

# SIMULTANEOUSLY DISSIPATIVE OPERATORS AND THE INFINITESIMAL WRAPPING EFFECT IN INTERVAL SPACES

A. N. GORBAN

*Krasnoyarsk Computing Center SB RAS, Russia*

YU. I. SHOKIN

*Institute of Computational Technologies SB RAS, Novosibirsk, Russia*

e-mail: shokin@adm.ict.nsc.ru

V. I. VERBITSKII

*Krasnoyarsk Computing Center SB RAS, Russia*

Работа посвящена приложениям теории совместно диссипативных операторов к интервальному анализу и химической кинетике. Главным объектом исследования является нежелательный “эффект упаковывания”, широко проявляющийся при численном решении на ЭВМ эволюционных дифференциальных уравнений с интервальными параметрами. Основной результат работы — доказательство типичности эффекта упаковывания в малом, что объясняет низкую эффективность традиционных пошаговых методов численного решения интервальных дифференциальных задач.

## Introduction

For solving systems of ordinary differential equations different classes of numerical methods with guaranteed error estimation including interval methods are used. In solving a system by an interval method the approximate solution at any considered moment of time  $t$  represents a set (called interval) containing the exact solution at the moment  $t$ . The detailed account of interval methods can be found in monographs by R. Moore [1] and S. A. Kalmykov, Yu. I. Shokin, Z. Kh. Yuldashev [2].

As a rule, all kinds of rectangular parallelepipeds with sides parallel to coordinate axes [1, 2] are used as intervals, less frequently — ellipsoids [3], balls of fixed norm [4, 5] etc.

One of shortcomings of stepwise interval methods is the following. The intervals determining the solution of a system are often expanded in the course of time irrespective of the method and step used. The simplest example of strong expansion of intervals during a short time, belonging to R. Moore, is given in [1]. The phenomenon of interval expansion, called the wrapping effect, essentially decreases the efficiency of interval methods. In the present work the notions of the interval and the wrapping effect are formalized and the wrapping effect is studied for autonomous systems on positively invariant convex compact.

Formally, one can get rid of the interval expansion for any globally stable system (i. e. such a system any solution of which is stable after Lyapunov). To demonstrate that, let us consider a smooth autonomous system:

$$\frac{dx}{dt} = f(x) \quad (1)$$

on the positively invariant compact  $B \subset \mathbb{R}^n$ . We construct a metric  $\rho$  on the set  $B$ , assuming for any  $x \in B, y \in B$ :

$$\rho(x, y) = \sup_{t \geq 0} \|x(t) - y(t)\|,$$

where  $x(t), y(t)$  are solutions of the system (1) with the initial conditions  $x(0) = x, y(0) = y$ . This metric is constricting for (1), i. e., for any pair  $x(t), y(t)$  of the solutions of (1) with the initial conditions in  $B$

$$\rho(x(t), y(t)) \leq \rho(x(s), y(s)) \quad \text{at } t \geq s.$$

The metric  $\rho$  is topologically equivalent to norm if and only if the system (1) is globally stable in  $B$ . If one considers as intervals all balls of the metric  $\rho$ , then in a definite sense the wrapping effect is absent. That is, there is no interval expansion when constructing the exact interval solution with any step  $h > 0$ . The exact interval solution of  $X(t)$  is defined in the following way:  $X(0) = X_0$ , where  $X_0$  is the initial interval with the center at the point  $x(0) = x_0$ ;  $X((n+1)h)$  is the minimal interval with the center at the point  $x((n+1)h)$  containing  $T_h X(nh)$ , where  $T_t$  is the transformation of the phase flow of (1) during the time  $t \geq 0$ . Indeed, the radius  $X((n+1)h)$  does not exceed the radius  $X(nh)$  at any  $n$ .

If the system is not globally stable, then metric is not topologically equivalent to the norm. It means that small, in usual sense, intervals became large in the metric  $\rho$ . This circumstance makes one refuse from consideration of similar metrics. Moreover, if the system (1) is absolutely unstable (for example, a system with mixing), then there is no reasonable way to get rid of the wrapping effect.

Unfortunately, the method that we describe to eliminate the wrapping effect for globally stable system is non-constructive. That can be demonstrated as follows: for constructing the constricting metric  $\rho$  one must know all exact solutions of the system (1). But then it is unreasonable to solve the system numerically. We must have constructively verifiable conditions of absence of the wrapping effect and a way of construction of corresponding intervals. This is what we deal with in the present paper. The conditions of absence of the wrapping effect are of local character and formulated in terms of Jacobi matrices of the system. Except that the causes of frequent appearance of the wrapping effect will be pointed out.

The authors are grateful to Sergey Shary for fruitful discussions and valuable comments.

## 1. Interval spaces and the wrapping effect

### 1.1. Interval Spaces

Before starting to study the wrapping effect, it is necessary to define what we mean by intervals. Generalizing some well-known constructions, we give the following definition.

**Definition 1.** We call the family  $\mathbb{J}$  of convex compacts in  $\mathbb{R}^n$  *the interval space* (and its elements – *intervals*), if it satisfies the following conditions:

- a)  $\mathbb{J}$  is closed with respect to multiplication by non-negative scalars:

$$\text{if } W \in \mathbb{J}, \alpha \geq 0, \quad \text{then } \alpha W = \{\alpha x \mid x \in W\} \in \mathbb{J};$$

b)  $\mathbb{J}$  is closed with respect to intersection:

$$\text{if } W_1 \in \mathbb{J}, W_2 \in \mathbb{J}, \quad \text{then } W_1 \cap W_2 \in \mathbb{J};$$

c)  $\mathbb{J}$  is closed according to Hausdorff (i. e. in the Hausdorff metric);

d) if  $W \in \mathbb{J}$ ,  $W \neq \{0\}$ , then  $0 \in \text{ri}W$ .

Remind [6] that the Hausdorff metric on the set of all compacts in  $\mathbb{R}^n$  is introduced as follows:

$$\rho_H(x, y) = \max\left\{\max_{x \in X} \min_{y \in Y} \|x - y\|, \max_{y \in Y} \min_{x \in X} \|x - y\|\right\},$$

where  $x, y$  are the compacts in  $\mathbb{R}^n$ ,  $\|\cdot\|$  is a fixed norm in  $\mathbb{R}^n$ . All Hausdorff metrics in  $\mathbb{R}^n$  are equivalent.

Further on by  $\lim_{H, i \rightarrow \infty} W_i$  we denote the Hausdorff limit of the sequence  $\{W_i\}_{i=1}^{+\infty}$  at  $i \rightarrow \infty$ .

Below, we give several examples of interval spaces.

**Example 1.**  $\mathbb{J}$  is the set of all convex compacts symmetric with respect to 0. It satisfies all the properties from a) to d).

**Example 2.**  $\mathbb{J}$  is the set of all symmetric with respect to 0 rectangular parallelepipeds (including non-singular), i. e. sets of the form

$$\{x = (x_1, \dots, x_n) \in \mathbb{R}^n \mid |x_k| \leq a_k \ (k = 1, \dots, n)\},$$

where  $a_k \geq 0$  ( $k = 1, \dots, n$ ). It satisfies the properties a), b) and d).

Let now  $\{W_i\}_{i=1}^{\infty} \subset \mathbb{J}$ ,  $\lim_{H, i \rightarrow \infty} W_i = W$  with  $a_k^{(i)}$  being  $a_k$ , corresponding to  $W_i$ .

If  $\rho_H(W_i, W) < \varepsilon$ , then  $W \subset W_i + P_\varepsilon$ ,  $W_i \subset W + P_\varepsilon$ , where  $P_\varepsilon = \{x \in \mathbb{R}^n \mid |x_k| \leq \varepsilon \ (k = 1, \dots, n)\}$  (here a norm in the definition of the Hausdorff metric is the  $l^\infty$ -norm). Then for any  $x \in W$

$$|x_k| \leq a_k^{(i)} + \varepsilon \quad (k = 1, \dots, n)$$

is true and for any  $x \in W_i$

$$\begin{aligned} |x_k| &\leq \underline{\lim}_{i \rightarrow \infty} a_k^{(i)}; \\ \lim_{i \rightarrow \infty} a_k^{(i)} &\leq \overline{\lim}_{i \rightarrow \infty} a_k^{(i)}, \end{aligned}$$

i. e. there exist the limits

$$\tilde{a}_k = \lim_{i \rightarrow \infty} a_k^{(i)} \quad (k = 1, \dots, n).$$

If  $x \in W$ , then

$$|x_k| \leq \lim_{i \rightarrow \infty} a_k^{(i)} \quad (k = 1, \dots, n). \quad (2)$$

Let  $b_i = \min_{1 \leq k \leq n} (a_k^{(i)} / \tilde{a}_k)$ . If for some  $x \in \mathbb{R}^n$  the inequalities (2) are satisfied, then  $\lim_{i \rightarrow \infty} b_i = 1$ .

Obviously,  $x^{(i)} = b_i x \in W_i$ , i. e. there exists such a subsequence of  $\{x^{(i)}\}_{i=1}^{\infty}$  that  $x^{(i)} \in W_i$ ,  $x = \lim_{i \rightarrow \infty} x^{(i)}$ . Hence  $W = \{x \in \mathbb{R}^n \mid |x_k| \leq \tilde{a}_k \ (k = 1, \dots, n)\}$ , i. e.  $W \in \mathbb{J}$  and the property c) is also satisfied.

In constructing interval methods of solving different problems it is, as a rule, the considered interval space that is made use of [1, 2].

**Example 3.** Let  $\|\cdot\|$  be a norm in  $\mathbb{R}^n$ ,  $\mathbb{III}_r = \{x \in \mathbb{R}^n \mid \|x\| \leq r\}$ , where  $r \geq 0$ . Let

$$\mathbb{J} = \{\mathbb{III}_r \mid r \geq 0\},$$

i. e.  $\mathbb{J}$  is the set of all closed balls (further on we omit the word "closed") of the norm. All the properties from a) to d) are satisfied. These interval spaces are used, for example, in [4, 5].

**Example 4.** The construction of example 3 can be generalized as follows. Let  $\|\cdot\|_1, \dots, \|\cdot\|_m$  be the finite set of norms in  $\mathbb{R}^n$ ,

$$\mathbb{III}_{r_k}^{(k)} = \{x \in \mathbb{R}^n \mid \|x\|_k \leq r_k \quad (k = 1, \dots, m)\}$$

where  $r_k \geq 0$  ( $k = 1, \dots, m$ ) and

$$W_{r_1, \dots, r_m} = \bigcap_{1 \leq k \leq m} \mathbb{III}_{r_k}^{(k)}.$$

Let  $\mathbb{J} = \{W_{r_1, \dots, r_m} \mid r_k \geq 0 (k = 1, \dots, m)\}$ . Obviously,  $\mathbb{J}$  possesses the properties a), b), and d).

Note that the same element of  $\mathbb{J}$  can be associated with different sets of  $\{r_k\}$ . To demonstrate that, let  $m = 2$ ,  $\sup_{x \neq 0} (\|x\|_2 / \|x\|_1) = C$ . Then  $\mathbb{III}_1^{(1)} = W_{1, C'}$ , where  $C'$  is any number not less than  $C$ . Also, even if one of  $r_k$  is equal to 0, then

$$W_{r_1, \dots, r_m} = \{0\}.$$

To each compact  $W \subset \mathbb{R}^n$  can be juxtaposed the set

$$\{r_k(W)\}_{k=1}^m \mid r_k(W) = \max_{x \in W} \|x\|_k \quad (k = 1, \dots, m).$$

If  $W \in \mathbb{J}$  then  $W = W_{r_1} \cap \dots \cap W_{r_m}$ .

Let now the sequence  $\{W_i\}_{i=1}^\infty$  converge according to Hausdorff to the compact  $W$ , with  $W_i \in \mathbb{J}$  for all  $i$ . Similarly to example 2, from the inclusions

$$W_i \subset W + \mathbb{III}_\varepsilon^{(k)}, \quad W \subset W_i + \mathbb{III}_\varepsilon^{(k)}$$

satisfied for each  $\varepsilon > 0$  for all  $i > i_0(\varepsilon)$  derive the existence of the limits:

$$\tilde{r}_k = \lim_{i \rightarrow \infty} r_k(W_i) \quad (k = 1, \dots, m)$$

and conclude that

$$W = W_{\tilde{r}_1, \dots, \tilde{r}_m},$$

i. e.  $W \in \mathbb{J}$ , and the property c) is satisfied.

**Example 5.** Let  $Q$  be a compact convex body without symmetry center (for instance, a triangle in  $\mathbb{R}^2$ ),  $0 \in \text{int } Q$ . Assume

$$\mathbb{J} = \{\alpha Q \mid \alpha \geq 0\}.$$

$\mathbb{J}$  possesses the properties from a) to d).

**Remark 1.** Example 5 can be generalized. For this purpose it is necessary to consider compact convex bodies  $Q_1, \dots, Q_m$ , the interior of each of them contains 0, and to take as  $\mathbb{J}$  a family of all the sets of the form  $\bigcap_{1 \leq k \leq m} \alpha_k Q_k$  where  $\alpha_k \geq 0$  ( $k = 1, \dots, m$ ).

## 1.2. Dissipative Operators

In this section the properties of the operators dissipative with respect to compact are studied. First, let remind some notations.

The affine hull of the convex set  $W$  is denoted by  $\text{Aff}W$ , the relative interior  $W$  (the interior of  $W$  in  $\text{Aff}W$ ) is denoted by  $\text{ri}W$ , the relative boundary of  $W$  (the boundary of  $W$  in  $\text{Aff}W$ ) is denoted by  $\text{r}\partial W$ . For the boundary of the set  $X$  we use the notation  $\partial X$ ,  $\text{int}X$  — for the interior of  $X$ ,  $\text{co}X$  — for the convex hull of  $X$ . By the sum of the sets of  $X$  and  $Y$  from  $\mathbb{R}^n$  we mean the set  $\{x + y \mid x \in X, y \in Y\}$ , by  $I$  — the unit operator.

Let us introduce a new notion.

**Definition 2.** The linear operator  $A$  in the space  $\mathbb{R}^n$  is called *dissipative with respect to the family of sets*  $\{W_\nu\} \subset \mathbb{R}^n$  if every set  $W_\nu$  is positively invariant with respect to the system

$$\frac{dx}{dt} = Ax. \quad (3)$$

In other words, every  $W_\nu$  is invariant with respect to the semi-group of the operators  $\exp(At)$  ( $t \geq 0$ ).

Below we consider operators dissipative with respect to families of convex compacts. In particular, the operator is dissipative with respect to the families of all balls of some norm (for this, dissipativity with respect to only one ball is sufficient) if and only if  $\|\exp(At)\| \leq 1$  at all  $t \geq 0$ . Thus, in this case we come to the known definition of dissipativity with respect to the norm [7].

The set of all operators dissipative with respect to  $\{W_\nu\}$  is denoted by  $K(\{W_\nu\})$ .

**Remark 2.** If an operator is dissipative with respect to the family of compacts and the interior of at least one of them is not empty, then it is dissipative with respect to some norm.

Indeed, any symmetric with respect to 0 compact convex body is a ball of some norm (see, for example, [7]). Choose as a ball the following set:

$$S = \text{co}\{W \cup (-W)\} \quad (4)$$

where  $W$  is any set of the considered family of  $\{W_\nu\}$ , for which  $\text{int}W \neq \emptyset$ .

However, if  $W$  is a compact and the operator is dissipative with respect to the norm whose ball is  $S$  (4), then it does not yet mean that the operator is dissipative with respect to  $W$  (see also example 8).

**Remark 3.** From the invariance of a family of compacts with respect to the linear operator follows the invariance of the Hausdorff closure (i.e. closure in the Hausdorff metric) of this family. Therefore from dissipativity of the operator with respect to the family of compacts follows the dissipativity with respect to Hausdorff closure of this family.

Let  $W$  be a convex compact in  $\mathbb{R}^n$  with  $0 \in \text{ri}W$ . In this case  $\text{Aff}W$  is a linear subspace, and if the operator  $A$  is dissipative with respect to  $W$ , then  $\text{Aff}W$  is invariant with respect to  $A$ . Introduce the following functional on the subspace  $L(W)$  of the space  $L(\mathbb{R}^n)$  (of linear operators in  $\mathbb{R}^n$ ), consisting of the operators, with respect to which  $\text{Aff}W$  is invariant:

$$\mu_W(A) = \sup_{x \in W} \mu_W(Ax). \quad (5)$$

Here  $\mu_W$  is the Minkowski functional of the set  $W$  (defined, for example, [8]) in the subspace  $\text{Aff}W$ .

It is easy to see that  $A \in K(W)$  if and only if

$$\mu_W(\exp(At)) \leq 1$$

for all  $t \geq 0$ .

In particular, the operator  $A \in L(W)$  is *strongly dissipative with respect to convex compact*  $W$  if exists such  $\varepsilon > 0$  that  $\mu_W(\exp(At)) \leq \exp(-\varepsilon t)$  at all  $t \geq 0$ . In general, the operator  $A$  is strongly dissipative with respect to convex compact  $W$  if and only if  $A + \varepsilon I \in K(W)$  for some  $\varepsilon > 0$ .

If  $W$  is a ball of the norm  $\|\cdot\|$ , then strong dissipativity with respect to  $W$  means the existence of such  $\varepsilon > 0$  that  $\|\exp(At)\| \leq \exp(-\varepsilon t)$  for all  $t \geq 0$ . We come to the definition of *stable dissipativity with respect to the norm* [11, 12, 19].

Introduce in  $L(\mathbb{R}^n)$  the following functional:

$$\gamma_W(A) = \lim_{h \rightarrow +0} \frac{\mu_W(I + hA) - 1}{h}.$$

In the case, when  $W$  is a ball of some norm (i. e.  $\mu_W$  is a norm), arrive at the known definition of the logarithmic Lozinsky norm [9, 10].

**Lemma 1.** *The operator  $A \in L(\mathbb{R}^n)$  is dissipative (strongly dissipative) with respect to  $W$ , if and only if the inequality  $\gamma_W(A) \leq 0$  ( $\gamma_W(A) < 0$ ) is satisfied.*

**Proof.** *Sufficiency.* The following inequality is obtained in [9]

$$\|\exp(At)\| \leq \exp(\gamma(A)t)$$

where  $\gamma(A)$  is the Lozinsky norm of the operator  $A$ , corresponding to the norm  $\|\cdot\|$ . By literal repetition of the reasoning from [9] (with a substitution of the norm by Minkowski functional), one can obtain the inequality

$$\mu_W(\exp(At)) \leq \exp(\gamma_W(A)t)$$

for all  $t \geq 0$ , from which immediately follows the sufficiency.

*Necessity.* Evidently,

$$\mu_W(\exp(At)) = \mu_W(I + At) + o(t) \quad (t \rightarrow 0).$$

Therefore,

$$\gamma_W(A) = \lim_{h \rightarrow +0} \frac{\mu_W(e^{Ah}) - 1}{h}.$$

Let  $\varepsilon \geq 0$ . If  $\mu_W(\exp(At)) \leq \exp(-\varepsilon t)$  at all  $t \geq 0$ , then

$$\gamma_W(A) \leq \lim_{h \rightarrow +0} \frac{\exp(-\varepsilon h) - 1}{h} = -\varepsilon,$$

which proves the necessity. The lemma is proved.

Assign a relatively open convex cone  $Q_x(W)$  to every point  $x \in r\partial W$  according to the rule:  $y \in Q_x(W)$  if and only if there exists such  $\varepsilon > 0$  that

$$x + \varepsilon y \in \text{ri } W.$$

**Lemma 2.** *For strong dissipativity of  $A$  with respect to convex compact  $W$  it is necessary and sufficient that for every point  $x \in r\partial W$  the inclusion*

$$Ax \in Q_x(W)$$

be true. For dissipativity of  $A$  with respect to  $W$  it is necessary and sufficient that for every point of  $X \in r\partial W$  the inclusion

$$Ax \in \overline{Q_x(W)}$$

be true.

**Proof.** Note that the operator  $A$  is strongly dissipative with respect to  $W$  if and only if there exists such  $t_0 > 0$  that  $\mu_W(I + At_0) < 1$ . Indeed, the existence of such  $t_0$  for a strongly dissipative operator follows immediately from the negativeness of  $\gamma_W(A)$ . Conversely, if  $\mu_W(I + At_0) < 1$ , then there exists such  $\varepsilon > 0$  that  $\mu_W(I + (A + \varepsilon I)t_0) < 1$ . But then  $\gamma_W(A + \varepsilon I) \leq 0$ , the operator  $(A + \varepsilon I)$  is dissipative. It means that  $A$  is strongly dissipative.

If the operator  $A$  is strongly dissipative with respect to  $W$ , then, according to the above, for each  $x \in r\partial W$  there exists such  $t_x > 0$  that  $(I + t_x A)x \in \text{ri} W$ . It means that the vector  $Ax$  belongs to the cone  $Q_x(W)$ .

Conversely, let the latter condition be satisfied. According to the hypothesis of the theorem and convexity of  $W$ , for each  $x \in r\partial W$  there exists the only positive number  $s = s(x)$  such that  $(I + sA)x \in r\partial W$ . Show that  $s_0 = \inf_{x \in r\partial W} s(x) > 0$ . Let it be not so. Then there exists such a subsequence  $\{x_n\}_{n=1}^{+\infty}$  that  $\lim_{n \rightarrow \infty} s(x_n) = 0$ . Choose from  $\{x_n\}$  a converging subsequence  $\{x'_n\}$ . Let  $\tilde{x} = \lim_{n \rightarrow \infty} x'_n$ . For every  $n \in N$  and for every  $\varepsilon > 0$

$$[I + (s(x'_n) + \varepsilon)A]x'_n \notin W.$$

Passing to the limit, obtain

$$(I + \varepsilon A)\tilde{x} \notin \text{ri} W$$

which contradicts the hypothesis of the theorem.

Thus,  $s_0 > 0$ . For any  $t_0 \in (0; s_0)$  is true  $\mu_W(I + At_0) < 1$ , i. e. the operator  $A$  is strongly dissipative.

If  $A \in K(W)$ , then for any  $\varepsilon > 0$  we have  $AX - \varepsilon x \in Q_x(W)$  (for any  $x \in r\partial W$ ), i. e.  $Ax \in \overline{Q_x}$ . Conversely, if  $Ax \in \overline{Q_x}$ , then  $Ax - \varepsilon x \in Q_x$  at any  $\varepsilon > 0$ , and  $A$  represents a limit point of the family of dissipative operators, i. e.  $A \in K(W)$ . The lemma is proved.

**Remark 4.** Immediately from the Krein—Milman theorem [8] follows that it is sufficient to require from the lemma conditions that inclusions be satisfied not for all points  $x \in r\partial W$ , but for extremal points of  $W$  only. In particular, if  $W$  is a polyhedron, then it is sufficient to test its vertices only. Thus, to elucidate the question about dissipativity (strong dissipativity) of the operator with respect to the polyhedron, one should test only the fulfillment of finite number of linear inequalities.

**Remark 5.** In the proof lemma 2 we have used the obvious fact: the closure of the set  $K(W)$ .

One more fact follows directly from lemma 2.

**Lemma 3.** The set  $K(W)$  is a closed convex cone. The cone of all strongly dissipative with respect to  $W$  operators coincides with  $\text{ri} K(W)$  and with  $\bigcup_{\varepsilon > 0} (K(W) - \varepsilon I)$ . If  $\{W_\nu\}$  is a family of convex compacts with  $0 \in \text{ri} W_\nu$  for all  $\nu$ , then  $K(\{W_\nu\})$  is a closed convex cone.

**Remark 6.** If  $\text{int} W = \emptyset$ , then  $\text{int} K(W) = \emptyset$ . Indeed, if  $\text{Aff} W$  is invariant with respect to the operator  $A_1$ , then  $A + \varepsilon A_1 \notin K(W)$  at  $\varepsilon \neq 0$ . If  $\text{int} W \neq \emptyset$ , then  $\text{int} K(W)$  is also non-empty and coincides with  $\text{ri} K(W)$ .

**Definition 4.** The operator  $A \in K(W)$  is called *stable (or roughly) dissipative with respect to  $W$* , if  $A \in \text{int} K(W)$ .

Definition 4 generalize the definition of the stable dissipativity with respect to the norm [11, 19].

Pass to the consideration of operators dissipative with respect to interval spaces. Let find out for which interval spaces  $\mathbb{J}$  the interior of the cone  $K(\mathbb{J})$  is not empty.

Let  $V$  be a set of all compact convex bodies in  $\mathbb{R}^n$ . Fix some norm  $\|\cdot\|$  in  $\mathbb{R}^n$  and assume

$$d(W) = \min_{x \in \partial W} \|x\|.$$

**Lemma 4.** *The function  $d(W)$  is continuous according to Hausdorff on the set  $V$ .*

**Proof.** First note that if  $X \in V$ ,  $Y \in V$ , then  $\rho_H(\partial X, \partial Y) \leq \rho_H(X, Y)$ . Indeed, let  $\rho_H(X, Y) \leq \varepsilon$ . Then  $X \subset Y + S_\varepsilon$  where  $S_\varepsilon = \{x \in \mathbb{R}^n \mid \|x\| \leq \varepsilon\}$ . Let, further on, there exists such  $y_0 \in (\partial Y) \cap X$  that  $y_0 \notin S_\varepsilon + \partial X$ . Construct at the point  $y_0$  a tangent hyperplane  $L$  to  $Y$ . Let  $l$  be the direction of the external normal to  $\partial Y$  at the point  $y_0$  orthogonal to  $L$ . Draw a ray from the point  $y_0$  in the direction of  $l$  to the point  $x_0$  of crossing with  $\partial X$ . Construct such a ball  $S$  of the norm  $\|\cdot\|$  with the center at the point  $x$  that  $y_0 \in \partial S$ . The radius of  $S$  is larger than  $\varepsilon$  and  $S \cap Y = \{y_0\}$ . Thus, if one constructs a ball  $S' \subset S$  of the radius  $\varepsilon$  with the center at  $x_0$ , then

$$S' \cap Y = \emptyset.$$

But then  $x_0 \notin Y + S_\varepsilon$ , i. e.  $X \not\subset Y + S_\varepsilon$  what is contrary to the assumption.

The existence of such  $y_0 \in \partial Y$  that  $y_0 \notin X \cup (\partial X + S_\varepsilon)$  is also impossible, since then  $y_0 \notin X + S_\varepsilon$ , i. e.  $Y \not\subset X + S_\varepsilon$ . Consequently,  $\partial Y \subset S_\varepsilon + \partial X$ , and that means  $d(X) \leq d(Y) + \varepsilon$ . Similarly,  $d(Y) \leq d(X) + \varepsilon$ . It means that  $|d(X) - d(Y)| \leq \varepsilon$ , and the function  $d(W)$  is continuous on  $V$ . The lemma is proved.

**Lemma 5.** *For non-emptiness of  $\text{int } K(\mathbb{J})$  it is necessary and sufficient for all the elements of the interval space  $\mathbb{J}$ , except  $\{0\}$ , to possess non-empty interior.*

**Proof.** *Necessity.* Follows immediately from remark 6.

*Sufficiency.* Show that under the conditions of the theorem the inclusion

$$-I \in \text{int } K(\mathbb{J}) \tag{6}$$

takes place.

To each point  $x$  ( $\|x\| = 1$ ) we assign the set  $W_x$  according to the rule:

$$W_x = \bigcap_{W \ni x, W \in \mathbb{J}} W.$$

According to the conditions b) and c) from definition 1,  $W_x \in \mathbb{J}$ . The set

$$\widetilde{W} = \overline{\bigcup_{\|x\|=1} W_x}$$

is compact. Indeed,  $\widetilde{W}$  is contained in any element of  $\mathbb{J}$  containing unit ball of the norm  $\|\cdot\|$ ; such an element exists due to non-emptiness of the interior of all intervals (except  $\{0\}$ ) and the property a) from definition 1. Note that Hausdorff closure of the family  $\{W_x \mid \|x\| = 1\}$  represents a compact in the Hausdorff metric, contained in  $\mathbb{J}$  (this follows from compactness according to Hausdorff of the family of all compact subsets of the compact [6]). From the property a) (definition 1) follows (by virtue of lemma 4) the existence of such  $\varepsilon > 0$  that  $d(W_x) \geq \varepsilon$  for all such  $x$  that  $\|x\| = 1$  (indeed,  $d(W_x) > 0$ , since  $0 \in \text{int } W_x$ ).

Thus, there exists such  $\varepsilon > 0$  that for all  $x$  ( $\|x\| = 1$ ) the inclusion

$$Ax \in \text{int } W_x$$

is true if  $\|A\| < \varepsilon$ .

In other words,  $Ax - x \in Q_x(W_x)$  if  $\|A\| < \varepsilon$ ,  $\|x\| = 1$  (see lemma 1). The more so, as  $Ax - x \in Q_x(W)$  for all  $W \in \mathbb{J}$  ( $W \ni x, \|A\| < \varepsilon$ ) at all such  $x$  that  $\|x\| = 1$ . But then  $Ax - x \in Q_{\alpha x}(\alpha W)$  for all  $\alpha > 0$ ,  $\|A\| < \varepsilon$ . Hence,  $A - I \in K(\mathbb{J})$ , i. e. (6) is satisfied. The lemma is proved.

Thus, we have shown that under the conditions of lemma 5  $K(\mathbb{J})$  is a convex solid cone.

**Definition 5.** The operator is *stable dissipative with respect to the interval space*  $\mathbb{J}$  if it belongs to  $\text{int } K(\mathbb{J})$ .

For stable dissipative operators the remark 2 is true: if an operator is stable dissipative with respect to the family of compacts and the interior of at least one of them is not empty, then it is stable dissipative with respect to some norm.

### 1.3. The Wrapping Effect for Autonomous Systems

The results of the previous section can be applied to the study of the wrapping effect. First we give the exact definition of what we understand by the wrapping effect.

Let in the vicinity of a compact convex body  $B \subset \mathbb{R}^n$  be given a smooth autonomous system

$$\frac{dx}{dt} = f(x) \quad (7)$$

with  $B$  positively invariant with respect to (7), and let  $x(0)$  be determined inexactly, namely

$$x(0) \in x_0 + W_0,$$

where  $x_0 \in B$ ,  $W_0 \in \mathbb{J}$ ,  $x_0 + W_0 \in B$ ,  $\mathbb{J}$  is some interval space (see definition 1).

**Remark 7.** Irrespective of particular numerical method (i. e. dealing with the exact solution of the initial value problem for (7) with the initial conditions  $x(0) = x_0$ ) a stepwise interval solution with step  $h > 0$  can be described as follows.

Let  $T_h$  be the transformation of the phase flow of (7) during the time  $t$  (*shift over time*  $t$ ),  $W_0 \in \mathbb{J}$  is the initial interval (its sense is an uncertainty in initial data). Assume

$$X_0 = x_0 + W_0,$$

$$X_{m+1} = T_{(m+1)h}x_0 + W_{m+1},$$

$$W_{m+1} = \bigcap_{W \supset W_{m+1}(h), W \in \mathbb{J}} W,$$

$$W_{m+1}(h) = T_h(T_{mh}x_0 + W_m) - T_{(m+1)h}x_0.$$

The sequence  $\{X_m\}_{m=0}^{+\infty}$  is the *exact stepwise interval solution* of (7).

**Definition 6.** The absence of *infinitesimal wrapping effect* (IWE) means that for any  $h > 0$  the sequence  $\{W_m\}_{m=0}^{+\infty}$  is enclosed:  $W_m \supset W_{m+1}$  for all  $m$ , i. e. the obtained intervals do not expand.

With IWE the intervals expand along any trajectory (7) for any small step, and that means that when solving a system by a stepwise interval numerical method with any small step the interval expansion takes place for any initial data irrespective of the applied method (since it is true even for exact solutions).

Generalizing the construction [10] for norms, introduce the following functional:

$$N_W(x, y) = \lim_{h \rightarrow +0} \frac{\mu_W(x + hy) - \mu_W(x)}{h}.$$

Literally (with substitution of the norm for Minkowski functional) repeating the reasoning from [10] (pp.127, 426), come to the following statements.

**Statement 1.** If  $x(t)$  with values in  $\mathbb{R}^n$  is differentiable on connected subset  $T$  of the real axis, and  $W$  is a convex compact ( $0 \in \text{ri } W$ ), then the function  $\mu_W(x(t))$  is almost everywhere differentiable on  $T$  and the derivative (where it exists) coincides with the right-hand derivative, equal to  $N_W(x(t), \dot{s}(t))$ . The right-hand derivative of  $\mu_W(x(t))$  exists everywhere on  $T$  except the right-hand end.

**Statement 2.**

$$\gamma_W(A) = \sup_{x \in W} N_W(x, AX).$$

By  $f'(x)$  further on we denote the mapping derivative of  $f$ .

The main part of further results on IWE can be obtained from the following theorem.

**Theorem 1.** *Let in the region  $U \subset \mathbb{R}^n$  be given a smooth autonomous system (7),  $B \subset U$  be positively invariant with respect to (7) compact convex body. IWE is absent for compact  $B$ , system (7) and interval space  $\mathbb{J}$  if and only if*

$$f'(x) \in K(\mathbb{J}) \tag{8}$$

for all  $x \in B$ , i. e. for any  $x \in B$  the Jacobi matrix of system (7) in the point  $x$  is strongly dissipative with respect to  $\mathbb{J}$ .

**Proof. Sufficiency.** Let  $W \in \mathbb{J}$ . Consider two solutions  $x_1(t), x_2(t)$  of system (7) with initial conditions from  $B$ . Denote  $\Delta(t) = x_1(t) - x_2(t)$ . Using statements 1 and 2 and the theorem on finite increment, estimate the derivative of  $\mu_W(\Delta(t))$ :

$$\begin{aligned} \frac{d}{dt} \mu_W(\Delta(t)) &= N_W(\Delta(t), d\Delta(t)/dt) \leq \\ &\leq \sup_{0 \leq \Theta \leq 1} N_W(\Delta(t), f'(x_c(t))\Delta(t)) \leq \\ &\leq \sup_{0 \leq \Theta \leq 1} \gamma_W(f'(x_c(t)))\mu_W(\Delta(t)), \end{aligned}$$

where  $x_c(t) = x_1(t) + \Theta(x_2(t) - x_1(t))$ ,  $0 \leq \Theta \leq 1$  for all  $t \geq 0$ . By (8) and statement 1 we obtain

$$\frac{d}{dt} \mu_W(\Delta(t)) \leq 0.$$

Since the latter inequality holds for all  $t \geq 0$  and for all  $W \in \mathbb{J}$ , in system (7) on  $B$  IWE with respect to  $\mathbb{J}$  is absent.

**Necessity.** Let  $W \in \mathbb{J}$ ,  $x_0 \in \text{int } B$ ,  $t_0 \geq 0$ ,  $y \in \text{Aff } B$ ,  $y \neq 0$ . There exists such  $h_0 > 0$  that  $x_0 + h_0 y \in B$ . Due to smoothness of system (7) there exist and are unique the solutions  $x_1(t), x_2(t)$  of the initial value problem for (7) with the initial conditions  $x_1(t_0) = x_0$ ,  $x_2(t_0) = x_0 + h_0 y$ . Assume  $\Delta(t) = x_1(t) - x_2(t)$ . Then

$$\frac{d}{dt} \ln \mu_W(\Delta(t))|_{t=t_0} =$$

$$\begin{aligned}
&= N_W \left( \frac{\Delta(t_0)}{\mu_W(\Delta(t_0))}, f'(x_c) \frac{\Delta(t_0)}{\mu_W(\Delta(t_0))} \right) = \\
&= N_W \left( \frac{y}{\mu_W(y)}, f'(x_c) \frac{y}{\mu_W(y)} \right),
\end{aligned}$$

where  $x_c = x_0 + \Theta h_0 y$ ,  $0 < \Theta < 1$ .

By virtue of absence of IWE

$$\frac{d}{dt} \ln \mu_W(\Delta(t)) \leq 0$$

for all  $t \geq 0$ . Since if  $x_0 + h_0 y \in b$ , then:

- (a)  $x_0 + h y \in B$  for all  $h \in [0; h_0]$ ,
- (b) a set of those  $h \in [0; h_0]$ , for which

$$N_W \left( \frac{y}{\mu_W(y)}, f'(x) \frac{y}{\mu_W(y)} \right) \leq 0,$$

is dense on the segment  $[0; h_0]$ , and (c) due to its closeness coincides with this segment.

By virtue of the arbitrary choice of  $x_0$  for any  $x_0 \in \text{int } B$ ,  $t \geq 0$ ,  $y \in \text{Aff } B$ ,  $y \neq 0$  the inequality

$$N_W \left( \frac{y}{\mu_W(y)}, f'(x) \frac{y}{\mu_W(y)} \right) \leq 0$$

is satisfied. It holds also for any  $x \in B$ ,  $t \geq 0$ ,  $y \in \text{Aff } B$ ,  $y \neq 0$ . Hence, from lemma 1 and statement 2 immediately follows dissipativity of  $f'(x)$  with respect to  $\mathbb{J}$  for all  $x \in B$ . The theorem is proved.

**Definition 7.** The family of linear operators  $\{A_\alpha\}$  is called *simultaneously dissipative*, if there exists a norm relative to which all the operators are dissipative.

Simultaneously dissipative operators were studied in detail in [11, 12, 17–22].

From theorem 1, example 3, and remark 2 we obtain the following theorem.

**Theorem 2.** *For existence of interval space in which at least one interval possesses non-empty interior and with respect to which in system (7) there is no IWE on  $B$ , it is necessary and sufficient for the family  $\{f'(x) \mid x \in B\}$  to be simultaneously dissipative.*

Thus, the problem of existence of the interval space, with respect to which IWE is absent, is reduced to the problem of simultaneous dissipativity of Jacobi matrices. As sought for space one can choose a set of all balls of that norm relative to which all Jacobi matrices are dissipative. This norm is *constricting* for (7) on  $B$  (i.e. the distance between two solutions with initial conditions from  $B$  will not expand with time). Hence, all systems without IWE (with respect to some interval space) on  $B$  are globally stable in  $B$  (see introduction).

Bellow by  $C^1(B)$  we denote the Banach space of smooth mappings of  $B$  in  $\mathbb{R}^n$  with the norm

$$\|f\|_{C^1(b)} = \max_{x \in B} \|f(x)\| + \sum_{k=1}^n \max_{x \in B} \left\| \frac{\partial f}{\partial x_k} \right\|$$

where  $\|\cdot\|$  is a fixed norm in  $\mathbb{R}^n$ .

Further on, speaking about properties of autonomous systems, we mean the properties of the vector fields generating them.

Immediately from lemma 3 and theorem 1 the following statement can be obtained.

**Theorem 3.** *The set of systems on  $B$  without IWE with respect to  $\mathbb{J}$  is closed convex cone in  $C^1(B)$ .*

For this cone we use the notation  $F_B(\mathbb{J})$ .

Further on, speaking about the vicinity of an autonomous system in  $C^1(B)$  we mean a part of the vicinity, consisting only of those systems for which the set  $B$  is positively invariant.

Let us study under what conditions the interior of the cone  $F_B(\mathbb{J})$  is non-empty.

**Theorem 4.** *For non-emptiness of  $\text{int } F_B(\mathbb{J})$  in  $C^1(B)$  it is necessary and sufficient for all elements of  $\mathbb{J}$ , except  $\{0\}$ , to possess non-empty interior.*

**Proof.** *Necessity.* Let exist such a set  $W \in \mathbb{J}$  that  $\text{int } W = 0$ . Consider any system (7) without IWE with respect to  $\mathbb{J}$  on  $B$ . Since  $\text{int } B \neq 0$ , there exist two different concentric balls  $S_1$  and  $S_2$  of usual  $l^2$ -norm, belonging to  $\text{int } B$  with  $S_1 \subset S_2$ . Construct such a function  $g \in C_\infty(\mathbb{R}^n)$  that  $g(x) = 1$  for all  $x \in S_1$  and  $g(x) = 0$  at all  $x \notin S_2$ . Since  $\text{Aff } W \neq \mathbb{R}^n$ , one can construct a linear operator  $A \in L(\mathbb{R}^n)$  mapping  $\text{Aff } W$  into such a subspace  $E_0 \neq \{0\}$  that  $(\text{Aff } W) \cap E_0 = \{0\}$ .

Consider the system

$$\frac{dx}{dy} = f(x) + \varepsilon g(x)AX, \quad (9)$$

where  $\varepsilon > 0$  is arbitrary. The set  $B$  is positively invariant with respect to (9), since the vector field generating (9) coincides with  $f$  in the vicinity of  $\partial B$ . On the other hand, there exist Jacobi matrices (9) relative to which  $\text{Aff } W$  is not invariant, i. e. in (9) exist IWE with respect to  $\mathbb{J}$  on  $B$ . Since in any vicinity of  $f$  there is at least one vector field, generating (9), then

$$\text{int } F_B(\mathbb{J}) = 0.$$

*Sufficiency.* Consider the system  $dx/dt = -x$ . It is a system on  $B$  without IWE with respect to  $\mathbb{J}$ . Furthermore, if all elements of  $\mathbb{J}$ , except  $\{0\}$ , possess non-empty interior, then by lemma 5 the matrix of the system is stable dissipative with respect to  $\mathbb{J}$  (see definition 5).

Consider the system:

$$\frac{dx}{dy} = -x + v(x), \quad (10)$$

where  $\|v\|_{C_1(B)} < \varepsilon$  with  $\varepsilon$  chosen so that

$$A - I \in K(\mathbb{J})$$

if  $\|A\| < \varepsilon$  (see the proof of lemma 5). Then all Jacobi matrices (10) are dissipative with respect to  $\mathbb{J}$ , and if  $v$  is chosen so that  $B$  is positively invariant with respect to (10), then in (10) IWE is absent (by theorem 1). The theorem is proved.

Thus,  $\text{int } F_B(\mathbb{J}) \neq 0$  if and only if  $\text{int } K(\mathbb{J}) \neq 0$ .

It is easy to see that in the proof of sufficiency in theorem 4 one can instead of the system  $dx/dy = -x$  consider any system whose Jacobi matrices are stable dissipative with respect to  $\mathbb{J}$ .

**Remark 8.** By analogy with the space  $C^1(B)$  one can construct the Banach spaces  $C^k(B)$  ( $k \in \mathbb{N}$ ) with the norm

$$\|f\|_{C^k(B)} = \sum_{|\alpha|=0}^k \max_{x \in B} \|(D^\alpha f)(x)\|,$$

where  $\alpha = (\alpha_1, \dots, \alpha_n)$  is a multiindex:

$$|\alpha| = \alpha_1 + \dots + \alpha_n, D^\alpha f = \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \dots \partial x_n^{\alpha_n}}$$

and the metric space  $C^\infty(B)$  with the system of seminorms

$$\left\{ \max_{x \in B} \|(D_\alpha f)(x)\| \mid |\alpha| \leq m \right\}_{m=0}^{+\infty}.$$

Small  $C^k$ -additions ( $1 \leq k \leq +\infty$ ) are small and in the  $C^1$ -norm. Therefore, for  $C^k$ -smooth systems under the conditions of theorem 4 the interior of  $F_B(\mathbb{J})$  is non-empty and in  $C^k(B)$ . As shows the proof of theorem 4 (necessity), if conditions of the theorem are not satisfied, then the interior of  $F_B(\mathbb{J})$  in  $C^k(B)$  is empty.

Let clarify what autonomous system without IWE in specific interval spaces looks like.

**Theorem 5.** *Any system without IWE with respect to  $\mathbb{J}$  from example 1 has the form:*

$$\frac{dx}{dt} = ax + c,$$

where  $a \leq 0$ ,  $C \in \mathbb{R}^n$  is a constant vector.

**Proof.** Let  $A \in K(\mathbb{J})$ . All the segments symmetrical with respect to 0 belong to  $\mathbb{J}$ . Every such a segment has the form  $\{y \in \mathbb{R}^n \mid y = ax, |a| \leq 1\}$  for some  $x \in \mathbb{R}^n$ . The cone  $Q_x$  (see lemma 2) for each segment consists of vectors of the form  $ax$ , where  $a < 0$ . Thus, every non-zero vector  $x \in \mathbb{R}^n$  is eigenvector of the operator  $A$ , corresponding to non-positive eigenvalue. Thus:

$$K(\mathbb{J}) = \{aI \mid a \leq 0\}. \quad (11)$$

Let now a system without IWE have the form

$$\begin{cases} \frac{dx_1}{dt} = f_1(x_1, \dots, x_n); \\ \dots \\ \frac{dx_n}{dt} = f_n(x_1, \dots, x_n). \end{cases}$$

According to (11) and theorem 1

$$\frac{\partial f_i}{\partial x_j} \equiv 0 \quad (i \neq j); \quad (12)$$

$$\frac{\partial f_1}{\partial x_1} \equiv \frac{\partial f_2}{\partial x_2} \equiv \dots \equiv \frac{\partial f_n}{\partial x_n} \leq 0. \quad (13)$$

From (12) follows that  $f_k$  depends only on  $x_k$  ( $k = 1, \dots, n$ ). It means that  $\partial f_k / \partial x_k$  also depends only on  $x_k$ , i. e. by virtue of (13)  $\partial f_k / \partial x_k = \text{const}$  ( $k = 1, \dots, n$ ). Then

$$\frac{\partial f_1}{\partial x_1} \equiv \frac{\partial f_2}{\partial x_2} \equiv \dots \equiv \frac{\partial f_n}{\partial x_n} \equiv a \leq 0$$

and the system has the form:

$$\frac{dx}{dt} = ax + c,$$

where  $a \leq 0$ ,  $c = \text{const}$ . The theorem is proved.

Thus, whatever nonlinear (or even linear with non-scalar matrix) system we consider, if we take as  $\mathbb{J}$  the interval space of example 1 (or any wider space), IWE will be present in the system. From theorem 5 also follows that any dissipative with respect to all norms operator has the form  $aI$ , where  $a \leq 0$  (see also remark 3).

**Example 6.** Consider  $\mathbb{J}$  from example 2.  $\mathbb{J}$  contains all symmetrical with respect to 0 segments of coordinate axes (thus, the conditions of theorem 4 are not satisfied, i. e.  $\text{int } F_B(\mathbb{J}) = 0$ ). Let  $A \in K(\mathbb{J})$ . Reasoning like in proof of theorem 5, conclude that all coordinate axes are eigenspaces of the operator  $A$ , corresponding to non-positive eigenvalues. In other words, the matrix of the operator  $A$  is diagonal and non-positive. On the other hand, by virtue of lemma 2, all such operators belong to  $K(\mathbb{J})$ . Thus, systems without IWE with respect to  $\mathbb{J}$  on  $B$  have the form

$$\begin{cases} \frac{dx_1}{dt} = f_1(x_1); \\ \dots \\ \frac{dx_n}{dt} = f_n(x_n), \end{cases}$$

where

$$\frac{\partial f_k}{\partial x_k} \leq 0 \quad (k = 1, \dots, n)$$

for all  $x \in B$ .

From the considered example follows that when using standard intervals (rectangular parallelepipeds) IWE will be observed in almost all systems in  $\mathbb{R}^n$  if  $n \neq 1$ .

The systems without IWE with respect to  $\mathbb{J}$  from example 3 on  $B$  represent all systems for which the norm  $\|\cdot\|$  is constricting in  $B$  (see the text after theorem 2).

**Remark 9.** Note that testing of dissipativity (stable dissipativity) of the operator with respect to the norm is equivalent to non-positiveness (negativeness) of the corresponding Lozinsky norm. For some norms an explicit form of corresponding Lozinsky norm is known (see, for example, [9] or [10, p. 463–465]). In particular, for the Euclidean norm the Lozinsky norm of the operator  $A$  coincides with the largest eigenvalue of the operator  $(A^* + A)/2$ . The Lozinsky norm of the operator  $A$  represented by the matrix  $(a_{ij})_{i,j=1}^n$  with respect to  $l^1$ - and  $l^\infty$ -norms is given by the formulae, respectively:

$$\max_{1 \leq i \leq n} (\text{Re } a_{ii} + \sum_{j \neq i} |a_{ji}|);$$

$$\max_{1 \leq i \leq n} (\text{Re } a_{ii} + \sum_{j \neq i} |a_{ij}|).$$

In remark 9 it is assumed that the operator  $A$  acts in the space  $C^n$ . The definitions and used here properties of dissipative operators in complex spaces are analogous to those in real ones.

**Example 7.** Let  $\mathbb{J}$  be the interval space from example 4. Then

$$K(\mathbb{J}) = \bigcap_{1 \leq k \leq m} K_{\|\cdot\|_k},$$

where  $K_{\|\cdot\|_k}$  is the cone of all operators dissipative with respect to the norm  $K_{\|\cdot\|_k}$ . It follows from the closure under intersection of the family of all positively invariant sets of an autonomous system. Similarly,

$$\text{int } K(\mathbb{J}) = \bigcap_{1 \leq k \leq m} \text{int } K_{\|\cdot\|_k}.$$

Let, for example,  $\|\cdot\|_1$  be  $l^\infty$ -norm,  $\|\cdot\|_2$  be  $l^2$ -norm in  $\mathbb{R}^2$ . Then the conditions of stable dissipativity of the operator  $A$  with the matrix  $(a_{ij})_{i,j=1}^2$  with respect to  $\mathbb{J}$  according to remark 9 are of the form:

$$\begin{cases} 4a_{11}a_{22} > (a_{12} + a_{21})^2; \\ a_{11} + |a_{12}| < 0; \\ |a_{21}| + a_{22} < 0. \end{cases}$$

Thus, for the system of the form

$$\begin{cases} \frac{dx_1}{dt} = f_1(x_1, x_2); \\ \frac{dx_2}{dt} = f_2(x_1, x_2) \end{cases} \quad (14)$$

if the inequalities

$$\begin{cases} 4\frac{\partial f_1}{\partial x_1} \cdot \frac{\partial f_2}{\partial x_2} > \left(\frac{\partial f_1}{\partial x_2} + \frac{\partial f_2}{\partial x_1}\right)^2; \\ \frac{\partial f_1}{\partial x_1} + \left|\frac{\partial f_1}{\partial x_2}\right| < 0; \\ \left|\frac{\partial f_2}{\partial x_1}\right| + \frac{\partial f_2}{\partial x_2} < 0 \end{cases}$$

are satisfied and the compact convex body  $B$  is positively invariant with respect to (14), then in (14) IWE with respect to  $\mathbb{J}$  (from example 4) is absent on  $B$ . For example, such is the following system:

$$\begin{cases} \frac{dx_1}{dt} = -2x_1 + x_2; \\ \frac{dx_2}{dt} = 2x_1 - 3x_2 \end{cases}$$

if  $B$  is the square  $\{(x_1, x_2) \mid |x_1| \leq 1, |x_2| \leq 1\}$  or the circle  $\{(x_1, x_2) \mid x_1^2 = x_2^2 \leq 1\}$ .

**Example 8.** Consider  $\mathbb{J}$  from example 5. Let  $Q$  be rectangular triangle with vertices at the points  $(-1; 2); (-1; -1); (1; -1)$ .

From lemma 2 and theorem 1 follows that the cone  $F_B(\mathbb{J})$  consists of the systems of the form (14), with respect to which the compact  $B$  is positively invariant and for which

$$\begin{cases} \frac{\partial f_1}{\partial x_1} + \frac{\partial f_1}{\partial x_2} \leq 0; \\ \frac{\partial f_1}{\partial x_1} - 2\frac{\partial f_1}{\partial x_2} \leq 0; \\ \frac{\partial f_2}{\partial x_1} + \left|\frac{\partial f_2}{\partial x_2}\right| \leq 0; \\ -3\frac{\partial f_1}{\partial x_1} + 6\frac{\partial f_1}{\partial x_2} - 2\frac{\partial f_2}{\partial x_1} + 4\frac{\partial f_2}{\partial x_2} \leq 0; \\ 3\frac{\partial f_1}{\partial x_1} - 3\frac{\partial f_1}{\partial x_2} + 2\frac{\partial f_2}{\partial x_1} - 2\frac{\partial f_2}{\partial x_2} \leq 0 \end{cases} \quad (15)$$

is true.

Substituting all the inequality signs in (15) by strict ones, obtain  $\text{int } F_B(\mathbb{J})$ . For example, the system

$$\begin{cases} \frac{dx_1}{dt} = -x_1; \\ \frac{dx_2}{dt} = -6x_1 - 4x_2 \end{cases}$$

belongs to  $\text{int } F_B(\mathbb{J})$  for  $B = Q$ .

Corresponding ball  $S$  (see remark 2) is the parallelogram with the vertices in the points  $(-1; 2); (-1; -1); (1; -2); (1; 1)$ . From remark 2 follows that

$$K(\mathbb{J}) = K(Q) \subset K(S).$$

One can see that  $K(\mathbb{J}) \neq K(S)$ . For example, the operator given by the matrix

$$\begin{pmatrix} -4 & -1 \\ 2 & 0 \end{pmatrix}$$

is dissipative with respect to  $S$ , but it is not dissipative with respect to  $\mathbb{J}$ . In other words, in the systems without IWE with respect to  $\{\alpha S \mid \alpha \geq 0\}$  (i. e. constricting according to the norm whose ball is  $\mathbb{J}$ ) there can be observed IWE with respect to  $\mathbb{J}$ .

This example can be generalized as follows. Consider  $\mathbb{J}$  from remark 2. In system (7) on  $B$  IWE is absent with respect to  $\mathbb{J}$  if the operators  $f'(x)$  for all  $x \in B$  are dissipative with respect to all sets  $Q_k$  ( $k = 1, \dots, m$ ).

To sum up, one can say the following. When using sufficiently wide interval spaces, then in accordance with theorem 4, IWE is observed in almost all systems. In particular, IWE takes place almost for all systems when using standard intervals (see example 6). Expansion of the interval space results in the appearance of new systems with IWE: thus, in using a set of all symmetrical to 0 convex compacts IWE is absent only for linear systems with non-positive scalar matrices. And the most impotent: the question about the existence of interval space, with respect to which in the considered system IWE is absent, is reduced to the problem of simultaneous dissipativity of the Jacobi matrices. Therefore, there is no interval space with respect to which all (or even if in some sense almost all) globally stable systems would have no IWE. One has to solve individually problems of the existence and constructing of corresponding interval spaces for each particular system. These problems are solved constructively very rarely.

We have treated the wrapping effect in a very strong sense. The condition of boundedness of the sequence of intervals  $\{W_m\}_{m=0}^{+\infty}$  at any step  $h > 0$  (see remark 7) is weaker (and acceptable, generally speaking, for constructing sufficiently narrow interval solutions). This condition can be called the condition of absence of the *asymptotic wrapping effect* (AWE). It is the weakest from acceptable conditions, since with AWE it is impossible to use stepwise interval methods to obtain narrow interval solutions at large times. The study of AWE is still not completed. It is evident only that for a linear autonomous system in considering the interval space from example 3 AWE is equivalent to IWE. One can suggest a hypothesis: the problem of existence and constructing of the interval space with respect to which AWE is absent in the autonomous system is reduced to the question of simultaneous dissipativity of Jacobi matrices (and of constructing a constricting norm).

## 2. Conditions of Simultaneous Dissipativity of Operators

### 2.1. Some General Results

In the present section some conditions of simultaneous dissipativity of the operators will be considered (see definition 7).

A definition of a simultaneous dissipativity can be generalized in such a way.

**Definition 7'.** A family of linear operators  $\{A_\alpha\}$  is called *simultaneously stable dissipative* if there exists a norm with respect to which all operators  $A_\alpha$  are stable dissipative.

**Lemma 6.** *Let the space  $\mathbb{R}^n$  be expanded into direct sum of subspaces  $E_i$  ( $i = 1, \dots, k$ ) and each of them is invariant with respect to all operators of the family  $\{A_\alpha\}$ . Further on, let restriction of the family  $\{A_\alpha\}$  on any  $E_i$  be simultaneously (simultaneously stable) dissipative. Then  $\{A_\alpha\}$  is simultaneously (simultaneously stable) dissipative.*

**Proof.** Let  $\|\cdot\|_i$  ( $i = 1, \dots, k$ ) be the norms in  $E_i$  in which the restrictions of  $\{A_\alpha\}$  on  $E_i$  are simultaneously (simultaneously stable) dissipative. Define the norm in  $\mathbb{R}^n$  in this way:

$$\|x\| = \sum_{i=1}^k \|x_i\|_i,$$

where  $x = \sum_{i=1}^k x_i$  with  $x_i \in E_i$  ( $i = 1, \dots, k$ ).

In this norm all operators  $A_\alpha$  are simultaneously (simultaneously stable) dissipative. The lemma is proved.

It is known [7] that for one operator the norm with respect to which it is dissipative exists if and only if the spectrum of the operator lies in the closed left half-plane and the boundary part is diagonalizable (i. e. Jordan boxes corresponding to pure imaginary, including zero ones, eigenvalues are diagonal). The norm, with respect to which the operator is stable dissipative, exists if and only if the spectrum of the operator lies in the open left half-plane.

Several stable dissipative (in their own norms) operators not necessarily are simultaneously dissipative. To demonstrate that, consider operators represented by the matrices

$$A_1 = \begin{pmatrix} -1 & 3 \\ 0 & -1 \end{pmatrix}; \quad A_2 = \begin{pmatrix} -1 & 0 \\ 3 & -1 \end{pmatrix}.$$

Each of them is stable dissipative in its norm (due to the location of the spectrum). But

$$A_1 + A_2 = \begin{pmatrix} -2 & 3 \\ 3 & -2 \end{pmatrix}.$$

The spectrum of the operator  $(A_1 + A_2)$  contains the point  $\lambda = 1$  which does not belong to the closed left half-plane. Thus, the operator  $(A_1 + A_2)$  is not dissipative in any norm. By lemma 6 the operators  $A_1$  and  $A_2$  are not simultaneously dissipative.

The problem to find out necessary and sufficient conditions of simultaneous dissipativity of an arbitrary (even finite) family of operators seems to be very difficult. Nevertheless, one can obtain some sufficient conditions imposing different constraints on the operators. Obtain the sufficient condition of simultaneous dissipativity of the family generating a solvable Lee algebra. Remind [13] that a family of matrices generates solvable Lee algebra if and only if all elements of this family are simultaneously reducible to triangular form (generally speaking in complex basis).

**Theorem 6.** *Let the family  $\{A_\alpha\}$  be compact and generate solvable Lee algebra, and the spectrum of each operator  $A_\alpha$  lies in the open left half-plane. Then  $\{A_\alpha\}$  is simultaneously stable dissipative.*

**Proof.** First consider the case of complex space  $\mathbb{C}^n$ . Consider matrices of the operators  $A_\alpha$  in the basis where they are of triangular form.

Let each matrix  $A_\alpha$  have the form

$$A_\alpha = \begin{pmatrix} \lambda_1^{(\alpha)} & 0 & 0 & \cdots & 0 & 0 \\ \mu_{21}^{(\alpha)} & \lambda_2^{(\alpha)} & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ \mu_{n1}^{(\alpha)} & \mu_{n2}^{(\alpha)} & \mu_{n3}^{(\alpha)} & \cdots & \mu_{n,(n-1)}^{(\alpha)} & \lambda_n^{(\alpha)} \end{pmatrix}.$$

Show the existence of such a set of positive numbers  $\{c_k\}_{k=1}^n$  that all  $A_\alpha$  are stable dissipative in the norm

$$\|z\| = \max_{1 \leq k \leq n} \frac{|z_k|}{c_k} \tag{16}$$

(here  $z_k$  is the  $k$ -th coordinate of the vector  $z$  in the given basis), whose unit ball is the polycylinder

$$|z_k| \leq c_k \quad (k = 1, \dots, n). \tag{17}$$

If  $\{e_k\}_{k=1}^n$  is the considered basis, then, evidently, norm (16) coincide with the  $l^\infty$ -norm with respect to the basis  $\{c_k/e_k\}_{i,j=1}^n$  in the norm (16):

$$\operatorname{Re} a_{ii} + \sum_{j \neq i} \frac{c_j}{c_i} |a_{ij}| < 0 \quad (i = 1, \dots, n). \tag{18}$$

For the matrices  $A_\alpha$  the conditions (18) look like this:

$$\left\{ \begin{array}{l} \operatorname{Re} \lambda_1^{(\alpha)} < 0; \\ \operatorname{Re} \lambda_2^{(\alpha)} + \frac{c_1}{c_2} |\mu_{21}^{(\alpha)}| < 0; \\ \dots \\ \operatorname{Re} \lambda_n^{(\alpha)} + \frac{c_1}{c_n} |\mu_{n1}^{(\alpha)}| + \dots + \frac{c_{n-1}}{c_n} |\mu_{n,(n-1)}^{(\alpha)}| < 0. \end{array} \right. \tag{19}$$

Suppose  $\mu = \sup_{\alpha, k \neq l} |\mu_{kl}^{(\alpha)}|$ ;  $\lambda = -\sup_{\alpha, k} \operatorname{Re} \lambda_k^{(\alpha)}$ . From the conditions of the theorem follows that  $0 < \lambda < +\infty$ ,  $0 < \mu < +\infty$ . To fulfil (19) for all  $A_\alpha$ , it is sufficient that the inequalities

$$(c_1 + \dots + c_{k-1})\mu < c_k \lambda \quad (k = 1, \dots, n); \quad c_1 > 0 \tag{20}$$

be satisfied.

Show the solvability of system (20). Let  $c_1 = 1$ . Choose the others  $c_k$  so that

$$c_2 > \mu/\lambda; \quad c_3 > (1 + c_2)\mu/\lambda; \quad \dots;$$

$$c_n > (1 + c_2 + \dots + c_{n-1})\mu/\lambda.$$

Then the inequalities (20) are satisfied, i. e. all operators  $A_\alpha$  are stable dissipative in the norm (14).

Let now operators  $A_\alpha$  act in the space  $\mathbb{R}^n$ . In usual way complexify  $\mathbb{R}^n$  and the family  $A_\alpha$ . Then, as it has been described above, construct a cylinder (17). Intersection of (17) with the initial space  $\mathbb{R}^n$  produce a ball of the norm in which all  $A_\alpha$  are stable dissipative. The theorem is proved.

If instead of stable dissipative operators one considers dissipative operators, then the analog of theorem 6 is not true, starting from real dimension 4. Let

$$A_1 = \begin{pmatrix} i & 1 \\ 0 & 2i \end{pmatrix}; \quad A_2 = \begin{pmatrix} 2i & 1 \\ 0 & i \end{pmatrix}.$$

Each of the operators  $A_{1,2}$  is dissipative in its norm. The finite family is compact, the matrices  $A_1$  and  $A_2$  generate solvable Lee algebra. Nevertheless

$$A_1 + A_2 = \begin{pmatrix} 3i & 2 \\ 0 & 3i \end{pmatrix}.$$

The only eigenvalue of the operator  $(A_1 + A_2)$  is pure imaginary, with the matrix of this operator representing (up to a constant factor) non-trivial Jordan box. That means it is not dissipative in any norm, i. e.  $A_1$  and  $A_2$  are not simultaneously dissipative. To obtain a real example, one has to make the matrices  $A_1$  and  $A_2$  real:

$$A_1^R = \begin{pmatrix} 0 & -1 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & -2 \\ 0 & 0 & 2 & 0 \end{pmatrix}; \quad A_2^R = \begin{pmatrix} 0 & -2 & 1 & 0 \\ 2 & 0 & 0 & 1 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 \end{pmatrix}.$$

To keep true the statement about simultaneous dissipativity for nonstable dissipative operators, it is sufficient to strengthen the requirement of solvability up to nilpotency. Remind [13] that for each linear operator  $A$  in the space  $E$  the operator  $\text{ad } A$  in  $L(E)$  is defined:

$$(\text{ad } A)B = AB - BA.$$

The family  $\{A_\alpha\}$  generates the nilpotent Lee algebra if and only if there exists such a number  $m \in \mathbb{N}$  that for any set of  $\{A_{\alpha_k}\}_{k=1}^m$  (among the elements of which there may be the same ones) and for all  $\alpha$ :

$$\prod_{k=1}^m (\text{ad } A_{\alpha_k})A_\alpha = 0. \quad (21)$$

Nilpotent Lee algebra is always solvable. Commutative Lee algebra is nilpotent (for it  $m = 1$ ) and solvable.

**Theorem 7.** *Let the family  $\{A_k\}$  be finite and generate nilpotent Lee algebra, and for each operator  $A_k$  exist a norm with respect to which it is dissipative. Then  $\{A_k\}$  is simultaneously dissipative.*

**Proof.** Without loss of generality one can assume that among the operators  $A_k$  there are no scalar ones (if  $A = aI$ , where  $\operatorname{Re} a \leq 0$ , then  $A$  is dissipative in any norm) and exists at least one operator (denote it  $A_1$ ), among eigenvalues of which there are pure imaginary (otherwise we are under the conditions of theorem 6).

First assume that  $A_k$  operates in  $\mathbb{C}^n$ . We prove the theorem by induction on dimension of space. In dimension 1 the statement of the theorem is trivial. Show that one can expand all the space  $\mathbb{C}^n$  into a direct sum of two non-trivial subspaces invariant with respect to all  $A_k$ . Since in both of them the conditions of the theorem (for corresponding restrictions of  $\{A_k\}$ ) are satisfied, then to complete the proof one has to use lemma 6.

Let  $\lambda$  be an imaginary eigenvalue of  $A_1$ ;  $E'$  be the corresponding to  $\lambda$  eigen-subspace (by virtue of diagonalizability of boundary part of  $A_1$  it coincides with whole corresponding root subspace);  $E''$  be the sum of root subspaces corresponding to all the others eigenvalues of  $A_1$ . Evidently,  $\mathbb{C}^n = E' \oplus E''$  (the sign  $\oplus$  means direct sum);  $E' \neq \mathbb{C}^n$ , otherwise the operator  $A_1$  is scalar. Show the invariance of  $E'$  and  $E''$  with respect to all  $A_k$ .

Let  $x \in E'$ . Then

$$A_1x = \lambda x.$$

On the other hand, in accordance with (21) there exists such  $m \in \mathbb{N}$  that  $(\operatorname{ad} A_1)^m A_k = 0$  for all  $k$  and

$$(A_1 - \lambda I)^m A_k x = 0.$$

A more general fact is true: if  $Ax = 0$  and  $(\operatorname{ad} A)^m B = 0$ , then  $A^m Bx = 0$ . For  $m = 0$  the fact is obvious. Let that be true for  $m = r$ . Assume

$$Ax = 0; (\operatorname{ad} A)^{r+1} B = 0.$$

Then  $(\operatorname{ad} A)^r (\operatorname{ad} A) B = 0$ , and according to the inductive hypothesis

$$A^r (\operatorname{ad} A) Bx = 0.$$

But  $A^{r+1} Bx = A^r (BAx + (\operatorname{ad} A) Bx)$ , i. e.  $A^{r+1} Bx = 0$ , as was to be proved.

As a consequence of coincidence of  $E'$  with the whole root subspace, corresponding to  $\lambda$ , we have:

$$(A_1 - \lambda I) A_k x = 0,$$

i. e.  $A_1 A_k x = \lambda A_k x, A_k x \in E'$ .

Show now the invariance of  $E''$ . Let  $\{e_j\}_{j=1}^n$  be the Jordan basis of the operator  $A_1$  with  $E'$  being corresponded to the vectors  $\{e_j\}_{j=j_1}^{j_2}$ . One has to show that for any  $j$  less then  $j_1$  or more then  $j_2$  the coordinates of  $A_k e_j$  with the numbers from  $j_1$  to  $j_2$  with respect to the assigned basis are equal to zero. Let it be not so and exist such  $j'$  that  $e_{j'} \in E''$ , but the  $j_1$ -th coordinate ( $j_1 \leq j_0 \leq j_2$ ) of the vector  $A_k e_{j'}$  is  $a \neq 0$ . Write it like this:

$$A_k e_{j'} = \dots = a e_{j_0}.$$

Let  $e_j$  be an eigenvector of  $A_1$  corresponding to the eigenvalue  $\mu \neq \lambda$ . Then

$$(\operatorname{ad} A_1) A_k e_{j'} = A_1 (\dots + a e_{j_0}) - \mu (\dots + a e_{j_0}) = \dots + (\lambda - \mu) a e_{j_0}.$$

Verify that

$$(\operatorname{ad} A_1)^m A_k e_{j'} = \dots + (\lambda - \mu)^m a e_{j_0}.$$

For  $m = 0$  it is obvious. Let it be satisfied for  $m = r$ . Then

$$\begin{aligned} (\operatorname{ad} A_1)^{r+1} A_k e_{j'} &= (\operatorname{ad} A_1)(\operatorname{ad} A_1)^r A_k e_{j'} = A_1(\operatorname{ad} A_1)^r A_k e_{j'} - (\operatorname{ad} A_1)^r A_k A_1 e_{j'} = \\ &= A_1(\dots + (\lambda - \mu)^r a e_{j_0}) - \mu(\operatorname{ad} A_1)^r A_k e_{j'} = \dots + (\lambda - \mu)^r a \lambda e_{j_0} - \mu(\lambda - \mu)^r a e_{j_0} = \\ &= \dots + (\lambda - \mu)^{r+1} a e_{j_0}, \end{aligned}$$

i. e. that is true also for  $m = r + 1$ , and, hence, for all  $m \in N$ .

Thus,

$$(\operatorname{ad} A_1)^m A_k e_{j'} = \dots + (\lambda - \mu)^m a e_{j_0} \neq 0$$

for any  $m \in N$ , which contradicts (21).

Let now  $e_{j'}$  be a root (but not eigen) vector, corresponding to the eigenvalue  $\mu$ , with the  $j_0$ -th coordinate of the vector  $A_k e_{j'-1}$  equal to 0. Then

$$(\operatorname{ad} A_1) A_k e_{j'} = A_1(\dots + a e_{j_0}) - A_k(e_{j'-1} + \mu e_{j'}) = \dots + (\lambda - \mu) a e_{j_0}.$$

Analogously

$$(\operatorname{ad} A_1)^m A_k a_{j'} = \dots + (\lambda - \mu)^m a e_{j_0} \neq 0$$

for any  $m \in N$ , which contradicts (21).

Since the sequence of basis vectors belonging to the root subspace begins with the eigenvector, the required statement for complex space is proved.

The transfer onto the case of real space can be done in the same way as in the proof of theorem 6 (the ball of corresponding norm in the complexified  $\mathbb{R}^n$  intersects with  $\mathbb{R}^n$ ). The theorem is proved.

From theorems 6 and 7 follows, in particular, that a finite (compact) commutative family consisting of operators dissipative (stable dissipative) in their own norms is simultaneously dissipative (simultaneously stable dissipative).

## 2.2. The Mass Action Law and Dissipative Mechanisms

Some constructive conditions of simultaneous dissipativity can be obtained for finite families of operators of rank 1. The problem of the absence of wrapping effect in the system constructed in accordance with the *Mass Action Law* (MAL) is reduced to the problem on simultaneous dissipativity of such operators.

MAL systems appear from mathematical description of systems of chemical and biological kinetics and in some other problems. To the considered process is assigned an algebraic object, called *reaction mechanism* and having the form:



Speaking in terms of chemical kinetics, the reaction mechanism is a list of stoichiometric equations of elementary reactions (22). In this case  $A_1, \dots, A_n$  are the substances taking part in the reaction;  $\alpha_{ri}, \beta_{ri}$  are the non-negative integers called stoichiometric coefficients and showing in what amount the particles of  $A_i$  enter into the  $r$ -th elementary reaction as the initial substance ( $\alpha_{ri}$ ) or product ( $\beta_{ri}$ ). The following notations are accepted:  $\gamma_{ri} = \beta_{ri} - \alpha_{ri}$ ,  $\gamma_r$  is the vector with the components  $\gamma_{ri}$  ( $i = 1, \dots, n$ ) — so-called stoichiometric vector of the  $r$ -th elementary reaction.

In accordance with MAL [14,15], to the mechanism (22) corresponds the following system of ordinary differential equations:

$$\frac{dc_i}{dt} = \sum_{r=1}^d \gamma_{ri} w_r \quad (23)$$

where  $c_i(t)$  is the concentration of substance  $A_i$  at the moment of time  $t \geq 0$ ,

$$w_r = k_r(t) \prod_{j=1}^n c_j^{\alpha_{rj}}$$

is the rate of the  $r$ -th elementary reaction, continuously depending on time. In particular, if reaction proceeds under constant external conditions, then  $k_r(t) = \text{const}$  ( $r = 1, \dots, d$ ) and  $k_r$  is called rateant of the  $r$ -th elementary reaction. In the latter case (23) represents an autonomous system with polynomial right sides.

Let  $L$  be a linear hull of the family  $\{\gamma_r\}_{r=1}^d$ . If  $L \neq \mathbb{R}^n$ , then there exist such  $a_i$  ( $i = 1, \dots, n$ ), not all equal to zero, that for all  $r = 1, \dots, d$  the equalities

$$\sum_{i=1}^n a_i \gamma_{ri} = 0$$

are satisfied, from which for system (23) follows that

$$\sum_{i=1}^n a_i c_i(t) = \text{const}. \quad (24)$$

Relationships (24) are called stoichiometric conservation laws. If all  $a_i$  are positive, then the corresponding stoichiometric law is called *the positive conservation law* [15]. In MAL positive conservation laws takes place rather often (but not always).

As it is known [15], balance polyhedrons are intersections of affine subspaces of the form  $(L + c)$ , where  $c$  is a constant vector, with a cone of non-negative vectors (first orthant) in  $\mathbb{R}^n$ . Balance polyhedrons represent positive invariant with respect to (23) convex sets (one can find the proof of their positive invariance in [15]). If there exists at least one of positive conservation law, they are compact.

The question arises: under what conditions does the norm exist in  $\mathbb{R}^n$  according to which the system (23) is constricting in all balance polyhedrons and independent of rate constants?

**Definition 8.** Mechanism (22) is called *dissipative*, if for system (23) there exists a norm, constricting in all balance polyhedrons irrespective of rate constants (in other words, the constricting norm depends on the mechanism only).

We use the notation  $M_{ri}$  for the operator in  $\mathbb{R}^n$ , represented by the matrix, in the  $i$ -th column of which there are components of the vector  $\gamma_r$ , and on other places — zeros. The subspace  $L$  is invariant with respect to all  $M_{ri}$  [15]. The notation  $M'_{ri}$  stays for restriction of  $M_{ri}$  on  $L$ .

**Theorem 8.** *Let for mechanism (22) exist at least one positive conservation law. This mechanism is dissipative if and only if the family  $\{M'_{ri} \mid \alpha_{ri} > 0\}$  is simultaneously dissipative.*

**Proof. Sufficiency.** It is known [15] that the Jacobi matrix  $J_c$  of system (22) at the point  $c$ , whose coordinates are positive, has the form

$$J_c = \sum_{\alpha_{ri} > 0} \alpha_{ri} \frac{w_r}{c_i} M_{ri}. \quad (25)$$

Matrices  $J_c$  belong to the convex cone produced by the family  $\{M_{ri} | \alpha_{ri} > 0\}$ . Besides, the difference of any two solutions (23) from one balance polyhedron belongs to the subspace  $L$ . Under the conditions of lemma 3 and theorem 2 obtain the existence of constricting norm in the subspace  $L$ . It can be expanded onto all  $\mathbb{R}^n$ .

*Necessity.* Matrices  $M_{ri}$  ( $\alpha_{ri} > 0$ ) belong to the closure of the family of matrices  $J_c$  for arbitrary non-negative vectors  $c$  and rate constants  $k_r$ . To prove this, first let consider the case when  $c_j$  ( $j = 1, \dots, n$ ) and  $k_r$  are fixed and all  $k_l$  ( $l \neq r$ ) tend to zero. In the limit in (25) only the sum for given  $r$  is left. Further on, fix all  $c_j > 0$  ( $j \neq i$ ) and let  $c_i$  tend to zero, changing  $k_r$  so that the equality  $\alpha_{ri}w_r/c_i = 1$  holds true. Then all the terms except one tend to zero and in the limit we obtain  $M_{ri}$ .

Thus, the matrices  $M_{ri}$  ( $\alpha_{ri} > 0$ ) belong to the closure of the family of restrictions of the matrices  $J_c$  on the subspace  $L$ . Hence, according to lemma 3 and theorem 2 the necessity follows. The theorem is proved.

Note that matrices  $M_{ri}$  represent matrix-columns (in each matrix there is only one non-zero column) and that means that the rank of each of them is equal to unity. We come to the problem of simultaneous dissipativity of the finite family of operators of rank 1.

Note that dissipative mechanisms of reactions were studied in details in [12]. In particular, some classes of dissipative mechanisms are pointed out and all dissipative mechanisms for  $n = 3$ ,  $\sum_{i=1}^3 \alpha_{ri} \leq 3$ ,  $\sum_{i=1}^3 \beta_{ri} \leq 3$  ( $r = 1, \dots, d$ ),  $c_1 + c_2 + c_3 = \text{const}$  enumerated.

In the next subsection are obtained necessary and sufficient conditions of simultaneous dissipativity of the operators of rank 1 in  $\mathbb{R}^2$  (corresponding to the case  $\dim L = 2$ ) and some sufficient conditions of simultaneous dissipativity of matrix-columns.

### 2.3. Constructive Conditions of Simultaneous Dissipativity of One-Dimensional Operators

Before consideration of simultaneous dissipativity of operators of rank 1, find out what can be said about dissipativity of one such operator. From necessary and sufficient conditions (see the paragraph after lemma 6) follows that the norm in which the given operator of rank 1 is dissipative exists if and only if it has a negative eigenvalue.

Positive semi-trajectories of system (3) corresponding to the initial condition  $x(0) = x_0$  are in this case rectilinear segments parallel to the image of  $A$  and connecting  $x_0$  with  $\text{Ker } A$ . Operator  $A$  of rank 1 is dissipative in the given norm if and only if for any point  $x$  ( $\|x\| = 1$ ) there exists such  $\varepsilon > 0$  that  $\|x + \varepsilon Ax\| \leq 1$ . It means that the negative number belongs to the spectrum of  $A$ , and the image of  $A$  is orthogonal to its kernel (in the given norm, the subspace  $E_2$  is orthogonal to  $E_1$  if  $\|x + y\| \geq \|x\|$  for any  $x \in E_1, y \in E_2$  [7]).

Let now be given a family  $\{M_k\}_{k=1}^m$  of operators of rank 1 in  $\mathbb{R}^n$ . Each of them can be represented in the form  $(\cdot; \psi_k)\varphi_k$ , i. e.  $M_k x = (x; \psi_k)\varphi_k$  where  $(\cdot; \cdot)$  is the standard scalar product in  $\mathbb{R}^n$ . The vectors  $\varphi_k$  and  $\psi_k$  are determined by the operator  $M_k$  unambiguously (up to scalar factors). Let  $\lambda_k = (\varphi_k; \psi_k)$ , i. e.  $\lambda_k$  is an eigenvalue of  $M_k$  (either it is the only non-zero eigenvalue, or 0, if the operator  $M_k$  is nilpotent). As it has already been mentioned, for simultaneous dissipativity of  $\{M_k\}$  the conditions

$$\lambda_k < 0 \quad (k = 1, \dots, m) \quad (26)$$

are necessary.

Assign to each operator  $M_k$  the projector  $P_k$  projecting parallel to the image of  $M_k$  on the kernel of  $M_k$ . It is easy to see that  $P_k = I - M_k/\lambda_k$ . By virtue of the mentioned above condition

of dissipativity of the operator of rank 1 in the given norm the operator  $M_k$  is dissipative in some norm if and only if  $P_k$  is constriction in this norm.

All  $P_k$  can be constrictions in one norm if and only if all products of the form  $\prod_{j=1}^q P_{k_j}$  ( $q \in N$  is arbitrary;  $k_j \in \{1, \dots, m\}$  and they are not necessarily different) are jointly bounded. We come to the following conclusion.

**Lemma 7.** *The family  $\{M_k\}_{k=1}^m$  of operators of rank 1 is simultaneously dissipative if and only if the conditions (26) are satisfied and all products of the form  $\prod_{j=1}^q P_{k_j}$  ( $q \in N$  is arbitrary;  $k_j \in \{1, \dots, m\}$  and they are not necessarily different) are jointly bounded. As a constricting norm one can take*

$$\|x\| = \sup_{q \in N, 1 \leq k_j \leq m} \{ \|x\|_0, \|(\prod_{j=1}^q P_{k_j})x\|_0 \}, \tag{27}$$

where  $\|\cdot\|_0$  is any norm in  $\mathbb{R}^n$ .

**Proof.** All statements of the lemma, except the latter, follow immediately from the above reasoning. Further on, if all products  $\prod_{j=1}^q P_{k_j}$  are jointly bounded, then

$$\sup_{q \in N, 1 \leq k_j \leq m} \{ \|x\|_0, \|(\prod_{j=1}^q P_{k_j})x\|_0 \} < \infty$$

for each  $x \in \mathbb{R}^n$ . This expression possesses all properties of norm and all operators  $P_k$  are constrictions in such norm, i. e. all  $M_k$  are simultaneously dissipative. The lemma is proved.

From lemma 7 follows a simple consequence.

**Corollary 1.** *If all  $\varphi_k$  are collinear (images of  $M_k$  coincide) or all  $\psi_k$  are collinear (kernels of  $M_k$  coincide) and  $(\varphi_k; \psi_k) < 0$  for all  $k = 1, \dots, m$ , then the operators  $M_k$  ( $k = 1, \dots, m$ ) are simultaneously dissipative. As corresponding constricting norm one can take*

$$\sup_{q \in N, 1 \leq k_j \leq m} \{ \|x\|_0, \|P_k x\|_0 \}.$$

To demonstrate this, it is sufficient to note that in these cases

$$\prod_{j=1}^q P_{k_j} = P_{k_1}$$

or

$$\prod_{j=1}^q P_{k_j} = P_{k_q},$$

respectively.

**Remark 10.** If not all  $\varphi_k$  are collinear, then as a norm in lemma 7 one can take

$$\sup_{q \in N, 1 \leq k_j \leq m} \|(\prod_{j=1}^q P_{k_j})x\|_0. \tag{28}$$

The criterion established in lemma 7 is not constructive. Constructive criteria of simultaneous dissipativity of finite family of operators of rank 1 in  $\mathbb{R}^n$  have been obtained only at  $n = 2$  (for arbitrary  $n$  there exist sufficient conditions for one class of operators; they are given at the end of the section). Pass to the consideration of the case  $n = 2$ .

Consider the family  $\{M_k\}_{k=1}^m$  of the operators of rank 1 in  $\mathbb{R}^2$ . As before, represent each operator  $M_k$  in the form  $(\cdot; \psi_k)\varphi_k$ . Let first  $m = 2$ .

**Lemma 8.** *The operators  $M_1 = (\cdot; \psi_1)\varphi_1$  and  $M_2 = (\cdot; \psi_2)\varphi_2$  are simultaneously dissipative in  $\mathbb{R}^2$  if and only if the condition*

$$\left| \frac{(\varphi_1; \psi_2) \cdot (\varphi_2; \psi_1)}{(\varphi_1; \psi_1) \cdot (\varphi_2; \psi_2)} \right| \leq 1 \quad (29)$$

is satisfied together with the conditions

$$(\varphi_1; \psi_1) < 0; \quad (\varphi_2; \psi_2) < 0.$$

As a corresponding constricting norm one can take

$$\|x\| = \max\{\|x\|_0, \|P_1x\|_0, \|P_2x\|_0, \|P_1P_2x\|_0, \|P_2P_1x\|_0\}. \quad (30)$$

**Proof.** In  $\mathbb{R}^2$  the projectors  $P_1$  and  $P_2$  have rank 1 and are represented in the form

$$P_1 = (\cdot; \eta_1)\chi_1; \quad P_2 = (\cdot; \eta_2)\chi_2,$$

where  $\eta_1, \eta_2, \chi_1, \chi_2$  are some vectors in  $\mathbb{R}^2$ .

The operators  $\prod_{j=1}^q P_{k_j}$  are bounded when the spectrum of the operator  $(P_1P_2)$  lies on the segment  $[-1; 1]$ :

$$|(\chi_1; \eta_2) \cdot (\chi_2; \eta_1)| \leq 1. \quad (31)$$

In a standard orthonormalized basis  $P_k$  acts like this:

$$P_k x = \frac{1}{(\varphi_k; \psi_k)} \cdot \left( \left( \begin{array}{c} x^{(1)} \\ x^{(2)} \end{array} \right); \left( \begin{array}{c} \varphi_k^{(2)} \\ -\varphi_k^{(1)} \end{array} \right) \right) \cdot \left( \begin{array}{c} \psi_k^{(2)} \\ -\psi_k^{(1)} \end{array} \right),$$

where

$$\left( \begin{array}{c} a^{(1)} \\ a^{(2)} \end{array} \right)$$

denotes the vector with the coordinates  $a^{(1)}$  and  $a^{(2)}$ . Hence

$$(\eta_1; \chi_2) = \frac{(\varphi_1; \psi_2)}{(\varphi_2; \psi_2)},$$

$$(\chi_1; \eta_2) = \frac{(\varphi_2; \psi_1)}{(\varphi_1; \psi_1)},$$

i. e. condition (31) takes the form (29).

To complete the proof, use lemma 7. To check a possibility of choosing corresponding norm in the form (30), note that

$$(P_1P_2)^r P_1 = (\eta_1; \chi_2)^r \cdot (\eta_2; \chi_1)^r \cdot P_1;$$

$$(P_2P_1)^r P_2 = (\eta_1; \chi_2)^r \cdot (\eta_2; \chi_1)^r \cdot P_2$$

for any  $r \in N$ . It means that with the account of (31), in (27) one can restrict oneself to finite number of products. The lemma is proved.

**Remark 11.** If  $\varphi_1$  and  $\varphi_2$  are non-collinear, then as required norm we can take

$$\max\{\|P_1x\|_0, \|P_2x\|_0, \|P_1P_2x\|_0, \|P_2P_1x\|_0\}.$$

This follows from remark 10. Then the ball of the norm is determined by the inequalities

$$|(x; \eta_1)| \leq \min\left\{\frac{1}{\|\chi_1\|_0}, \frac{1}{|(\chi_1; \eta_2)| \cdot \|\chi_2\|_0}\right\};$$

$$|(x; \eta_2)| \leq \min\left\{\frac{1}{\|\chi_2\|_0}, \frac{1}{|(\chi_2; \eta_1)| \cdot \|\chi_1\|_0}\right\},$$

i. e. it is parallelogram.

Also notice that for simultaneous dissipativity of a family the dissipativity of each operator from convex hull of the family is insufficient. To see this, consider the operators represented by the matrices

$$M_1 = \begin{pmatrix} -1 & 1 \\ 0 & 0 \end{pmatrix}; \quad M_2 = \begin{pmatrix} 0 & 0 \\ -2 & -1 \end{pmatrix}.$$

Each of them is dissipative in its norm. It is easy to show that spectrum of any non-trivial convex combination of  $M_1$  and  $M_2$  lies in open left half-plane. Nevertheless

$$\frac{(\varphi_1; \psi_2) \cdot (\varphi_2; \psi_1)}{(\varphi_1; \psi_1) \cdot (\varphi_2; \psi_2)} = -2,$$

i. e. condition (29) is not satisfied.

Reasoning like in proof of lemma 8, it is easy to obtain a criterion of a simultaneous dissipativity for arbitrary  $m$ . The result is a set of conditions of the form

$$(\varphi_k; \psi_k) < 0 \quad (k = 1, \dots, m); \quad (32)$$

$$\left| \frac{(\varphi_{k_1}; \psi_{k_2}) \cdot (\varphi_{k_2}; \psi_{k_3}) \cdot \dots \cdot (\varphi_{k_q}; \psi_{k_1})}{(\varphi_{k_1}; \psi_{k_1}) \cdot (\varphi_{k_2}; \psi_{k_2}) \cdot \dots \cdot (\varphi_{k_q}; \psi_{k_q})} \right| \leq 1, \quad (33)$$

where  $\{k_j\}_{j=1}^q$  is a set of different numbers from 1 to  $m$ , and inequalities (33) holds for all such sets. The number of conditions has the order  $O((m-1)!)$  and for any large  $m$  testing of these conditions becomes unrealizable. It turns out, however, that among inequalities (33) there are dependent ones and the number of conditions can be reduced.

**Theorem 9.** *Let the vectors  $\psi_k$  ( $k = 1, \dots, m$ ) lie in one half-plane clockwise. Then the family of operators  $\{M_k\}_{k=1}^m$ , where  $M_k = (\cdot; \psi_k)\varphi_k$  is simultaneously dissipative if and only if the vectors  $\varphi_k$  ( $k = 1, \dots, m$ ) lie in one half-plane clockwise and the conditions (32) and the followings ((34), (35)) are satisfied:*

$$\left| \frac{(\varphi_k; \psi_{k+1}) \cdot (\varphi_{k+1}; \psi_k)}{(\varphi_k; \psi_k) \cdot (\varphi_{k+1}; \psi_{k+1})} \right| \leq 1 \quad (k = 1, \dots, m \text{ with } \varphi_{m+1} = -\varphi_1; \psi_{m+1} = -\psi_1); \quad (34)$$

$$\left\{ \begin{array}{l} \left| \frac{(\varphi_1; \psi_2) \cdot (\varphi_2; \psi_3) \cdot \dots \cdot (\varphi_m; \psi_1)}{(\varphi_1; \psi_1) \cdot (\varphi_2; \psi_2) \cdot \dots \cdot (\varphi_m; \psi_m)} \right| \leq 1; \\ \left| \frac{(\varphi_1; \psi_m) \cdot (\varphi_m; \psi_{m-1}) \cdot \dots \cdot (\varphi_2; \psi_1)}{(\varphi_1; \psi_1) \cdot (\varphi_2; \psi_2) \cdot \dots \cdot (\varphi_m; \psi_m)} \right| \leq 1. \end{array} \right. \quad (35)$$

The corresponding norm can be chosen polyhedral (a norm, whose ball is polygon).

**Proof. Necessity.** Let  $l_k$  be the kernels of the operators  $M_k$  (i. e. straight lines orthogonal to  $\psi_k$ ). Straight lines  $l_k$  divide the plane into  $2m$  sectors. If among the vectors  $\psi_k$  there are collinear, then some sectors are singular, but this does not change the further reasoning. In each sector  $G$  and for each  $p \in \{1, \dots, m\}$

$$\text{sign}(x_1; \psi_p) = \text{sign}(x_2; \psi_p)$$

for all  $x_1 \in \text{int } G$ ,  $x_2 \in \text{int } G$ .

Let  $G_r$  be a sector lying between corresponding rays of straight lines  $l_r$  and  $l_{r+1}$  (where  $l_{m+1} = l_1$ ). It is enough to consider the sectors  $\{G_k\}_{k=1}^m$  into which one half-plane is divided, since for sectors lying in vertical angles to  $G_k$  the reasons are the same.

Note that by inequality (32) for each operator  $M_k$  the projector  $P_k$  is determined, which operates in each sector  $G_k$  as a projector in the direction  $v_{kr} = \text{sign}(x; \psi_k) \cdot \varphi_k$  ( $x \in G_r$ ) onto the straight line  $l_k$ .

A norm with respect to which all  $M_k$  are dissipative exists if and only if there exist a convex body  $Q$  symmetrical with respect to 0 and positively invariant with respect to all systems of the following form

$$\frac{dx}{dt} = \sum_{k=1}^m h_k(t)(x; \psi_k)\varphi_k, \quad (36)$$

where  $h_k(t)$  is any function piecewise continuous and non-negative for  $t \geq 0$ . The sufficiency is evident (suppose  $h_k(t) \equiv 1$ ,  $h_j(t) \equiv 0$  for  $j \neq k$  and come to dissipativity of  $M_k$  with respect to  $Q$ ). To prove the necessity, it is sufficient to make an estimation analogous to that made in the proof of theorem 1:

$$\begin{aligned} \frac{d}{dt} \|x(t)\|_Q &= N_Q(x(t), \sum_{k=1}^m h_k(t)M_k x(t)) \leq \\ &\leq \gamma_Q \left( \sum_{k=1}^m h_k(t)M_k x(t) \right) \cdot \|x(t)\|_Q \leq 0. \end{aligned}$$

Here  $\|\cdot\|_Q$  is a norm whose unit ball is  $Q$ .

Since  $(x; \psi_k)\varphi_k = |(x; \psi_k)| \cdot v_{kr}$  at  $x \in G_r$ , then (36) can be rewritten as follows:

$$\frac{dx}{dt} = \sum_{k=1}^m y_k(t)v_{kr} \quad (37)$$

where  $y_k(t)$  is piecewise continuous and non-negative for  $t \geq 0$ . Thus, it is sufficient to construct such a polygon  $W$  that from each point of its boundary  $\partial W$  all the vectors  $v_{kr}$  are not directed into the exterior of  $W$ . Then one can take

$$Q = \text{co} \{W \cup (-W)\}.$$

Let (37) have at least one unbounded solution, whose positive semi-trajectory lies inside one of sectors. Then (36) has an unbounded solution, i. e. the operators  $M_k$  are not simultaneously dissipative.

The notation  $C\{v_{kr}\}$  is used for a convex cone produced by  $\{v_{kr}\}_{r=1}^m$ .

Let this cone coincide with  $\mathbb{R}^2$  at least in one sector  $G_r$  (i. e. the vectors generating it do not lie in one half-plane). Then as  $y_k(t)$  one can choose such constants that  $v = \sum_{k=1}^m y_k v_{kr} \in G_r$ , and then, drawing a ray from the point  $x_0 \in \text{int} G_r$  in the direction of  $v$ , obtain a positive semi-trajectory of unbounded solution (37) lying inside  $G_r$ .

Thus, for simultaneous dissipativity of  $\{M_k\}$  it is necessary to satisfy the conditions

$$C\{v_{kr}\} \neq \mathbb{R}^2 \quad (k = 1, \dots, m). \tag{38}$$

If  $C\{v_{kr}\}$  in some sector is a half-plane, then it must contain the vertical angle to  $G_r - \hat{G}_r$  (and thus intersect with  $G_r$  only at zero); otherwise (37) has an unbounded solution. For each sector  $G_r$  consider the boundary of the cone  $C\{v_{kr}\}$ . It consists of two directions. Show that for  $G_j$  it is  $v_{jj}, v_{(j+1),j}$ . It is sufficient to show that for  $j = 1$ .

Let  $v_{1,1}$  and  $v_{2,1}$  be collinear and oppositely directed. Then to satisfy (38) it is necessary that the other  $v_{k1}$  lie on one side of the straight line, stretched on  $v_{1,1}$ . But if  $v_{1,1}$  and  $v_{2,1}$  are non-collinear, then all other  $v_{k1}$  can be expanded in terms of the basis  $v_{1,1}, v_{2,1}$ .

Let, for example,  $v_{3,1} = c_1 v_{1,1} + c_2 v_{2,1}$ , and  $v_{3,1}$  be collinear to one of the basis vectors (for example,  $v_{1,1}$ ; the case with  $v_{2,1}$  is considered analogously). Then  $c_2 = 0$ . If  $c_1 > 0$  then  $v_{3,1} \in C\{v_{1,1}, v_{2,1}\}$ . Let  $c_1 < 0$ . Then to satisfy (38) in  $G_1$  it is necessary for  $v_{1,1}$  and  $v_{3,1}$  to be boundary directions in  $C\{v_{k1}\}$ . Since  $v_{k,(l+1)} = v_{kl}$ , if  $k \neq l + 1$ , and  $v_{(l+1),(l+1)} = -v_{l,(l+1)}$ , then the same directions are boundary for  $C\{v_{k2}\}$  as well, otherwise (38) is not satisfied in  $G_2$ . Simultaneously  $v_{2,1} \in C\{v_{k1}\}$ ,  $v_{2,2} = -v_{2,1} \in C\{v_{k2}\}$ . It means  $C\{v_{k1}\}$  and  $\{v_{k2}\}$  represent half-plane whose join is all  $\mathbb{R}^2$ , what is impossible. That means  $c_1 > 0$ .

Let now  $v_{3,1}$  be non-collinear neither to  $v_{1,1}$  nor to  $v_{2,1}$ . If  $c_1 < 0, c_2 < 0$ , then in  $G_1$  (38) is not satisfied. If  $c_1 < 0, c_2 > 0$ , then in  $G_2$  there  $v_{3,2} = c_1 v_{1,2} + (-c_2) v_{2,2}$ , i. e. again (38) is not satisfied. Analogous reasoning hold for the case  $c_1 > 0, c_2 < 0$ , i. e. the only possible case is  $c_1 \geq 0, c_2 \geq 0$  and therefore  $v_{3,1} \in C\{v_{1,1}, v_{2,1}\}$  (where  $C\{x, y\}$  is a convex cone, stretched on the vectors  $x$  and  $y$ ).

The case is left when the directions  $v_{1,1}$  and  $v_{2,1}$  coincide.

Without loss of generality one can assume non-collinearity of  $v_{3,1}$  and  $v_{1,1}$ . Then  $v_{2,2}$  and  $v_{3,2}$  are boundary directions in  $C\{v_{k2}\}$ . Consequently,  $v_{1,1} \in C\{-v_{1,1}, v_{3,1}\}$  i. e. the directions  $v_{3,1}$  and  $v_{1,1}$  coincide contrary to the assumption. It means that if  $v_{rr}$  and  $v_{(r+1),r}$  are co-directed, all  $v_{kr}$  are collinear, i. e. all  $\varphi_k$  are collinear. In this case the directions  $v_{rr}$  and  $v_{(r+1),r}$  are also boundary.

We call the obtained fact *the boundary condition*.

Since all  $\psi_k$  lie clockwise in one half-plane, then it is easy to check that in sector  $G_m$  either all  $(x; \psi_k) \geq 0$  for all  $k$  or  $(x; \psi_k) \leq 0$  for all  $k$ . Thus, by virtue of (32), all  $\varphi_k$  lie in one half-plane. From the boundary condition follows that  $\varphi_k \in C\{\varphi_{k-1}, \varphi_{k+1}\}$ , i.e vectors  $\varphi_k$  are arranged either clockwise, or anti-clockwise.

Let, for example,  $v_{1,1} = \varphi_1$  (the case  $v_{1,1} = -\varphi_1$  is considered analogously). Then  $v_{2,1} = \varphi_2$  lies in the half-plane bounded by the straight line stretched on  $v_1$  and containing  $\hat{S}_1$ . Therefore the direction from  $\varphi_1$  to  $\varphi_2$  in the half-plane containing all  $\varphi_k$  is the same as from  $\psi_1$  to  $\psi_2$ , i. e. clockwise.

The necessity of the other conditions is obvious, since (34) – (35) is simply a part of conditions (33).

*Sufficiency.* Let the family  $\{\varphi_k\}_{k=1}^m$  be arranged clockwise in one half-plane and the conditions (32) and (34) – (35) be satisfied. Assume that among  $\varphi_k$  there are non-collinear vectors, and among  $\psi_k$  there are no collinear ones.

The condition of clockwise arrangement of  $\varphi_k$  in one half-plane means that the angle (counted from  $\varphi_1$  clockwise) between  $\varphi_1$  and the vectors  $\varphi_1, \varphi_2, \dots, \varphi_m, \varphi_{m+1} = -\varphi_1$  monotonically increases from 0 to  $\pi$ . Taking into account that the angle between  $\varphi_{k1}$  and  $(-\varphi_{k2})$  is the angle between  $\varphi_{k1}$  and  $\varphi_{k2}$ , taken with opposite sign, it is easy to conclude that systems  $\{v_{kr}\}$  (in each sector) lie in one half-plane and are arranged clockwise (to avoid exiting from corresponding half-plane we start counting in sector  $G_r$  from  $v_{(r+1),r}$ ).

From conditions (34) follows that in each sector there is a "convex configuration", i. e. there is vector  $x \in G_r$ , representable in the form

$$x = - \sum_{k=1}^m c_k v_{kr},$$

where all  $c_k > 0$ .

It means that if from one point  $\tilde{x} \in \text{int } G_r$  one draws segments  $\bar{a}$  and  $\bar{b}$  in the directions of  $v_{rr}$  and  $v_{(r+1),r}$  up to the crossing with  $l_r$  and  $l_{r+1}$ , respectively, then these segments together with the segments connecting 0 with the point of crossing  $\bar{a}$  with  $l_r$  and  $\bar{b}$  with  $l_{r+1}$ , respectively, form a convex polygon (if  $v_{rr}$  and  $v_{(r+1),r}$  are oppositely directed, it will be a triangle, and if they are non-collinear — a quadrangle; as we have seen before they cannot be co-directed).

Due to the same orientation of  $\{\varphi_k\}$  and  $\{\psi_k\}$  all the other  $v_{kr}$  are directed (from point  $\tilde{x}$ ) into this polygon, i. e. for any cone  $C\{v_{kr}\}$  directions on the straight lines  $v_r$  and  $v_{r+1}$  are boundary.

Fix now the point  $x_0 \in l_1$  ( $x_0 \neq 0$ ) on the boundary ray of sector  $G_1$  (actually, one can begin from any straight line  $l_k$ ; we begin from  $l_1$ ). Due to the boundary condition either direction from  $x_0$  on  $l_2$  goes into sector  $G_1$ , or direction from  $x_0$  on  $l_m$  goes into  $\hat{G}_m$ .

If one and only one of these statements is true, continue moving in the corresponding direction (to the neighboring straight line) till the direction on the neighboring straight line goes into the neighboring sector. In other words, move from  $l_r$  to  $l_{r+1}$  in the direction parallel to  $\varphi_r$ , if this direction goes into sector  $G_r$  (or, into  $\hat{G}_{r-1}$ , respectively). As a polygon  $W$  mentioned after (37) one should take a polygon formed by the segments which we moved along, and the segments of those straight lines on which the movement broke (if exit on the initial ray did not occur, in our case it is a part of  $l_1$  corresponding to  $G_1$ , then it is a segment connecting  $x_0$  with 0, and a segment of that straight line on which the movement broke, connecting the point of breaking with zero; if exit on the initial ray occurred, then it is a segment connecting  $x_0$  with the point of exit).

If both statements are satisfied, then as  $W$  one can take a join of two such polygons formed in moving to both sides from  $x_0$ .

This algorithm is easy to check proceeding from boundary conditions, "convex configuration", and (35) (the latter condition means that if exit on the initial ray occurred in moving in either side, then the point of exit is no farther from the beginning of coordinates than the initial point; in particular, if the point of exit coincides with the initial point, then the formed polygon can be taken as  $W$ ). The ball of the sought for norm is a polygon.

If some of  $\psi_k$  are collinear, then some sectors  $G_k$  are singular. This, however, does not change the results. The reasoning are analogous to the case when among  $\psi_k$  there are no collinear vectors. The only difference here is the following: some straight lines  $l_k$  correspond to several directions  $\{\varphi_j\}_{j=k_0}^{k_1}$ . Then in constructing  $W$  one needs to move along  $\varphi_{k_0}$ .

In the case when all  $\varphi_k$  (or all  $\psi_k$ ) are collinear (see corollary 1), all the same one can regard that  $\{\varphi_k\}$  and  $\{\psi_k\}$  have the same orientation, starting from (32).

Conditions (34)–(35) are satisfied in this case. The norm can be chosen polyhedral, if one chooses a polyhedral norm as  $\|\cdot\|_0$  in (30). The theorem is proved.

**Remark 12.** One can obtain the arrangement of vectors  $\psi_k$  required by the conditions of theorem 9 by renumbering vectors and (if it is necessary) changing signs of some of them.

Thus, the problem of simultaneous dissipativity of a family of operators of rank 1 in  $\mathbb{R}^2$  is solved completely. The number of conditions to be checked now, in contrast to (33), is only of the order  $O(m)$ .

With theorem 9 one can study the MAL mechanism on dissipativity (and, respectively, on the absence of IWE). For example, let the mechanism be

$$\begin{aligned} A_1 &\rightarrow A_2, A_1 \rightarrow A_3, A_2 \rightarrow A_1, A_2 \rightarrow A_3, \\ 3A_2 &\rightarrow A_1 + 2A_3, 2A_1 \rightarrow A_2 + A_3, 2A_2 \rightarrow A_1 + A_3, \\ 2A_3 &\rightarrow A_1 + A_2, 3A_1 \rightarrow A_2 + 2A_3, 3A_2 \rightarrow 2A_1 + A_3, \\ A_1 + A_2 &\rightarrow 2A_3. \end{aligned} \tag{39}$$

This mechanism possesses positive conservation law  $c_1 + c_2 + c_3 = \text{const}$ . The corresponding subspace is the plane

$$c_1 + c_2 + c_3 = 0.$$

Obviously,  $\dim L = 2$ , and one can use theorem 9. Writing matrices  $M'_{ri}$  and using theorem 9, let make sure that mechanism (39) is dissipative. The corresponding norm in the subspace  $L$  has the form

$$\|c\| = |c_1| + |c_2|.$$

It can be expanded onto all  $\mathbb{R}^3$ , for example, in this way:

$$\|c\| = |c_1| + |c_2| + |c_1 + c_2 + c_3|.$$

To complete the section, consider the question of simultaneous dissipativity of the finite family of operators of rank 1 of special form in  $\mathbb{R}^n$  for arbitrary  $n$ . Namely, we consider operators represented by matrix-columns. Let obtain sufficient conditions of simultaneous dissipativity of such operators.

Let the basis  $\{e_k\}_{k=1}^n$  and the norm

$$\|x\| = \sum_{k=1}^n p_k |x_k| \tag{40}$$

be given in  $\mathbb{R}^n$ , where  $p_k > 0$  ( $k = 1, \dots, n$ ),  $x_k$  is the  $k$ -th coordinate of vector  $x$  in the basis  $\{e_k\}$ . Norm (40) coincides with  $l^1$  norm with respect to the basis  $\{e_k/p_k\}$ . Therefore, the necessary and sufficient dissipativity conditions of the operator  $A$  represented by the matrix  $(a_{ij})_{i,j=1}^n$  according to remark 9 have the form

$$p_i a_{ii} + \sum_{j \neq i} p_j |a_{ji}| \leq 0 \quad (i = 1, \dots, n). \tag{41}$$

Let now there be a family of operators, represented by the matrix-columns  $A_{kl_k}$  ( $k=1, \dots, n$ ;  $l_k=0, \dots, r_k$ ), where  $A_{kl_k}$  is the  $l_k$ -th matrix with non-zero  $k$ -th column:

$$A_{kl_k} = \begin{pmatrix} 0 & \dots & 0 & a_{1k}^{(l_k)} & 0 & \dots & 0 \\ 0 & \dots & 0 & a_{2k}^{(l_k)} & 0 & \dots & 0 \\ & & \dots & \dots & & & \\ 0 & \dots & 0 & a_{nk}^{(l_k)} & 0 & \dots & 0 \end{pmatrix}. \quad (42)$$

Coming from (41), write dissipativity conditions of all operators in norm (40) with some constants  $p_k$ :

$$p_k a_{kk}^{(l_k)} + \sum_{j \neq k} p_j |a_{jk}^{(l_k)}| \leq 0 \quad (k=1, \dots, n; l_k=0, \dots, r_k). \quad (43)$$

**Theorem 10.** *If the system of linear inequalities (43) complemented by the inequalities*

$$p_k > 0 \quad (k=1, \dots, n) \quad (44)$$

*has a solution, then the family of operators represented by matrices (42) is simultaneously dissipative.*

**Proof.** Solvability of the systems (43), (44) means the existence of positive constants  $p_k$  ( $k=1, \dots, n$ ) for which inequalities (43) are satisfied, and that is dissipativity condition of all operators of the family in norm (40). Obviously, in this case the family is dissipative. The theorem is proved.

Thus, for simultaneous dissipativity of finite family of operators represented by matrix-columns the solvability of above written finite system of linear inequalities proves to be sufficient. To check solvability, one can use algorithms of linear programming [16].

**Remark 13.** The solution of the system (43)–(44) exists if there exists solution of the system of  $(n-d)$  linear inequalities complemented by inequalities (44) (where  $d$  is the number of those  $k$  for which  $r_k=0$ ; evidently  $0 \leq d \leq n-1$ ). To prove this, assume

$$a_{kk} = \max_{0 \leq l_k \leq r_k} a_{kk}^{(l_k)}; \quad a_{jk} = \max_{0 \leq l_k \leq r_k} |a_{jk}^{(l_k)}| \quad (j \neq k) \\ (k=1, \dots, n).$$

Consider the system

$$p_k a_{kk} + \sum_{j \neq k} p_j a_{jk} \leq 0 \quad (k=1, \dots, n). \quad (45)$$

Obviously, if the set  $\{p_k\}$  satisfies the system (44)–(45), then it satisfies the system (43)–(44) as well. Numbers  $k$  for which  $r_k=0$  are excluded. Therefore, in system (45) there are  $(n-d)$  inequalities.

**Remark 14.** For  $n=2$  theorem 10 provides necessary and sufficient conditions of simultaneous dissipativity. To demonstrate that, note that for operator  $M_k$  of the considered form the vector  $\psi_k$  (see the notation at the beginning of the subsection) is directed along one of the coordinate axes. Therefore (see the proof of sufficiency in theorem 9), if the family is simultaneously dissipative, then one can choose parallelogram as a ball of the corresponding norm, with vertices

on coordinate axes, i. e. the norm is of the form (40). In the case of arbitrary  $n$  the conditions of theorem 10 are already not necessary. To see this, let

$$A_1 = \begin{pmatrix} -1 & 0 & 0 \\ -1 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}; \quad A_2 = \begin{pmatrix} 0 & -1 & 0 \\ 0 & -1 & 0 \\ 0 & 1 & 0 \end{pmatrix}.$$

The system of linear inequalities

$$\begin{cases} -p_1 + p_2 + p_3 \leq 0; \\ p_1 - p_2 + p_3 \leq 0 \end{cases}$$

has no positive solutions. Nevertheless, simultaneous dissipativity exists, since each of the operators is dissipative in its norm and  $\varphi_1 = \varphi_2$  (see corollary 1).

## Conclusion

Let us resume. The *infinitesimal wrapping effect* (IWE) in the interval space for smooth autonomous system on positively invariant convex compact is studied. The local conditions of absence of IWE in terms of Jacobi matrices field of the system are obtained. The relation between the absence of IWE and simultaneous dissipativity of the Jacobi matrices is established, and some sufficient conditions of simultaneous dissipativity are obtained.

On the basis of the conducted analysis the reason of weak efficiency of interval stepwise methods is pointed out. The main reason is that to solve the problem of absence of IWE in the system and to construct corresponding interval space one needs analysis of simultaneous dissipativity of Jacobi matrices of system and constructing a constricting norm. The latter questions are rarely solved constructively. Besides, in sufficiently rich interval spaces (for example, in using standard intervals – rectangular parallelepipeds) IWE is almost always present. One should, however, remember that the notion of the wrapping effect in the work is treated sufficiently strongly. The final conclusion on the efficiency of stepwise interval methods can be drawn only after studying *asymptotic wrapping effect* (AWE). It should also be noted that there may be definitions of interval spaces, different from definition 1.

Some particular classes of systems without IWE and corresponding interval spaces are pointed out. These results can be used in solving by interval methods particular systems from the pointed out classes.

## References

- [1] MOORE R. E. *Interval analysis*. Prentice-Hall, N.-Y., 1966.
- [2] KALMYKOV S. A., SHOKIN YU. I., YULDASHEV Z. KH. *Methods of interval analysis*. Nauka, Novosibirsk, 1986.
- [3] CHERNOUSKO F. L. Optimal guaranteed estimations of uncertainty by means of ellipsoids. *Izv. AN SSSR. Tekhnich. kibernetika*, No. 5, 1980, 5–10.
- [4] KRACHT M., SCHOEDER G. Zur Intervallrechnung in linear Raumen. *Computing*, **11**, 1973, 73–79.

- [5] RATSCHKE H. Nichtnumerische aspekte der intervallararithmetik. *Interval Mathematics*, Springer — Verlag, Berlin-Heidelberg, 1975, 48–74.
- [6] KURATOVSKII K. *Topology*. Vol. 1, Mir, Moscow, 1966.
- [7] LOZINSKY S. M. Error estimation of numerical integration of ordinary differential equations. *Izv. vuzov. Ser. matem.*, No. 5, 1958, 52–90.
- [8] BYLOV B. F., VINOGRAD P. A., GROBMAN D. M., NEMYTSKII V. V. *Lyapunov Index Theory and its Applications to the Problems of Stability*. Nauka, Moscow, 1966.
- [9] VERBITSKII V. I., GORBAN A. N. *Simultaneously Dissipative Operators and Their Applications in Dynamical Systems*. Krasnoyarsk, 1987 (Prepr. /AS USSR, SB, Computing Center).
- [10] VERBITSKII V. I., GORBAN A. N. Thermodynamical restrictions and quasi-thermodynamicity conditions in chemical kinetics. *Mathematical Problems of Chemical Kinetics*, Ed. by K. I. Zamaraev and G. S. Yablonskii, Nauka, Novosibirsk, 1989, 42–83.
- [11] BURBAKI N. *Lie Groups and Algebras*. Mir, Moscow, 1976.
- [12] VOLPERT A. I., KHUDYAEV S. I. *Analysis in the Classes of Discontinuous Functions and Equations of Mathematical Physics*. Nauka, Moscow, 1975.
- [13] GORBAN A. N., BYKOV V. I., YABLONSKII G. S. *Essays on Chemical Relaxation*. Nauka, Novosibirsk, 1986.
- [14] GOLSHTEIN YE. G., YUDIN D. B. *New Tendencies in Linear Programming*. Sov. radio, Moscow, 1966.
- [15] VERBITSKII V. I., GORBAN A. N., UTJUBAEV G. SH., SHOKIN YU. I. Moore effect in interval spaces. *Dokl. AN SSSR*, **304**, No. 1, 1989, 17–21.
- [16] BYKOV V. I., VERBITSKII V. I., GORBAN A. N. On one estimation of solution of Cauchy problem with uncertainty in initial data and right part. *Izv. vuzov. Ser. matem.*, No. 12, 1991, 5–8.
- [17] VERBITSKII V. I., GORBAN A. N. Simultaneously dissipative operators and their applications. *Sib. matem. jurnal*, **33**, No. 1, 1992, 26–31.
- [18] VERBITSKII V. I., GORBAN A. N. Simultaneously dissipative operators and quasi-thermodynamicity of the chemical reactions systems. *Advances in Modelling and Simulation*, **26**, No. 1, 1991, 13–21.
- [19] VERBITSKII V. I., GORBAN A. N. On one approach to the analysis of stability of nonlinear systems and differential inclusions. *Advances in Modelling and Analysis, A*, **19**, No 4, 1994, 15–27.
- [20] VERBITSKII V. I., GORBAN A. N. Stability analysis and solution evaluation for nonlinear systems by “Jacobian fields” and Liapunov norms. *AMSE Transactions, Scientific Siberian, A, Vol. 4, Dynamics*, 1992, 104–133.