\Pi(g) = \begin{pmatrix} -a^{-1}b & a^{-1} \\ d - ca^{-1}b & ca^{-1} \end{pmatrix} \quad \textit{is the Potapov transform of} \quad \begin{pmatrix} a & b \\ c & d \end{pmatrix} \tag{8.3}$$

REMARK. The domain of definitemess of the Potapov transform is $\det a \neq 0$.

REMARK. The size of the matrix $\Pi(g)$ is $(p+q) \times (q+p)$, i.e., the blocks on the second diagonal are squares, i.e., $p \times p$, $q \times q$.

8.2. The inverse transform. It can be readily checked that

$$\Pi(g) = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \qquad \Longrightarrow \qquad g = \begin{pmatrix} \beta^{-1} & -\beta^{-1}\alpha \\ \delta\beta^{-1} & \gamma - \delta\beta^{-1}\alpha \end{pmatrix} \tag{8.4}$$

We will refer to this operation as *inverse Potapov transform* $\Pi^{-1}(g)$.

Formulae for the direct transform (8.3) and for the inverse transform (8.4) seem to be suspiciously similar. An explanation of this similarity is very simple. Let us represent the system (8.2) in the form

$$\begin{pmatrix} v_1 \\ w_2 \end{pmatrix} = \begin{pmatrix} a^{-1} & -a^{-1}b \\ ca^{-1} & d - ca^{-1}b \end{pmatrix} \begin{pmatrix} w_1 \\ v_2 \end{pmatrix} \tag{8.5}$$

and denote the matrix in the right hand side by $\widetilde{\Pi}(g)$. Then

$$\widetilde{\Pi}\big(\widetilde{\Pi}(g)\big) = g$$

since the second transformation $\widetilde{\Pi}$ means that we solve the system (8.5) with respect to the variables w_1, w_2 considering v_1, v_2 as parameters. Obviously, we revert to (8.1)

8.3. Coordinates-I. We realize the group $\mathrm{U}(p,q)$ in the standard way (2.3).

Theorem 8.1 *a) If* $g \in \mathrm{U}(p,q)$*, then* $\Pi(g)$ *is well defined. Moreover,* $\Pi(g) \in \mathrm{U}(p+q)$.

b) Conversely, the image $\Pi(\mathrm{U}(p,q))$ *consists of matrices* $r = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \in \mathrm{U}(p+q)$ *satisfying the four following equivalent conditions*

1. β is invertible　　　*3.* $\|\alpha\| < 1$

2. γ is invertible　　　*4.* $\|\delta\| < 1$

PROOF. a) Consider a pseudo-unitary transformation g given by

$$\begin{pmatrix} w_1 \\ w_2 \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$$

The pseudo-unitarity condition means

$$-\|w_1\|^2 + \|w_2\|^2 = -\|v_1\|^2 + \|v_2\|^2 \tag{8.6}$$

therefore

$$\|v_1\|^2 + \|w_2\|^2 = \|v_2\|^2 + \|w_1\|^2 \tag{8.7}$$

However, this is precisely the condition of unitarity of $\Pi(g)$.

Problem 8.1 Show that a is invertible. We can refer to the defining relations (2.4). We can also refer to (8.6). It sufficies to take $v_1 \in \ker a$, $v_2 = 0$ and to look on the left-hand side.

b) First, we must show that $\Pi^{-1}(r) \in \mathrm{U}(p,q)$. The inverse transform is defined if β is invertible. It remains to note that (8.7) implies (8.6).

Secondly, we wish to verify the equivalence of the conditions 1-4. A unitary matrix preserves the norm of any vector, in particular, it preserves the norm of $x \oplus 0 \in \mathbb{C}^{p+q}$,

$$\|\alpha x \oplus \gamma x\|^2 = \|x\|^2 \quad \Longrightarrow \quad \|\alpha x\|^2 + \|\gamma x\|^2 = \|x\|^2 \tag{8.8}$$

If $\|\alpha\| = 1$, then there is a vector x such that $\|\alpha x\|^2 = \|x\|^2$. Hence $\|\gamma x\|^2 = 0$ and γ is not invertible. This consideration proves $2 \Leftrightarrow 3$.

Repeating the same trick to $0 \oplus y$, we obtain $1 \Leftrightarrow 4$. Applying the same considerations to vector-rows, we obtain the remaining equivalences. $\square$

Problem 8.2 Thus the proof seems to be trivial. Prove statement a) directly starting from the explicit formula (8.3) and the defining relation (2.4) for $\mathrm{U}(p,q)$.

Theorem 8.2 *a) If $g \in \Gamma\mathrm{U}(p,q)$, then $\|\Pi(g)\| \leqslant 1$.*

b) The image $\Pi(\mathrm{U}(p,q))$ consists of matrices $r = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$ satisfying $\|r\| \leqslant 1$ and the conditions 1-4 of the previous theorem.

c) Moreover, for a matrix r with $\|r\| \leqslant 1$, each of the conditions 1 or 2 implies the both 3 and 4. Therefore it is sufficient to require $r \in \overline{\mathrm{B}}_{p+q,p+q}$ and existence of β^{-1}, γ^{-1}.

d) $g \in \Gamma\mathrm{U}(p,q)$ is a strict contraction, iff $\|\Pi(g)\| < 1$.

PROOF of the previous theorem survives literally. To prove a) we observe that

$$-\|w_1\|^2 + \|w_2\|^2 \leqslant -\|v_1\|^2 + \|v_2\|^2 \qquad \Longleftrightarrow \qquad \|v_1\|^2 + \|w_2\|^2 \leqslant \|v_2\|^2 + \|w_1\|^2$$

The former condition is the condition of pseudo-contractivity, the latter is the condition of the usual (Euclidean) contractivity.

Further, look to (8.3). By the construction, $\beta = a^{-1}$, therefore it is invertible. Also g is invertible. By Theorem 1.1.4,

$$0 \neq \det(g) = \det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \det a \cdot \det(d - ca^{-1}b) \quad \Longrightarrow \quad \gamma \text{ is invertible.}$$

Also,

$$\|\alpha x \oplus \gamma x\|^2 \leqslant \|x\|^2 \quad \Longrightarrow \quad \|\alpha x\|^2 + \|\gamma x\|^2 \leqslant \|x\|^2$$

if $\|\alpha\| = 1$, then γ is not invertible. □

Problem 8.3 Reconstruct remaining details.

Corollary 8.3 *The semigroup* $\Gamma\mathrm{U}(p,q)$ *is connected.*

PROOF. The set described in Theorem 8.2.b is connected. □

8.4. Products of matrices in the Potapov coordinates.

Theorem 8.4 *Let* g_1, g_2 *be* $(p+q) \times (p+q)$*-matrices. Let*

$$\Pi(g_1) = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}, \qquad \Pi(g_2) = \begin{pmatrix} \varphi & \psi \\ \theta & \varkappa \end{pmatrix}$$

Then

$$\Pi(g_1 g_2) = \begin{pmatrix} \alpha + \beta(1 - \varphi\delta)^{-1}\varphi\gamma & \beta(1 - \varphi\delta)^{-1}\psi \\ \theta(1 - \delta\varphi)^{-1}\gamma & \varkappa + \theta\delta(1 - \varphi\delta)^{-1}\psi \end{pmatrix} \tag{8.9}$$

REMARK. Apparently, this formula is not the simplest way to multiply matrices. However, in the next section we observe that this formula corresponds to a more general operation.

PROOF. We write out

$$\begin{pmatrix} w_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \begin{pmatrix} w_2 \\ v_1 \end{pmatrix}, \qquad \begin{pmatrix} v_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} \varphi & \psi \\ \theta & \varkappa \end{pmatrix} \begin{pmatrix} v_2 \\ u_1 \end{pmatrix}$$

or,

$$\begin{aligned} w_1 &= \alpha\, w_2 + \beta\, v_1, \\ v_2 &= \gamma\, w_2 + \delta\, v_1, \\ v_1 &= \varphi\, v_2 + \psi\, u_1, \\ u_2 &= \theta\, v_2 + \varkappa\, u_1. \end{aligned}$$

We wish to express the variables w_1, u_2 in the terms of the variables w_2, u_1. For this purpose, we exclude v_1, v_2 from the second and third equations,

$$v_1 = (1 - \varphi\delta)^{-1}(\varphi\gamma\, w_2 + \psi\, u_1), \qquad v_2 = (1 - \delta\varphi)^{-1}(\varphi\gamma\, w_2 + \psi\, u_1)$$

and substitute these v_1, v_2 to the first and forth equation. □

8.5. Action on matrix balls.

Theorem 8.5 *For $g \in \Gamma\mathrm{U}(p,q)$ denote by $\Pi(g) = \begin{pmatrix} \alpha & \beta \\ \delta & \gamma \end{pmatrix}$ its Potapov transform Let $z \in \mathrm{B}_{p,q}$. Then g transfers z to*

$$g: \quad z \mapsto \alpha + \beta(1 - z\delta)^{-1} z\gamma \tag{8.10}$$

Such maps are called by Krein–Schmul'yan maps.

PROOF. Now we have the equations

$$w_1 = zw_2 \tag{8.11}$$
$$v_1 = \alpha v_2 + \beta w_1 \tag{8.12}$$
$$w_2 = \gamma v_2 + \delta w_1 \tag{8.13}$$

First, we substitute w_1 from (8.11) to (8.13) and get

$$w_2 = \gamma v_2 + \delta z w_2 \qquad \Longrightarrow \qquad w_2 = (1 - \delta z)^{-1} \gamma v_2$$

Using (8.11) again, we get

$$w_1 = z(1 - \delta z)^{-1} \gamma v_2$$

Substituting this w_1 to (8.12), we obtain

$$v_1 = \Bigl(\alpha + \beta z (1 - \delta z)^{-1} \gamma\Bigr) v_2 \qquad \square$$

8.6. Digression. Some remarks on formulae (8.9) and (8.10). It is interesting to compare the formula (8.10) for Krein–Schmul'yan transformations and product formula (8.9). First, the expression (8.10) is present in the left upper block of (8.9).

Secondly, the expression (8.9) for a product can be represented in the form

$$\begin{pmatrix} \alpha & 0 \\ 0 & \varkappa \end{pmatrix} + \begin{pmatrix} \beta & 0 \\ 0 & \theta \end{pmatrix} \left[\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \delta & 0 \\ 0 & \varphi \end{pmatrix} \right]^{-1} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \gamma & 0 \\ 0 & \psi \end{pmatrix} \tag{8.14}$$

and in the form

$$\begin{pmatrix} \alpha & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} \beta & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \varphi & \psi \\ \theta & \varkappa \end{pmatrix} \left[\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} - \begin{pmatrix} \delta & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \varphi & \psi \\ \theta & \varkappa \end{pmatrix} \right]^{-1} \begin{pmatrix} \gamma & 0 \\ 0 & 1 \end{pmatrix} \tag{8.15}$$

These formulae looks like a special case of the Krein–Schmul'yan map (8.10).

Thus the both formulas (8.9) and (8.10) looks as particular cases of one another. We return to this analogy in Subsection 9.4.

8.7. Coordinates-II. Let $p = q$. Realize the group $\mathrm{U}(q,q)$ as the group of matrices $g = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$ preserving the Hermitian form $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$, see (3.15). We introduce the following transform

$$\begin{pmatrix} w_1 \\ w_2 \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} \quad \Longleftrightarrow \quad \begin{pmatrix} w_1 \\ v_1 \end{pmatrix} = \begin{pmatrix} AC^{-1} & -B + AC^{-1}D \\ C^{-1} & C^{-1}D \end{pmatrix} \begin{pmatrix} w_2 \\ -v_2 \end{pmatrix}$$

For a matrix g in the left-hand side, we denote the resulting matrix by $\Pi^\circ(g)$. This matrix differs from $\Pi(g)$ by a permutation of blocks and changing signs.

Theorem 8.6 *a) Let* $g = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \in \mathrm{U}(q,q)$, *let* C *be invertible. Then* $\Pi^\circ(g)$ *is an anti-Hermitian* $(q+q)\times(q+q)$*-matrix.*

c) The image $\Pi^\circ(\mathrm{U}(q,q))$ *consists of anti-Hermitian matrices* $\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$ *such that the block* γ *is invertible.*

Thus we obtain an open dense map on $\mathrm{U}(q,q)$.

Problem 8.4 Verify the identity $\bigl[\Pi^\circ(g)\bigr]^* = -\Pi^\circ(g)$ directly from the defining equations (3.15).

Theorem 8.7 *a) Let* $g = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \in \Gamma\mathrm{U}(q,q)$, *let* C *be invertible. Then* $\mathrm{Re}\,\Pi^\circ(g) \leqslant 0$.

b) The map Π° *is a bijection of the semigroup of strict contractions into the subset in the open wedge* $\mathrm{Re}\,r > 0$ *distinguished by the costrain:* β, γ *are invertible.*

We omit proofs since they are repetitions of the same ideas.

2.9 Krein–Shmul'yan category. Compressivity

Starting this section, we use linear relations, see above Section 1.7 which is independent on other sections of Chapter 1.

9.1. Contractive linear relations. Let V, W be pseudo-Euclidean spaces with inertia indices (p,q), (r,s) (in our notation, p, r are negative indices). A linear relation $P : V \rightrightarrows W$ is called *contractive* if

a) If $v \oplus w \in P$, then $\langle v, v\rangle_V \geqslant \langle w, w\rangle_W$.

b) $\dim P = q + r$, i.e., this is the maximal possible dimension of a linear satisfying the condition a).

Let us say the same in other words. Consider the space $V \oplus W$ equipped with the Hermitian form

$$\langle v \oplus w, v' \oplus w'\rangle_{V\oplus W} := \langle v, v'\rangle_V - \langle w, w'\rangle_W \tag{9.1}$$

A contractive linear relation $V \rightrightarrows W$ is a maximal semi-positive subspace in $V \oplus W$.

Theorem 9.1 *A product of contractive linear relations is a contractive linear relation*

Proof. Let $P : V \rightrightarrows W$, $Q : W \rightrightarrows Y$ be contractive. Let $v \oplus y \in QP$. Then one can find $w \in W$ such that $v \oplus w \in P$, $w \oplus y \in Q$. Hence

$$\langle v, v\rangle_V \geqslant \langle w, w\rangle_W \geqslant \langle y, y\rangle_Y$$

Therefore the condition a) is valid for QP.

Lemma 9.2 *For a contractive linear relation $P : V \rightrightarrows Y$,*

$$\operatorname{dom} P \subset (\mathit{zero} \ker P)^{\perp}, \qquad \operatorname{im} P \subset (\mathit{zero} \operatorname{indef} P)^{\perp}$$

PROOF OF LEMMA. Let $\xi \in \mathit{zero} \ker P$. Since the isotropic vector $\xi \oplus 0$ is contained in a semi-positive subspace P, it follows that $\xi \oplus 0 \in \mathit{zero}\, P$. Let $v \oplus w \in P$. Then

$$0 = \langle \xi \oplus 0, v \oplus v \rangle_{V \oplus W} = \langle \xi, v \rangle_V$$

Thus, $v \perp \xi$. □

END OF PROOF OF THE THEOREM. By the lemma,

$$\operatorname{im} P \subset (\mathit{zero} \operatorname{indef} P)^{\perp}, \qquad \operatorname{dom} Q \subset (\mathit{zero} \ker Q)^{\perp}$$

Hence

$$\operatorname{dom} Q + \operatorname{im} P \subset \bigl(\mathit{zero} \ker Q \cap \mathit{zero} \operatorname{indef} P\bigr)^{\perp}$$

By (7.7),

$$\dim QP = \dim P + \dim Q - \dim\bigl(\ker Q \cap \operatorname{indef} P\bigr) - \dim\bigl(\operatorname{im} P + \operatorname{dom} Q\bigr)$$

Since $\ker Q$ is non-negative and $\operatorname{indef} P$ is non-positive, it follows that their intersection is $(\mathit{zero} \ker Q) \cap (\mathit{zero} \operatorname{indef} P)$. Thus

$$\begin{aligned} &\dim QP = \\ &= \dim P + \dim Q - \dim\bigl(\mathit{zero} \ker Q \ \cap \ \mathit{zero} \operatorname{indef} P\bigr) - \dim\bigl(\operatorname{im} P + \operatorname{dom} Q\bigr) \geqslant \\ &\qquad \geqslant \dim Q + \dim P - \dim W \end{aligned}$$

However, the last number is the maximal possible dimension of a semi-positive subspace in $V \oplus Y$. □

9.2. The Krein–Shmul'yan category. In spite of Theorem 9.1, it is natural to reduce just a little the category of contractive linear relations.

We define the *Krein–Shmul'yan category* as a category, whose objects are pseudo-Euclidean spaces and morphisms are contractive linear relations $P : V \rightrightarrows W$ satisfying the two additional conditions

c) $\ker P \supset V$ is strictly positive,

c)′ $\operatorname{indef} P \subset W$ is strictly negative.

REMARK. The condition a) implies semi-positivity of $\ker P$ and semi-negativity of $\operatorname{indef} P$.[19]

As above, we denote by $\operatorname{Mor}(V, W)$ the set of morphisms $V \to W$.

[19] This reduction is a matter of test, but there are some natural reasons for it. First of all, a product of contractive linear relations does not depend continuously on the factors; in either case the continuity can be saved by a single condition *c*) or by a single condition *c*)′. A more important reason is the L^2-boundedness conditions for Gaussian operators from Section 1.8. Also the central extension of the Krein–Schmul'yan category (see Subsection 3.7.3) does not survive for the category of contractive linear relations.

Now let $V = V_1 \oplus V_2$, $W_1 \oplus W_2$ be objects of the Krein–Shmul'yan category, let the subspaces V_1, W_1 be negative, and the subspaces V_2, W_2 positive. Then, by Theorem 3.1 a morphism $P : V \rightrightarrows W$ is a graph of a contractive operator in Euclidean spaces

$$\Pi(P) = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} : \; V_2 \oplus W_1 \to V_1 \oplus W_2$$

(here we change signs of scalar products on V_1, W_1 and regards both spaces as Euclidean).

Theorem 9.3 *$P \in \mathrm{Mor}(V, W)$ iff $\|\Pi(P)\| \leqslant 1$ and $\|\alpha\| < 1$, $\|\delta\| < 1$*

We leave a proof as an exercise. It is a rephrasing of Theorem 3.1.

Thus we get a parametrization of the set $\mathrm{Mor}(V, W)$; also we can say that we have a coordinate system on $\mathrm{Mor}(V, W)$.

Theorem 9.4 *The product in these coordinates is given by the formula (8.9)*

PROOF. Proof of Theorem 8.4 survives literally. □

REMARK. Certainly, formula (8.9) is not the best way to multiply block matrices g_1, g_2. Nevertheless, a product of linear relations is a more general operation than a matrix multiplication; our formula allows to express this product in the terms of matrix operations. □

9.3. Digression. The involution in Krein-Shmul'yan category. Let $P : V \rightrightarrows W$ be a morphism of the Krein–Shmul'yan category. We define the linear relation $P^\sigma : W \rightrightarrows V$ by

$$P^\sigma := (P^\perp)^\square$$

where $P^\square$ is the pseudo-inverse linear relation.

Proposition 9.5 *a) P^σ is a morphism $W \rightrightarrows V$.*

b) For $P : V \rightrightarrows W$, $Q : W \rightrightarrows Y$, $(QP)^\sigma = P^\sigma Q^\sigma$

We omit a proof (see parallel considerations in Subsection 1.9.7).

This operation is an imitation of the operation $A \mapsto A^{\circledast}$.

9.4. The Krein–Shmul'yan functor. Now, let $R \subset V$ be a negative subspace, let z be its coordinate. Let $P : V \rightrightarrows W$ be a morphism of the Krein–Shmul'yan category, denote by $\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}$ its Potapov transform.

Theorem 9.6 *a) The map[20] $R \mapsto RP$ is given by the formula*

$$\tau_P : z \mapsto \alpha + \beta(1 - z\delta)^{-1} z\gamma$$

In particular,

$$\|\alpha + \beta(1 - z\delta)^{-1} z\gamma\| < 1 \tag{9.2}$$

[20] We apply a linear relation to a linear subspace

b) This is a functor, i.e.,

$$\tau_P \tau_Q = \tau_{PQ} \tag{9.3}$$

PROOF. This is a special case of the formula (8.9). Indeed, R can be considered as a morphism $0 \to V$. $\square$

Denote by (p, q) the inertia indices of V, and by (r, s) the inertia indices of W. We get a map $\tau_P : \mathrm{B}_{p,q} \to \mathrm{B}_{r,s}$, or

$$\tau_P : \mathrm{U}(p,q)/\mathrm{U}(p) \times \mathrm{U}(q) \to \mathrm{U}(r,s)/\mathrm{U}(r) \times \mathrm{U}(s)$$

Maps of this form are called *Krein–Shmul'yan maps.*

REMARK. This explains a half of the strange observation from 8.6; the formula for Krein–Shmul'yan maps is a special case of the product-formula in the Potapov coordinates. We also are ready to discuss the another half.

Let V, W, Y be arbitrary pseudo-Euclidean spaces, let $P : V \rightrightarrows W$, $Q : W \rightrightarrows Y$ be linear relations. Define the linear relation

$$\widetilde{Q} := 1 \oplus Q : V \oplus W \rightrightarrows V \oplus Y$$

Applying the linear relation $\widetilde{Q}$ to the subspace $P \subset V \oplus W$, we obtain the subspace $\widetilde{Q}P \subset V \oplus Y$, which is nothing but QP. This explains formula (8.15).

Formula (8.14) arises from another rephrasing of a product of linear relations. Consider the linear relation

$$P^{\square} \oplus Q : W \oplus W \rightrightarrows V \oplus Y$$

and the diagonal $S \subset W \oplus W$. Applying $P^{\square} \oplus Q$ to S we obtain a subspace in $V \oplus Y$ which is nothing but QP. $\square$

9.5. Compression of angles.

Theorem 9.7 *Consider a Krein–Shmul'yan map $\tau : \mathrm{B}_{p,q} \to \mathrm{B}_{r,s}$. For z, $u \in \mathrm{B}_{p,q}$, denote by $\varphi_1 \geqslant \varphi_2 \geqslant \dots$ the hyperbolic angles between z, u. Let $\psi_1 \geqslant \psi_2 \geqslant \dots$ be the hyperbolic angles between $\tau(z)$ and $\tau(u)$. Then*

$$\varphi_1 \geqslant \psi_1, \quad \varphi_2 \geqslant \psi_2, \quad \varphi_3 \geqslant \psi_3 \dots \tag{9.4}$$

Proof is given in Subsection 9.7.

9.6. Example. Contractive maps of unit disk. Consider the space $\mathbb{C}^{1,1}$. The 'matrix ball' $B_{1,1}$ is the unit disk D: $|z| < 1$ on $\mathbb{C}$. The semigroup $\Gamma\mathrm{U}(1,1)$ acts on D by the (injective) linear fractional transformations

$$z \mapsto \frac{b + zd}{a + zc} \tag{9.5}$$

This semigroup contains the Möbius group $\mathrm{SU}(1,1)/\{\pm 1\}$ of biholomorphic transformations of the disk; this corresponds to the case $d = \overline{a}$, $c = \overline{b}$, $|a|^2 - |b|^2 = 1$, see Subsection 2.9.

The Krein–Schmul'yan maps $D_1 \to D_1$ are

$$z \mapsto \alpha + \frac{\beta\gamma z}{1 - \delta z}$$

If $\beta\gamma \neq 0$, then this transformation can be reduced to (9.5). Otherwise, we obtain Krein–Schmulian maps that are not linear fractional, namely $z \mapsto \alpha$.

The disk D is endowed with a $\mathrm{U}(1,1)$-invariant Riemannian metric

$$(1 - |z|^2)^{-2}\, dz\, d\overline{z}$$

A geodesic distance is given by the formula

$$\rho(z, u) = \varphi(z, u)$$

where φ is the hyperbolic angle, see (5.6).

Lemma 9.8 *Transformation $g \in \Gamma\mathrm{U}(1,1)$ are contractive, i.e.,*

$$\rho(z^{[g]}, u^{[g]}) \leqslant \rho(z, u) \tag{9.6}$$

PROOF. It sufficies to prove the statement for strict contractions. By the Potapov–Olshanski Theorem 7.7, it follows that each strict contraction is a product $g = h_1 r_t h_2$, where h_1, h_2 are in the Möbius group, and r_t is a homothety

$$z \mapsto tz, \qquad \text{where } 0 < t < 1$$

The following problem finishes a proof. □

Problem 9.1 Show, that r_t contracts the Riemannian metric.

9.7. Proof of Theorem 9.7. We must look at the minimax expression (5.4) for angles. Let L_1, $L_2 \in \mathcal{M}_{p,q}$ be negative subspaces, let P be a linear relation

$$a_1 \in L_1,\ a_2 \in L_2,\ a_1 \oplus b_1 \in P,\ a_2 \oplus b_2 \in P \Longrightarrow b_1 \in PL_1,\ b_2 \in PL_2$$

The angles φ_j between L_1, L_2 and the angles ψ_j between PL_1, PL_2 are given by (5.4).

We have $\cosh \psi_j \geqslant 1$. For this reason, in minimax expression we watch only pairs b_1, b_2 generating indefinite planes. Hence a_1, a_2 must also generate an indefinite plane (since L_j are negative, the vectors a_j are negative; there is a logical possibility that a_1, a_2 generate a negative plane, but after the contraction P it can not became indefinite).

By Lemma 9.8, we have $\angle(b_1, b_2) \leqslant \angle(a_1, a_2)$. Looking to (5.4) we obtain the desired inequality. □

9.8. Remarks. For maps $\mathrm{B}_{1,1} \to \mathrm{B}_{1,1}$, Theorem 9.7 means only the contractivity in the usual sense, i.e., a Krein–Shmul'yan map is a contraction of the geodesic distance. Certainly, one can find a vast number of contractive maps of the disk into itself.

The same holds for maps of a ball $\mathrm{B}_{1,q}$ into itself. Nevertheless for $p>1$, $q>1$ a picture changes drastically; the compressivity becomes a very rigid property.

There are several variants of the inverse theorem, for instance,

Let τ be a holomorphic map $\mathrm{B}_{p,q} \to \mathrm{B}_{r,s}$ satisfying the conclusion of Theorem 9.7. If there are two points z, u such that $\psi_2(\tau(z),\tau(u)) \ne 0$[21], then $\tau(z)$ is either a Krein–Shmul'yan map or a map of a form $z \to \gamma(z^t)$, where γ is a Krein–Shmul'yan map.

2.10 Isotropic category. Inverse limits

10.1. Isotropic relations. Fix $m = 0,1,2,\dots$. Let us consider pseudo-Euclidean spaces of the form $\mathbb{C}^{p,p+m}$.

Let $V = \mathbb{C}^{p,p+m}$, $W = \mathbb{C}^{q,q+m}$. We say that a linear relation $P : \mathbb{C}^{p,p+m} \rightrightarrows \mathbb{C}^{q,q+m}$ is *form-preserving* if

a) For each $v \oplus w$, $v' \oplus w' \in P$, we have

$$\langle v, v'\rangle_V = \langle w, w'\rangle_W$$

b) $\dim P = p+q+m$, i.e., it is the maximal possible dimension of a subspace satisfying the condition a)..

We can also introduce the Hermitian form in $V \oplus W$ as above (9.1), then a form-preserving linear relation is nothing but an isotropic subspace. In particular, a form-preserving linear relation is a contractive linear relation in the sence of the previous section.

Theorem 10.1 *Let $P : V \rightrightarrows W$, $Q : W \rightrightarrows Y$ be form-preserving linear relations. Then $QP : V \rightrightarrows Y$ is form-preserving.*

PROOF. Let $v \oplus y$, $v' \oplus y' \in QP$. Then one can find w, $w' \in W$ such that

$$v \oplus w \in P,\ v' \oplus w' \in P \qquad w \oplus y \in Q,\ w' \oplus y' \in Q$$

Hence

$$\langle v, v'\rangle = \langle w, w'\rangle = \langle y, y'\rangle$$

This yeilds the condition a).

But P and Q are also contractive linear relations in the sense of the previous section. By Theorem 9.1, the dimension of the product is maximal possible. $\square$

The Potapov transform of an isotropic linear relation is a *unitary* matrix

$$S = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} \tag{10.1}$$

[21] this condition means that the image of a map is not contained in a totally geodesic subball of type $B_{1,k}$; omiting this constrain, we obtain a vast number of contractive maps of $\mathrm{B}_{p,q}$ into a small subball.

of the size $[p+(q+m)]\times[(p+m)+q]$. There are no additional constrains on a matrix S.

10.2. Formulae. Since isotropic linear relations also are contractive, we can use the formulae of the previous Section.

10.3. Action of the isotropic category on the Stiefel manifolds. As we seen above (4.1), the Stiefel manifold $\mathrm{Sti}(p,q)$ is in one-to-one correspondence with the set $\mathrm{IGr}(\mathbb{C}^{p,q}$ of all maximal isotropic subspaces in $\mathbb{C}^{p,q}$

Let $V\simeq\mathbb{C}^{p,p+m}$, $W\simeq\mathbb{C}^{q,q+m}$, $P:V\rightrightarrows W$ a form preserving linear relation. Let R range in $\mathrm{IGr}(\mathbb{C}^{p,p+m})$ Then $R\mapsto PR$ is a well-defined transformation $\mathrm{IGr}(p,p+m)\to\mathrm{IGr}(q,q+m)$.

The corresponding Krein–Shmul'yan map $\mathrm{Sti}(q,q+m)\to\mathrm{Sti}(p,p+m)$ is given by the usual formula

$$z\mapsto\alpha+\beta(1-z\delta)^{-1}z\gamma$$

(this map has some points of discontinuity).

In particular, for $m=0$, we obtain a family of natural maps $\mathrm{U}(q)\to\mathrm{U}(p)$.

10.4. The Livshits map. Let $m=0$, let us write $q=p+j$. Consider $[p+(p+j)]\times[p+(p+j)]$ matrix S given by

$$S:=\begin{pmatrix}0&1_p&0\\1_p&0&0\\0&0&-1_j\end{pmatrix}\tag{10.2}$$

The corresponding Krein–Shmul'yan map is

$$\begin{pmatrix}a&b\\c&d\end{pmatrix}\mapsto 0_p+\begin{pmatrix}1_p&0\end{pmatrix}\begin{pmatrix}a&b\\c&d\end{pmatrix}\left\{\begin{pmatrix}1_p&0\\0&1_j\end{pmatrix}-\begin{pmatrix}0_p&0\\0&-1_j\end{pmatrix}\begin{pmatrix}a&b\\c&d\end{pmatrix}\right\}^{-1}\begin{pmatrix}1_p\\0\end{pmatrix}=$$
$$=\begin{pmatrix}1_p&0\end{pmatrix}\begin{pmatrix}a&b\\c&d\end{pmatrix}\begin{pmatrix}1_p&0\\c&1_j+d\end{pmatrix}^{-1}\begin{pmatrix}1_p\\0\end{pmatrix}=$$
$$=\begin{pmatrix}1_p&0\end{pmatrix}\begin{pmatrix}a&b\\c&d\end{pmatrix}\begin{pmatrix}1_p&0\\-(1_j+d)^{-1}c&(1_j+d)^{-1}\end{pmatrix}^{-1}\begin{pmatrix}1_p\\0\end{pmatrix}=$$
$$=\begin{pmatrix}1_p&0\end{pmatrix}\begin{pmatrix}a-b(1_j+d)^{-1}c&*\\ *&*\end{pmatrix}\begin{pmatrix}1_p\\0\end{pmatrix}=a-b(1+d)^{-1}c$$

Thus, we define the *Livshits map* $\Upsilon^j:\mathrm{U}(p+j)\to\mathrm{U}(p)$ given by

$$\Upsilon^j\begin{pmatrix}a&b\\c&d\end{pmatrix}=a-b(1+d)^{-1}c\tag{10.3}$$

Our considerations imply the following statement.

Theorem 10.2 Υ^j *is a map* $\mathrm{U}(p+j)\to\mathrm{U}(p)$.

Another proof. We write the equation

$$\begin{pmatrix}\xi\\ -x\end{pmatrix}=\begin{pmatrix}a&b\\ c&d\end{pmatrix}\begin{pmatrix}\eta\\ x\end{pmatrix}$$

or equivalently,

$$\xi=a\eta+bx,\qquad -x=c\eta+dx$$

Excluding x, we get $x=-(1+d)^{-1}c\eta$ or

$$\xi=(a-b(1+d)^{-1}c)\eta \tag{10.4}$$

On the other hand, the matrix $\begin{pmatrix}a&b\\ c&d\end{pmatrix}$ is unitary, therefore

$$\|\xi\|^2+\|-x\|^2=\|\eta\|^2+\|x\|^2 \qquad\Longrightarrow\qquad \|\xi\|^2=\|\eta\|^2$$

Thus the operator (10.4) is unitary. □

Proposition 10.3 *For any h_1, $h_2\in\mathrm{U}(p)$,*

$$\Upsilon^j\left[\begin{pmatrix}h_1&0\\ 0&1\end{pmatrix}\begin{pmatrix}a&b\\ c&d\end{pmatrix}\begin{pmatrix}h_2&0\\ 0&1\end{pmatrix}\right]=h_1\left\{\Upsilon^j\begin{pmatrix}a&b\\ c&d\end{pmatrix}\right\}h_2 \tag{10.5}$$

This statement is obvious. □

This shows that the Livshits map is distinguished among relatively numerous[22] Krein–Shmul'yan maps $\mathrm{U}(p+j)\to\mathrm{U}(p)$.

Proposition 10.4 *The composition of the maps $\Upsilon^i:\mathrm{U}(n)\to\mathrm{U}(n-i)$, $\Upsilon^j:\mathrm{U}(n-j)\to\Upsilon(n-i-j)$ is Υ^{i+j}.*

This can be proved by straightforward multiplication of linear relations or by calculation of 'strange product' for two matrices of the type (10.2).

10.5. The Livshits map and the Cayley transform. Recall that the *Cayley transform*

$$\mathrm{Cay}_n(g):=(1-g)/(1+g)$$

is a bijection of the space $\mathrm{U}(n)$ onto the space $\mathrm{AHerm}(n)$ of anti-Hermitian matrices.

Define the 'cut-off map' $\Xi:\mathrm{AHerm}(p+j)\to\mathrm{AHerm}(p)$ given by

$$\Xi^j\begin{pmatrix}r_{11}&r_{12}\\ r_{21}&r_{22}\end{pmatrix}=r_{11}$$

Theorem 10.5 *For $g\in\mathrm{U}(p+j)$, the identity*

$$\mathrm{Cay}_p\circ\Upsilon^j(g)=\Xi^j\circ\mathrm{Cay}_{p+j}(g)$$

holds. Equivalently

$$\Upsilon^j(g)=\mathrm{Cay}_p^{-1}\circ\Xi^j\circ\mathrm{Cay}_{p+j}$$

[22] they are enumerated by pints of $\mathrm{U}(2p+j)$.

PROOF-I. Obviously, for any vectors v, u

$$u - v = g(u+v) \quad \Longleftrightarrow \quad u = (1-g)^{-1}(1+g)v = \mathrm{Cay}(g)v \tag{10.6}$$

The following row can be considered as a definition of the map Ξ:

$$u_1 = r_{11}v_1 \quad \Longleftrightarrow \quad \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \end{pmatrix} \begin{pmatrix} v_1 \\ 0 \end{pmatrix} \quad \text{for some } u_2$$

We are going to evaluate $\Xi(\mathrm{Cay}(g))$ in this way. Thus, let $u_1 = \Xi(\mathrm{Cay}(g))v_1$. Then there is u_2 such that

$$\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} - \begin{pmatrix} v_1 \\ 0 \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \left[\begin{pmatrix} v_1 \\ 0 \end{pmatrix} + \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \right]$$

or

$$\begin{aligned} (u_1 - v_1) &= a(u_1 + v_1) + bu_2 \\ u_2 &= c(u_1 + v_1) + du_2 \end{aligned}$$

The second equation implies $u_2 = -(1+d)^{-1}c(u_1+v_1)$; substitution u_2 to the first equation, we obtain

$$(u_1 - v_1) = \big(a - b(1+d)^{-1}c\big)\big(v_1 + u_1\big)$$

But this is an equation of form (10.6). Thus $\xi(\mathrm{Cay}(g))$ is the Cayley transform of $\Upsilon(g)$. □

PROOF. DIRECT FORCE ACTION. Recall the Frobenius trick for an inversion of a block matrix. First,

$$\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix} = \begin{pmatrix} 1 & \beta\delta^{-1} \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \alpha - \beta\delta^{-1}\gamma & 0 \\ 0 & \delta \end{pmatrix} \begin{pmatrix} 1 & 0 \\ \delta^{-1}\gamma & 1 \end{pmatrix}$$

this is a paraphrasing of (1.2.13). Hence

$$\begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 0 \\ -\delta^{-1}\gamma & 1 \end{pmatrix} \begin{pmatrix} (\alpha - \beta\delta^{-1}\gamma)^{-1} & 0 \\ 0 & \delta^{-1} \end{pmatrix} \begin{pmatrix} 1 & -\beta\delta^{-1} \\ 0 & 1 \end{pmatrix} = \\ = \begin{pmatrix} (\alpha - \beta\delta^{-1}\gamma)^{-1} & \dots \\ \dots & \dots \end{pmatrix}$$

(only the left upper block is interesting for us).

Further, calculate the Cayley transform

$$(1-g)/(1+g) = -1 + 2(1+g)^{-1}$$

of a block matrix $g = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$,

$$-\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + 2\left[\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} a & b \\ c & d \end{pmatrix}\right]^{-1} = \begin{pmatrix} -1 + 2(1 + a - b(1+d)^{-1}c)^{-1} & \dots \\ \dots & \dots \end{pmatrix}$$

By definition, the left upper block is $\Xi(\mathrm{Cay}(g))$. But this block is also the Cayley transform of $\Upsilon(g)$. □

Problem 10.1 The group $\mathrm{U}(n) \times \mathrm{U}(n)$ acts on itself by left and right multiplications $g \mapsto u^{-1}gv$, where $g \in \mathrm{U}(n)$ and $(u, v) \in \mathrm{U}(n) \times \mathrm{U}(n)$. Show that the corresponding transformation in the Cayley coordinates is the linear fractional transformation $r \mapsto (\alpha + r\gamma)^{-1}(\beta + r\delta)$, where

$$S := \begin{pmatrix} \alpha & \gamma \\ \beta & \delta \end{pmatrix} = \frac{1}{2} \begin{pmatrix} v+u & v-u \\ v-u & v+u \end{pmatrix} \tag{10.7}$$

10.6. Inverse limits of the spaces $\mathrm{U}(n)$.

Theorem 10.6 *The pushforward of the Haar measure μ_{p+j} on $\mathrm{U}(p+j)$ by the Livshits map is the Haar measure μ_p on $\mathrm{U}(p)$.*

Recall the definition of the direct image (push-forward) of a measure. Let M be a space with a measure μ. Let $f : M \to N$ be a map. The measure ν on the target space N is defined by the condition: for each $C \subset N$, we assume $\nu(C) := \mu(f^{-1}(C))$.

PROOF. The Haar measure on $\mathrm{U}(p+j)$ is invariant with respect to left and right shifts. In particular, it is invariant with respect to the left-right action of $\mathrm{U}(p) \times \mathrm{U}(p)$. Proposition 10.3 implies the invariance of the push-forward of the Haar measure. Thus, we obtain a $\mathrm{U}(p) \times \mathrm{U}(p)$-invariant measure on $\mathrm{U}(p)$, by definition it is the Haar measure on $\mathrm{U}(p)$. □

Normalize each Haar measure in such a way that the measure of the whole group is 1. We have the chain of the Livshits maps

$$\leftarrow \mathrm{U}(p) \leftarrow \mathrm{U}(p+1) \leftarrow \mathrm{U}(p+2) \leftarrow \dots \tag{10.8}$$

The pushforward of the Haar measure on each step of the chain is the Haar measure. By the Kolmogorov Theorem about the projective limits, there exists a space $\mathfrak{U}$ equipped with a probability measure μ_∞ and maps $\pi_n : \mathfrak{U} \to \mathrm{U}(n)$ consistent with arrows (10.8) such that the pushforward of μ_∞ under π_n is the Haar measure on[23] $\mathrm{U}(n)$.

Certainly, $\mathfrak{U}$ is not a group, however we are going to show that $\mathfrak{U}$ inherits a certain algebraic structure.

Denote by $\mathrm{U}(\infty)$ the group of all infinite unitary matrices g such that $g - 1$ has only finite number of nonzero matrix elements.

Theorem 10.7 *The group $\mathrm{U}(\infty) \times \mathrm{U}(\infty)$ acts on $\mathfrak{U}$ by the left and right multiplications. This action preserves the measure μ_∞.*

We leave this statements as an exercise. (it follows from the equivariance property (10.5).

Thus there arises a problem of harmonic analysis on the space $\mathfrak{U}$.

[23] and also $(\mathfrak{U}, \mathfrak{u})$ is a Lebesgue space and the σ-algebra of $\mathfrak{U}$ is generated by π_n-pullbacks of the σ-algebras on $\mathrm{U}(n)$.

10.7. Another point of view to the inverse limits. Again, consider the Cayley transform $\mathrm{U}(n) \to \mathrm{AHerm}(n)$. It can be shown[24], that the push-forward of the Haar measure is

$$d\mu_n := \mathrm{const}\,\det(1 - T^2)^{-n} dT \tag{10.9}$$

Theorem 10.8 (Hua Loo Keng) *The push-forward of the measure $d\mu_{p+j}$ with respect to cut-off map Ξ^j is $d\mu_p$.*

The statement is obvious since it is a paraphrasing of Theorem 10.6. But actually the Hua Theorem was obtained[25] as a result of heavy calculations which repelled specialists in semi-simple groups during several decades.

10.8. Digression. Hua-type measures. Projectivity theorem in the Hua Loo Keng's form suggests that the Haar measure can be replaced by measures $\det(1 + T)^{-n+\lambda} \det(1 - T)^{-n+\mu}\, dT$;

Theorem 10.9 *Fix a, $b \in \mathbb{C}$ such that $\mathrm{Re}(a + b) > -1$. Consider the complex-valued measures (charges)*

$$d\mu_n^{a,b} = \prod_{k=1}^{n} \frac{\Gamma(k+\lambda)\Gamma(k+\mu)}{\Gamma(k)\Gamma(k+\lambda+\mu)} \cdot \det(1+g)^{\lambda} \det(1+\overline{g})^{\mu} dg \tag{10.10}$$

on $\mathrm{U}(n)$; if $b = \overline{a}$, then $\mu_n^{a,b}$ is a positive measure.

a) The total charge of $\mathrm{U}(n)$ is 1.

b) The system is consistent with the Livshits maps (i.e., the pushforward of $\mu_n^{a,b}$ under Υ^1 is $\mu_{n-1}^{a,b}$).

Lemma 10.10 *Denote by D the disk $|z| < 1$. Consider the map $\mathrm{U}(n) \to \mathrm{U}(n-1) \times D$ given by*

$$g \mapsto (\Upsilon^1(g), g_{nn}) \tag{10.11}$$

where d_{nn} in the mostright lowest matrix element of g. The pushforward of the Haar measure under this map is

$$\pi^{-1}(n-1) \cdot (1 - |z|^2)^{n-2} \mu_{n-1}(h)\, dz\overline{dz}$$

Proof of Lemma. Since Υ^1 is equivariant, the desired measure has the form

$$\varkappa(d_{nn}) \mu_{n-1}(h)\, dz\overline{dz}$$

where $\varkappa$ is a function. To find $\varkappa$, consider the pushforward of our measure under the projection $\mathrm{U}(n-1) \times D$ on D. Evidently, it is $\varkappa(z)\, dz\overline{dz}$.

Let us find $\varkappa(z)\, dz\overline{dz}$ directly as a pushforward of μ_n under the map $\mathrm{U}(n) \to D$ that takes g to g_{nn}. In another words, we must find the distribution of the matrix element g_{nn} on the unitary group.

Denote by S^{2n-1} the sphere in $\mathbb{C}^n$, by e_n the last basis element in $\mathbb{C}^n$. Consider the map $\mathrm{U}(n) \to S^{2n-1}$ given by

$$g \mapsto ge_n \qquad \text{or} \qquad g \mapsto (g_{1n}, \dots, g_{nn})$$

[24] We omit a formal proof. The tools of the next section and Problem 10.1 allow to verify the invariance of the expression (10.9).

[25] It was not formulated in this form, but in fact it was used in calculations.

Evidently, the pushforward of the Haar measure is the uniform measure on the sphere. Now we must find a distribution of the last coordinate on the sphere. We leave this problem as an exercise. $\square$

PROOF OF THE THEOREM. We represent $g = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ and write $\det(1+g)$ as the determinant of a block matrix, see (1.11),

$$\det(1+g) = (1+d)\det(1+a-b(1+d)^{-1}c) = (1+d)\det(1+\Upsilon^1(g))$$

Thus the pushforward of $\mu_n^{a,b}$ under the map (10.11) is

$$\pi^{-1}(n-1)(1+d)^\lambda(1+\overline{d})^\mu(1-|d|^2)^{n-2}\cdot\mu_{n-1}^{a,b}$$

Considering the push-forward of this measure to $\mathrm{U}(n-1)$ we get

$$C(a,b,n)\cdot\mu_{n-1}^{a,b}$$

It remains to evaluate the constant

$$C = \frac{n-1}{\pi}\int_{|z|<1}(1-|z|^2)^{n-2}(1+z)^a(1+\overline{z})^b\,dz\,\overline{dz}$$

Substituting $z = -1+2(1+u)^{-1}$ we come to the integral

$$\int_{\mathrm{Re}\,u>0}(\mathrm{Re}\,u)^{n-2}(1+u)^{-a-2}(1+\overline{u})^{-b-2}du\overline{du} \qquad (10.12)$$

It can be evaluated in a straightforward way, we write $u = x+iy$, and integrate dy; this is the Cauchy beta-integral, see [3], after the integration we get the usual beta-integral in x. We omit details[26]. $\square$

These measures also form a projective system (for fixed λ, μ) and it is more natural to consider the infinite-dimensional harmonic analysis in this generality.

10.9. Digression. Inverse limits of the symmetric groups. Denote by S_n the usual symmetric group. Define the map Ъ $: S_n \to S_{n-1}$ according to the following rule. Consider a substitution $\begin{pmatrix} 1 & 2 & \dots & n \\ i_1 & i_2 & \dots & i_n \end{pmatrix}$ and remove both entries of n in the first and the second row. For instance, let $n=6$,

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 3 & 1 & 5 & 4 & 6 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & \\ 2 & 3 & 1 & 5 & 4 & \end{pmatrix}$$

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 2 & 3 & 1 & 6 & 4 & 5 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & \\ 2 & 3 & 1 & & 4 & 5 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 2 & 3 & & 5 & \boxed{\begin{matrix} 4 \\ 5 \end{matrix}} \\ 2 & 3 & 1 & & 4 & \end{pmatrix}$$

In the first case, we immediately obtain an element of S_5. In the second case we do not know the image of '4' and the preimage of '5'. Thus, assume $4 \to 5$.

The Haar measure on S_n is defined by the condition: the measure of each point is $1/n!$.

[26] On another hand, (10.12) looks as an L^2-inner product of two functions with respect to the measure $(\mathrm{Re}\,z)^{n-2}dz\,\overline{dz}$. We can immediately write the final result applying the Plancherel formula for a weighted Laplace transform, see below Theorem 7.4.17.

Problem 10.2 a) Show that the map Ъ : $S_n \to S_{n-1}$ commutes with the left-right action of $S_{n-1} \times S_{n-1}$.

b) The Ъ-image of the Haar measure is the Haar measure.

c) (the analog of Hua measures) Fix $t > 0$, put $\varkappa(t) := t(t+1)\dots(t+n-1)$. Denote by $\{\sigma\}$ the number of cycles in $\sigma \in S_n$. Assume that the measure of a point σ is $\varkappa(t)^{-1} t^{\{\sigma\}}$. Show that the total measure of S_n is 1; moreover this system of measures is Ъ-projective.

d)* Invent a reasonable description for the inverse limit of the spaces S_n.

e)** Invent a reasonable description for the inverse limit of the spaces $\mathrm{U}(n)$ (the author does not know a solution).

2.11 Kähler structure on a matrix ball. Some matrix tricks

In next Chapter, we discuss different approaches to central extensions of the groups $\mathrm{Sp}(2n, \mathbb{R})$ and $\mathrm{U}(p, q)$. One of the approaches (see 8.3) require the Kähler structure on the matrix ball.

11.1. The differential of a linear fractional map. Let z range in $\mathrm{Mat}(p, q)$.

Lemma 11.1 *a) Consider a map* $z \mapsto z^{[g]} := (a+zc)^{-1}(b+zd)$, *where* $g := \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{GL}(p+q, \mathbb{C})$. *Its differential is*

$$\Delta \mapsto (a+zc)^{-1}\Delta(-cz^{[g]}+d) \tag{11.1}$$

b) If $g \in \mathrm{U}(p,q)$, *then the differential also equals*

$$\Delta \mapsto (a+zc)^{-1}\Delta(b^*z+d^*)^{-1} \tag{11.2}$$

PROOF. a) Let us calculate $(z+\varepsilon\Delta)^{[g]} \mod (\varepsilon^2)$,

$$\bigl(a+(z+\varepsilon\Delta)c\bigr)^{-1}\bigl(b+(z+\varepsilon\Delta)d\bigr) = \bigl((a+zc)+\varepsilon z\Delta\bigr)^{-1}\bigl((b+zd)+\varepsilon\Delta d\bigr)$$

We transform the first factor,

$$\dots = (a+zc)^{-1}\bigl(1+\varepsilon\Delta\, c(a+zc)^{-1}\bigr)^{-1} = (a+zc)^{-1}\bigl(1-\varepsilon\Delta\, c(a+zc)^{-1}\bigr)+O(\varepsilon^2)$$

and obtain

$$(a+zc)^{-1}(b+zd) - \varepsilon(a+zc)^{-1}\Delta\, c(a+zc)^{-1}(b+zd) + \varepsilon(a+zc)^{-1}\Delta\, d + O(\varepsilon^2) =$$
$$= z^{[g]} + \varepsilon(a+zc)^{-1}\Delta\bigl(-c(a+zc)^{-1}(b+zd)+d\bigr) + O(\varepsilon^2)$$

b) We transform the term $z^{[g]}$ in formula (11.1) by (3.5) and obtain

$$\begin{aligned}-cz^{[g]}+d=-c(a+zc)^{-1}(b+zd)+d=-c(a^*z+c^*)(b^*z+d^*)^{-1}+d=\\ =\Bigl[-c(a^*z+c^*)+d(b^*z+d^*)\Bigr](b^*z+d^*)^{-1}=\\ =\Bigl[(-cc^*+dd^*)+(-ca^*+db^*)z\Bigr](b^*z+d^*)^{-1}\end{aligned}$$

By defining relations (2.4) for $\mathrm{U}(p,q)$, the expression in the square brackets is 1. □

11.2. The Jacobian of a linear fractional map.

Proposition 11.2 *The complex Jacobian of a transformation $z \mapsto z^{[g]}$, where $g \in \mathrm{GL}(p+q,\mathbb{C})$, equals $\det(a+zc)^{-p-q}\det(g)^q$.*

Lemma 11.3 *Consider a linear map $\mathrm{Mat}(p,q) \to \mathrm{Mat}(p,q)$ given by*

$$z \mapsto azb, \qquad \text{where } a \in \mathrm{GL}(p),\ b \in \mathrm{GL}(q) \tag{11.3}$$

Its determinant is $\det(a)^q \det(b)^p$.

Proof of the lemma. It sufficies to consider the map $z \mapsto az$. We represent $\mathrm{Mat}(p,q) = \mathbb{C}^p \oplus \dots \mathbb{C}^p$ (q times); our operator acts as $a \oplus a \dots \oplus a$. □

Proof of Proposition. The differential is of form (11.3). It remains to evaluate $\det(-cz^{[g]}+d)$. By formula (1.1.11) for the determinant of a block matrix,

$$\begin{aligned}\det(a+zc)\det\bigl(-c(a+xz)^{-1}(b+zd)+d\bigr) = \det\begin{pmatrix}a+zc & b+zd\\ c & d\end{pmatrix} =\\ = \det\begin{pmatrix}1 & z\\ 0 & 1\end{pmatrix}\det\begin{pmatrix}a & b\\ c & d\end{pmatrix} = \det\begin{pmatrix}a & b\\ c & d\end{pmatrix}\end{aligned}$$ □

11.3. One useful formula.

Lemma 11.4 *For $g = \begin{pmatrix}a & b\\ c & d\end{pmatrix} \in \mathrm{U}(p,q)$,*

$$1 - z^{[g]}(u^{[g]})^* = (a+zc)^{-1}(1-zu^*)(a+uc)^{*-1} \tag{11.4}$$

Proof. The left hand side is

$$\begin{aligned}1-(a+zc)^{-1}(b+zd)(b^*+d^*u^*)(a^*+c^*u^*)^{-1} =\\ =(a+zc)^{-1}\Bigl[(a+zc)(a^*+c^*u^*)-(b+zd)(b^*+d^*u^*)\Bigr](a^*+c^*u^*)^{-1}=\\ =(a+zc)^{-1}\Bigl[(aa^*-bb^*)+z(ca^*-db^*)+(ac^*-bd^*)u^*+z(cc^*-dd^*)u^*\Bigr](a^*+c^*u^*)^{-1}\end{aligned}$$

We apply defining equations (2.4) for $\mathrm{U}(p,q)$ and obtain $(1-zu^*)$ in the square brackets. □

Problem 11.1 Solve Problem 5.6

Problem 11.2 For $g \in \mathrm{U}(p,q)$,

$$1-(z^{[g]})^* u^{[g]} = (b^*+d^*z)^{*-1}(1-z^*u)(b^*+d^*u)^{-1} \tag{11.5}$$

11.4. The invariant Hermitian metric. Let T, S be tangent vectors to $\mathrm{B}_{p,q}$ at a point z (in other words, T, S are $p\times q$ matrices). Define the Hermitian metric[27] on $\mathrm{B}_{p,q}$ by

$$\Omega(T,S) = \mathrm{tr}(1-zz^*)^{-1}T(1-z^*z)^{-1}S^* \tag{11.6}$$

Theorem 11.5 *The Hermitian metric Ω is $\mathrm{U}(p,q)$-invariant*

Proof. The rules of transformation of all the factors are (11.2), (11.4, (11.5),

$$\begin{array}{ccl} (1-zz^*)^{-1} & \mapsto & (a+zc)^*\,(1-zz^*)^{-1}(a+zc) \\ T & \mapsto & (a+zc)^{-1}\,T\,(b^*z+d^*)^{-1} \\ (1-z^*z)^{-1} & \mapsto & (b^*z+d^*)(1-z^*z)^{-1}(b^*z+d^*)^* \\ S^* & \mapsto & (b^*z+d^*)^{*-1}S^*(a+zc)^{*-1} \end{array}$$

Hence

$$(1-zz^*)^{-1}T(1-z^*z)^{-1}S^* \;\mapsto\; (a+zc)^*\Big[(1-zz^*)^{-1}T(1-z^*z)^{-1}S^*\Big](a+zc)^{*-1}$$

and the traces of the left and right expressions coincide. □

Problem 11.3 For the coordinate expression Ξ for angles given by (5.7),

$$\Xi(z, z+\varepsilon T) = 1 + \varepsilon^2(1-zz^*)^{-1}T(1-z^*z)^{-1}T^* + O(\varepsilon^2)$$

11.5. The invariant measure on $\mathrm{B}_{p,q}$.

Theorem 11.6 *There is a unique to within a factor $\mathrm{U}(p,q)$-invariant measure*

$$d\mu(z) = \det(1-zz^*)^{-p-q}\prod_{ij} d\,\mathrm{Re}\, z_{ij} \prod_{ij} d\,\mathrm{Im}\, z_{ij} \tag{11.7}$$

on $\mathrm{B}_{p,q}$.

Proof. Due to Proposition 11.2 and Lemma 11.4 we know the rules of the transformation of $\det(1-zz^*)$ and $\prod_{ij} d\,\mathrm{Re}\, z_{ij} \prod_{ij} d\,\mathrm{Im}\, z_{ij}$ under the action of $\mathrm{U}(p,q)$. This easily implies the invariance of (11.7). Uniqueness is more-or-less obvious, see Addendum C.10. □

11.6. The Kähler potential.

[27] A Hermitian metric on a complex manifold is a Euclidean inner product on each tangent space smoothly depending on a point

Theorem 11.7 *The metric Ω is Kähler. Its Kähler potential is*

$$-\ln\det(1 - zz^*) = -\operatorname{tr}\ln(1 - zz^*)$$

We define the *logarithm of a matrix* in the usual way,

$$\ln(1 + X) = \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} X^k, \qquad \text{if } \|X\| < 1 \tag{11.8}$$

On Kähler metrics see, for instance [16], [5]. Recall, that a Hermitian metric Ω is *Kähler* if the *2-form* $\operatorname{Im}\Omega$ *is closed*. In particular, *our* $\mathrm{B}_{p,q}$ *is a symplectic manifold*. A function $\mathcal{K}$ is a *Kähler potential* of a Kähler metric Ω if

$$\Omega(dz, d\overline{z}) = \sum_{k,l} \frac{\partial^2 \mathcal{K}(z, \overline{z})}{\partial z_k \partial \overline{z}_l} \, dz_k \wedge d\overline{z}_l$$

here

$$\frac{\partial}{\partial z_k} = \frac{1}{2}\Big(\frac{\partial}{\partial x_k} - i\frac{\partial}{\partial y_k}\Big), \qquad \frac{\partial}{\partial \overline{z}_k} = \frac{1}{2}\Big(\frac{\partial}{\partial x_k} + i\frac{\partial}{\partial y_k}\Big)$$
$$dz_k = dx_k + i dy_k, \qquad d\overline{z}_k = dx_k - i dy_k$$

Recall that a Kahler potential is not canonically determined by a metric.

Proof. Direct force action. We must find the term $\varepsilon^2 TS^*$ in the Taylor expansion of $\operatorname{tr}\ln(1 - (z + \varepsilon T)(z + \varepsilon S)^*)$

$$\begin{aligned}
&\operatorname{tr}\ln\big(1 - (z + \varepsilon T)(z + \varepsilon S)^*\big) = \\
&\qquad = \operatorname{tr}\ln\Big(1 - zz^* - \varepsilon(zS^* + Tz^*) - \varepsilon^2 TS^*\Big) = \\
&= \operatorname{tr}\Big[\ln(1 - zz^*) + \ln\Big(1 - \varepsilon(1 - zz^*)^{-1}(Tz^* + zS^*) - \varepsilon^2(1 - zz^*)^{-1}(TS^*)\Big)\Big] = \\
&\qquad = \operatorname{tr}\ln(1 - zz^*) + \operatorname{tr}\Big[-\varepsilon(1 - zz^*)^{-1}(Tz^* + zS^*) - \\
&-\frac{1}{2}\varepsilon^2(1-zz^*)^{-1}(Tz^*+zS^*)(1-zz^*)^{-1}(Tz^*+zS^*) - \varepsilon^2(1-zz^*)^{-1}TS^*\Big] + O(\varepsilon^3)
\end{aligned}$$

We must only watch the terms with both T, S^*:

$$\begin{aligned}
\varepsilon^2 \operatorname{tr}\Big[&-\frac{1}{2}(1 - zz^*)^{-1}Tz^*(1 - zz^*)^{-1}zS^* - \\
&- \frac{1}{2}(1 - zz^*)^{-1}zS^*(1 - zz^*)^{-1}Tz^* - (1 - zz^*)^{-1}TS^*\Big]
\end{aligned}$$

Traces of the first two summands coincide (since $\operatorname{tr} AB = \operatorname{tr} BA$) and we get

$$\begin{aligned}
&-\varepsilon^2 \operatorname{tr}(1 - zz^*)^{-1}\Big[Tz^*(1 - zz^*)^{-1}zS^* + TS^*\Big] = \\
&= -\varepsilon^2 \operatorname{tr}(1-zz^*)^{-1}T\Big[z^*(1-zz^*)^{-1}z+1\Big]S^* = -\varepsilon^2 \operatorname{tr}(1-zz^*)^{-1}T(1-z^*z)^{-1}S^*
\end{aligned}$$

and we obtain a desired expression. $\square$

Another proof. We are going to watch a behavior of $\operatorname{tr}\ln(1-zz^*)$ under the transformation $z\mapsto z^{[g]}$. By (11.4),

$$\operatorname{tr}\ln(1-z^{[g]}(z^{[g]})^*)= \tag{11.9}$$
$$=\operatorname{tr}\ln\Bigl((a+zc)^{-1}(1-zz^*)(a+zc)^{*-1}\Bigr)= \tag{11.10}$$
$$=\operatorname{tr}\ln\Bigl(a^{-1}(1+zca^{-1})^{-1}(1-zz^*)(1+zca^{-1})^{*-1}a^{*-1}\Bigr)= \tag{11.11}$$
$$=\operatorname{tr}\ln\Bigl(a^{*-1}a^{-1}(1+zca^{-1})^{-1}(1-zz^*)(1+zca^{-1})^{*-1}\Bigr)= \tag{11.12}$$
$$=\operatorname{tr}\ln(1-zz^*)-\operatorname{tr}\ln(aa^*)-\operatorname{tr}\ln(1+zca^{-1})-\operatorname{tr}\ln(1+zca^{-1})^* \tag{11.13}$$

Explain what happened. We intend to apply the formula

$$\operatorname{tr}\ln(XY)=\operatorname{tr}\ln X+\operatorname{tr}\ln Y \tag{11.14}$$

This formula is dangerous, since the logarithm is a many-valued function. The initial expression (11.9) is well defined. Nevertheless, we can not immediately transform (11.10) since we do know what is $\ln(a+zc)$. The passage from (11.11) to (11.12) is an application of $\operatorname{tr}\ln XY=\operatorname{tr}\ln YX$. At the last moment, we apply (11.14), because all the summands in (11.13) are well defined.

Thus, $\operatorname{tr}\ln(1-z^{[g]}(z^{[g]})^*)$ is

$$=\operatorname{tr}\ln(1-zz^*)+\Bigl\{\text{constant}\Bigr\}+\left\{\begin{matrix}\text{holomorphic}\\ \text{summand}\end{matrix}\right\}+\left\{\begin{matrix}\text{antiholomorphic}\\ \text{summand}\end{matrix}\right\}$$

After differentiating $\partial\overline{\partial}$ the three additional terms vanish. Thus

— *Kähler metrics defined by our potential is* $\mathrm{U}(p,q)$*-invariant.*

The metric Ω also is invariant. Therefore it suffices to compare our metrics at the point 0. $\square$

11.7. Exercises on matrix differentials and Jacobians. 1)

Problem 11.4 For the spaces $\mathrm{Symm}(n)$ and $\mathrm{ASymm}(n)$ of real symmetric and skew-symmetric matrices, consider the maps $z\mapsto a^t za$. Show that their determinants are respectively $(\det a)^{n+1}$ and $(\det a)^{n-1}$, see also Lemma 7.4.2.

Problem 11.5 b) Prove the Hua formula (10.9) for the Haar measure. One of possible ways is to verify the invariance of this expression with respect to linear fractional transformations from Problem 10.1.

Problem 11.6 a) Find the differential of the map $X\mapsto e^X$ on the space $\mathrm{Mat}(n)$ as follows. For a matrix T, define the linear maps $L=L_T$, $R=R_T:\mathrm{Mat}(n)\to\mathrm{Mat}(n)$ by

$$L_TY=TY,\qquad R_TY=YT$$

Show that the differential of the map $X\mapsto\exp(X)$ is

$$\Delta\mapsto\frac{\exp L-\exp R}{L-R}\,\Delta:=\sum_{n=1}^{\infty}\frac{1}{n!}\frac{L^n-R^n}{L-R}\,\Delta$$

b) Show that the Jacobian of the map $X \mapsto e^X$ is

$$\prod_{i\neq j} \frac{e^{\lambda_i} - e^{\lambda_j}}{\lambda_i - \lambda_j} = \prod_{i>j} \Big(\frac{\sinh(\lambda_i - \lambda_j)/2}{(\lambda_i - \lambda_j)/2} \Big)^2 \tag{11.15}$$

where λ_j are the eigenvalues of X. HINT: first, assume that X is diagonalizable; next, apply continuity arguments.

c) Extend a), b) to arbitrary maps $X \mapsto \sum_{k\geqslant 0} c_k X^k$, where c_k are scalars.

2) Let G be Lie group admiting a two-side-invariant Haar measure (see Addendum C.10), $\mathfrak{g}$ a Lie algebra. The exponential map $\mathfrak{g} \to G$ determines a local coordinate system on $\mathfrak{g}$.

Problem 11.7 Show that the density of the Haar measure in the exponential coordinates is

$$\mathrm{const} \cdot \prod_j \frac{\sinh \mu_j/2}{\mu_j/2} \tag{11.16}$$

where μ_j are the eigenvalues of $\mathrm{ad}(X)$[28]. HINT. By the Ado theorem, $\mathfrak{g} \subset \mathfrak{gl}(n)$. Let $h := \exp(P)$. Find the Jacobian of the map $X \mapsto g^{-1} \exp X$ at a point P.

2.12 Matrix balls as symmetric spaces

We occasionally refer to this section below and above, but it is not used seriously.

12.1. Riemannian metrics on matrix balls. In the previous section, we defined the invariant Hermitian metric on $\mathrm{B}_{p,q}$. Its real part

$$M(T,S) := \mathrm{Re}\,\mathrm{tr}(1 - zz^*)^{-1} \cdot T \cdot (1 - z^*z)^{-1} \cdot S^*$$

is an invariant Riemannian metric.

12.2. Symmetric spaces. Let M be a complete Riemannian manifold.[29]. Let $x \in M$. A *geodesic symmetry* θ_x is a map from a neighborhood of x to itself defined in the following way:

$$\theta_x(y) = z, \qquad \text{if } x \text{ is the midpoint of the geodesics connecting } y \text{ and } z$$

The space M is called *symmetric* if all the geodesic symmetries are isometries.

Theorem 12.1 *The space* $\mathrm{B}_{p,q}$ *is symmetric.*

PROOF. The map $z \mapsto (-z)$ is a geodesic symmetry in 0. Since the group of isometries acts transitively, all the geodesic symmetries are isometries. □

We also us present a coordinate-free description of geodesic symmetries. Namely, for a negative subspace $P \in \mathcal{M}_{p,q}$, the map θ_P is the reflection in P.

12.3. Description of geodesics.

[28] where ad is the adjoint map, see Addendum C.9.

[29] The definition given below can be extended pseudo-Riemannian spaces. For a representative collection of examples of pseudo-Riemannian symmetric spaces, see Problem 3.3.6

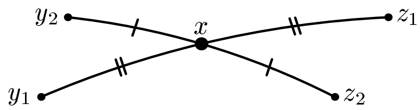

Fig. 2.18: A geodesic symmetry

Theorem 12.2 *a) The submanifold in* $\Xi \subset \mathrm{B}_{p,q}$ *consisting of matrices*

$$\begin{pmatrix} \lambda_1 & 0 & \dots \\ 0 & \lambda_2 & \dots \\ \vdots & \vdots & \ddots \end{pmatrix}, \qquad \text{where } \lambda_j \text{ are real, } |\lambda_j| < 1 \tag{12.1}$$

is totally geodesic and the Riemannian metric is flat on Ξ. *Moreover, the coordinates* $x_j = \operatorname{arctanh} \lambda_j$ *on (12.1) are flat*

b) Given $\sigma_j \in \mathbb{R}$*, the curve*

$$\gamma(t) := \begin{pmatrix} \tanh(\sigma_1 t) & 0 & \dots \\ 0 & \tanh(\sigma_2 t) & \dots \\ \vdots & \vdots & \ddots \end{pmatrix} \tag{12.2}$$

is a geodesic

c) Each geodesics can be reduced to this form by an element of $\mathrm{U}(p,q)$

d) The geodesic distance in $\mathrm{B}_{p,q}$ *is given by the formula*

$$\rho(L_1, L_2) = \mathrm{const} \cdot \sqrt{\sum\nolimits_j \psi_j(L_1, L_2)^2} \tag{12.3}$$

where ψ_j *are the hyperbolic angles.*

Lemma 12.3 *Any connected component of the set of fix points of any isometry is a totally geodesic subspace, i.e., a short geodesic joining two close points of the set is contained in the set.*

Proof of the lemma. Denote the isometry by τ, by $\mathrm{Fix}(\tau)$ the set of its fixed points. Given close point x, $y \in \mathrm{Fix}(\tau)$, denote by γ the short geodesics connecting these points. Then $\tau\gamma$ also is a geodesic having the same length. Since a short geodesic is unique, it follows that $\tau\gamma = \gamma$, and $\gamma \in \mathrm{Fix}(\tau)$. $\square$

Proof of Theorem. a) Consider an isometry $\mathrm{B}_{p,q} \to \mathrm{B}_{p,q}$ having the form

$$z \mapsto \Lambda z \begin{pmatrix} \Lambda^{-1} & 0 \\ 0 & \Theta \end{pmatrix}$$

where $\Lambda \in \mathrm{U}(p)$, $\Theta \in \mathrm{U}(q-p)$ are diagonal matrices with pairwise distinct entries $\neq 1$. The set of its fixed points consists of matrices having form (12.1)

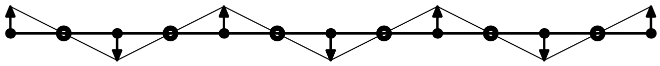

Fig. 2.19: A sequence of geodesic symmetries.

with $\lambda_j \in \mathbb{C}$. Therefore, the set of matrices (12.1) with complex entries[30] is totally geodesic.

Next, we consider the isometry $z \mapsto \overline{z}$ and apply the lemma again. Thus, Ξ is totally geodesic.

The Riemannian metric on Ξ is

$$\sum_{j=1}^{p} \frac{d\lambda_j^2}{(1-\lambda_j^2)^2} = \sum_{j=1}^{p} \bigl(d \operatorname{arctanh} \lambda_j\bigr)^2$$

Therefore the coordinates $x_j = \operatorname{arctanh} \lambda_j$ are flat.

b) This curve is a stright line in the flat coordinates.

c) We move a geodesic by an element of $\mathrm{U}(p, q)$ and put one of its points to 0. Next, we move the geodesic by an element of the group $\mathrm{U}(p) \times \mathrm{U}(q)$ (this group stabilizes 0). In this way, we can diagonalize a given point of the geodesics. Now two points of the geodesics are contained in the submanifold Ξ and we know the geodesic connecting these points.

d) We merely evaluate the length of a segment of a geodesic. $\square$

12.4. Parallel transport in $\mathrm{B}_{p,q}$. Let $\gamma : [a, b] \to M$ be a smooth curve in Riemannian manifold. Recall that it determines a canonical map (parallel transport) of tangent spaces $T_{\gamma(a)} \to T_{\gamma(b)}$.

Theorem 12.4 *For $M = \mathrm{B}_{p,q}$, one can find an isometry $g \in \mathrm{U}(p, q)$ taking $\gamma(a)$ to $\gamma(b)$ and inducing the parallel transport map $T_{\gamma(a)} \to T_{\gamma(b)}$.*

To prove this theorem, we reformulate a definition of parallel transport. Denote by t the natural parameter on γ, put $a = 0$. Given N, denote $x(k) := \gamma\bigl(a + \frac{2k+1}{2N+1}(b - a)\bigr)$ (i.e., we divide the curve into $2N$ segments having equal lengths). Next, consider the product

$$\theta_{x(2N)} \circ \theta_{x(2N-1)} \circ \cdots \circ \theta_{x(2)} \circ \theta_{x(1)}$$

and the induced map $T_{\gamma(a)} \to T_{\gamma(b)}$. Now tend $N \to \infty$. The limit is the parallel transport map.

After this reformulation, the theorem becomes obvious.

12.5. Digression. What is the geometry of symmetric spaces? Thus, the standard objects of differential geometry as geodesics, parallel transport, Riemann curvature tensor etc. admit a rather transparent description for $\mathrm{B}_{p,q}$ (and more generally, for all Riemannian symmetric spaces).

[30] It is a product of p Lobachevsky planes

But the main subjects of geometry of symmetric spaces are far beyond such questions. Some of important classical topics are

– the topology of compact symmetric spaces (as Grassmannians, the group $\mathrm{U}(n)$, the space $\mathrm{U}(n)/\mathrm{O}(n)$ discussed in the next Chapter, etc.)

– geometry of non-compact spaces at infinity

– discrete groups acting on symmetric spaces.

In spite of this, it seems that the geometry of symmetric spaces is not well-understood. It seems that the main geometric structure is angles. In this context, we have no geometrical intuition. Only recently it was obtained the triangle inequality. Recall this problem.

Let L_1, L_2, $L_3 \in \mathcal{M}_{p,q}$. Then we have three collections of pairwise angles, say μ_j, ν_j, and $\varkappa_j$.

We return to this problem below in ?????????????????

12.6. Digression. Problems of harmonic analysis on $\mathrm{B}_{p,q}$. For definiteness, assume $p \leqslant q$.

The first problem of harmonic analysis on $\mathrm{B}_{p,q}$ is the decomposition of the space $L^2(\mathrm{B}_{p,q}, d\mu)$ into an integral of irreducible representations, see, for instance [68].

There are two another types of inner products which lead to a reacher theory.

A) *Berezin kernels* on $\mathrm{B}_{p,q}$ are defined by

$$K_\alpha(z,u) := \left(\frac{\det(1-zz^*)\det(1-uu^*)}{|\det(1-zu^*)|^2}\right)^\alpha,$$

Let δ_z denote the delta-function supported by $z \in \mathrm{B}_{p,q}$. We introduce the inner products of finite linear combinations delta-functions by

$$\langle \delta_z, \delta_u \rangle_\alpha = K_\alpha(z,u)$$

Theorem 12.5 *This inner product is nonnegative definite iff $\alpha = 0, 1, \dots, p-1$ or $\alpha > p-1$.*

This statement is a rephrasing of the Berezin–Gindikin–Vergne–Rossi Theorem, see Section 7.4.

Therefore our inner product defines a structure of a pre-Hilbert space and we consider the associated Hilbert space $\mathcal{H}_\alpha$ (for a detailed discussion of constructions of this type see Section 7.1).

It turns out that the harmonic analysis in $\mathcal{H}_\alpha$ is a quite reach subject; the classical L^2-analysis is its limit as $\alpha \to \infty$.

B) *Stein–Sahi kernels.* Define the inner product on the space of smooth compactly supported functions on $\mathrm{B}_{p,q}$ by the singular integral

$$\langle f, g\rangle_\alpha = \iint\limits_{\mathrm{B}_{p,q}\times\mathrm{B}_{p,q}} \det{}^{-\alpha}\Bigl((1-zz^*)^{-1}(z-u)\det(1-u^*u)^{-1}(z^*-u^*)\Bigr)\times$$
$$\times f(z)\overline{g(u)}\,d\mu(z)\,d\mu(u)$$

this integral converges for $\alpha < 1$, further we consider its meromorphic continuation in the variable α.

Conditions of positivity of this inner product are not completely investigated. It is known that for $p = q$ the inner product is positive for $|\alpha - p| < 1/2$; for $\alpha = p$, we obtain the L^2-inner product. For certain $p \neq q$ there are some small islands of positivity.

C) Recall that one has several ways to write the Hua double ratio. Both kernels (Berein and Stein–Sahi) are powers of determinant of one of variants of the double ratio. It is interesting, that there is a third variant,

$$\det{}^{-\alpha}\Big((1-u^*z)^{-1}(u^*-z^*)(1-zu^*)^{-1}(u-z)\Big)$$

As far as I know, it is senseless.

12.7. Other Riemannian symmetric spaces.

Theorem 12.6 *Submanifold in* $\mathrm{B}_{n,n}$ *consisting of symmetric, matrices is totally geodesic.*

PROOF. The map $z \mapsto z^t$ is an isometry (see Lemma 12.3). □

This manifold is an important hero of the next chapter, namely it is $\mathrm{Sp}(2n,\mathbb{R})/\mathrm{U}(n)$. We can also consider skew-symmetric matrices ($\mathrm{SO}^*(2n)\mathrm{U}(n)$), Hermitian matrices ($\mathrm{GL}(n,\mathbb{C})/\mathrm{U}(n)$), real symmetric and skew-symmetric matrices ($\mathrm{GL}(n,\mathbb{R})/\mathrm{O}(n)$ and $\mathrm{O}(n,\mathbb{C})/\mathrm{O}(n)$), real $p\times q$-matrices in $\mathrm{B}_{p,q}$ ($\mathrm{O}(p,q)/\mathrm{O}(p)\times\mathrm{O}(q)$). We can also consider the quaternions as complex 2×2-matrices, this allows to realize the space $\mathrm{Sp}(p,q)/\mathrm{Sp}(p)\times\mathrm{Sp}(q)$ as a totally geodesic subspace in $\mathrm{B}_{2p,2q}$, and also $\mathrm{GL}(n,\mathbb{H})/\mathrm{Sp}(n)$, $\mathrm{Sp}(2n,\mathbb{C})/\mathrm{Sp}(n)$ as totally geodesic subspaces in $\mathrm{B}_{2n,2n}$ consisting of quaternion Hermitian and anti-Hermitian matrices. Modulo minor details, this is a complete list of classical noncompact Riemannian symmetric spaces.

In many cases, this remark allows to transfer information from one series of Riemannian symmetric spaces to another. For instance, we immediately get explicit realizations, formulae for geodesics and geodesic distance, expressions for angles, compresivity theorem etc.